Using Thermal Inertia of Buildings with Phase Change Materials as a Flexible Energy Resource

Zahra Rahimpour

Centre for Future Energy Networks
School of Electrical and Information Engineering
Faculty of Engineering
The University of Sydney

Candidate's ORCID

A thesis submitted to fulfil requirements for the degree of
Doctor of Philosophy

September 2022

In memory of my kind father, Mozaffar,
To my amazing husband, Majid,
To my lovely son, Ryan,
And to my beautiful mother, Marzieh.

Statement of originality

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 50,000 words including appendices, bibliography, footnotes, tables and equations, and has fewer than 100 figures.

Zahra Rahimpour
September 2022

Acknowledgments

First and foremost, I would like to express my heartfelt deepest gratitude and acknowledgement to my supervisors Dr. Gregor Verbič and Dr. Archie C. Chapman. Thanks for being an exceptional mentor, generous in devoting your time to review my work and provide your insightful feedback. Having just become a mother during the first year of my PhD, I couldn't have completed my PhD research without your continuous support, encouragement, and patience. There are not enough words to express my gratitude to you.

My gratitude extends to the Australian Government for the Research Training Program Stipend Scholarship I received during my PhD study. I am also grateful for the Norman I Price Scholarship and the Postgraduate Research Support Scheme that I was granted from the University of Sydney during my study.

I want to thank the Higher Degree Research centre, Faculty of Engineering, the School of Electrical and Information Engineering at the University of Sydney for providing me with all support and guidance that I needed during my study. I thank Dr. Jianguo Zhu, Dr. Jeremy Qiu, Melissa Yeung, Rita, Ping and Linda. Also, I would like to thank Dr. Jin Ma for his support as a research training director, a friendly chat during lunchtime, and his valuable parental advice.

I also want to acknowledge the Sydney Informatics Hub and the use of the University of Sydney's high-performance computing cluster, Artemis. In particular, I want to thank Stephen Kolmann for his help in using Artemis and ongoing support when I incurred an issue.

Additionally, I thank the examiners of this thesis and my papers' reviewers. I appreciate your invaluable insights towards improving my work.

I gratefully acknowledge my co-authors Nicholas, Alice, and Donald for their contributions to the work contained in this thesis.

I would like to thank all Lab 329 members during my study at the University of Sydney: Jaysson, Daniel, Yiju, Bowen, Shabir, Mohammad, Mehdi and Samir, for providing an excellent research environment that inspired me in many ways. Additionally, I would like to thank the visiting students during my studies; Kaveh, Mitja, Brice, and Fynn. Having the opportunity to chat with you and learn from your research and experiences was a pleasure.

And last but not least, my deep and sincere gratitude goes out to my husband, Majid, for all your love, care, inspiration, encouragement and unconditional support. It would not have been possible for me to reach this point without your sacrifices and patience. I am deeply grateful to my wonderful son, Ryan.

You are a miracle in my life; having you is the greatest achievement of my life. You always remind me not to give up! Whenever something goes wrong in your plays, you always say: ”I try again!” I have learned this lesson from you, my little hero!

And I am always truly thankful for having a supportive family, my parents, Marzieh and Mozaffar, my sister, Zeinab, my brothers, Hossein and Saeed. My parents, in particular, deserve special thanks for their unconditional love, support, and sacrifices that helped me become who I am today. I am forever indebted to you.

Abstract

The massive uptake of *distributed energy resources* (DERs) and advancement in infrastructure technologies have attracted more attention on demand response. Using photovoltaic (PV) systems and battery storage systems enable end-users to actively manage their energy consumption by shifting or reducing flexible loads. One of the flexible loads that contribute primarily to buildings' energy usage and overall energy consumption worldwide is space heating and cooling. In recent years, the thermal inertia of a building has been demonstrated as a potential alternative to battery energy storage. Thermal inertia is the ability of a building's envelope to store or release thermal energy. In some countries such as Australia, the predominant type of building has low thermal inertia. Therefore, using thermal inertia as an alternative to battery storage systems becomes infeasible for these lightweight buildings. However, the integration of *phase change materials* (PCMs) into the building's envelope improves the building's thermal inertia significantly. PCM has very high latent heat compared to construction materials with high thermal inertia, such as brick. The type of PCM used in this thesis is plant-based, with a long lifespan (almost 80 years). The effectiveness of PCM in reducing heating and cooling energy consumption depends on various factors such as the geographical location of the building, PCM melting point, and the thickness of the PCM layer, to name a few. Moreover, using PCM passively without using heating, ventilation, and air conditioning (HVAC) in most applications puts the building's indoor temperature partially or entirely out of the desired comfort range of the householders.

Given this context, the present dissertation expands frontiers of PCM application in residential buildings, offering advances on PCM integration into *home energy management systems* (HEMS) that can be installed in smart meters with low computational power. Moreover, through many case studies in Australia, it investigates the viability of PCM to use as an alternative to the battery storage system to reduce electricity cost and increase self-consumption of PV systems. HEMS in this work consist of an HVAC system, rooftop PV system and a layer of PCM. The corresponding optimisation problem aims to find optimal scheduling of the HVAC system that minimises the electricity cost while maintaining the building's indoor temperature in a desired comfort range of householders. Solving this problem requires a technique that can deal with the nonlinear nonconvex characteristic of the PCM. Using conventional optimisation techniques such as *linear programming* (LP), *mixed-integer linear programming* (MILP), *mixed-integer nonlinear programming* (MINLP), *particle swarm optimisation* (PSO) and *genetic algorithm* (GA) are impractical. These methods are either incapable of handling nonlinearity or getting trapped in a local optimum. Using the state of the art *dynamic programming* (DP) method can handle the nonlinearity. However, DP suffers from the *curse of dimensionality*. The algorithm's runtime grows exponentially by adding more variables or increasing the time horizon of the problem. Moreover, model-based methods such as DP require an explicit dynamic model of an underlying system. In more detail, to solve an optimisation problem in a building with PCM, applying the DP method needs an exact thermal dynamic model of a building. This makes a plug-and-play implementation of the DP in smart meters difficult due to various building design and construction types. In other words, we need not only a computationally efficient method but also a method such

as cutting-edge *model-free reinforcement learning* (RL) methods. Against DP, model-free RL has no internal model of the system (thermal dynamic model of a building) and learns through direct interaction with its environment (which may be simulated offline).

Given this, this work begins with an introductory chapter, explaining current methods in HEMS and justifying a need for more efficient methods in HEMS with PCM. This is followed by a literature review that describes the application of PCM in residential buildings, HEMS, existing shortcomings in the current state of art methods used to solve the HEMS optimisation problem, and the standard methods and available software for modelling thermal behaviour of the buildings.

Next, we discuss the thermal modelling of buildings for energy management applications. In particular, we present mathematical modelling to capture the thermal performance of a typical residential building in Australia. The thermal model of the building will be used in HEMS formulation using model-based methods such as our proposed method of *multi-time scale approximate dynamic programming* (MADP) and as a learning environment for the model-free RL methods.

Then in the next chapter, to study the different factors that affect the capability of PCM in reducing or shifting the cooling and heating load, a wide range of scenarios is considered. The investigated factors include the impacts of geographical location, PCM melting point, duration of precooling and preheating, setting points of the HVAC system, thickness, and location of PCM.

Afterwards, in the following chapters, as the main contributions of the thesis, first, we propose a novel MADP methodology to reduce the computational burden of DP while maintaining the quality of the solution. Specifically, the method incorporates multi-timescale Markov decision processes (MDP) and a neural network function approximator of the system's dynamic model, coupled with an underlying state-space approximation. The results demonstrate that the proposed method performs well with a computational speed-up of up to 157,600 times compared to the direct application of DP. Second, we adopted a model-free actor-critic on-policy reinforcement learning method based on *deep deterministic policy gradient* (DDPG) that can learn policies in continuous action spaces without access to the full dynamics of the building. We demonstrate the competitive performance of DDPG by benchmarking it against the proposed MADP method that has access to the full thermal dynamics of the building. Next, we used our developed method of MADP as a computational engine to analyse the viability of phase change material to reduce electricity costs and observe how it impacts the self-consumption of a PV system. And finally, using the developed DDPG model, we implement simulations of different scenarios. We investigate the impact of householders preference (trade-off between electricity cost and householders comfortability), different electricity tariffs design, and the different sizes of the PV system on PCM performance as a storage system in HEMS on summer days in Sydney.

To this end, this thesis aims to give researchers and energy policymakers a practical vision in integrating phase change material into HEMS as an alternative to the battery storage system to be used in conjunction with a PV system.

Publications included in this thesis

I made a number of contributions in the papers listed below, where I was the lead author in three journals and three conference publications. The numbers in the tables represent percentage shares of each author's contribution. In each work, authors are arranged in ascending order based on their contributions.
Moreover, I have contributed as the co-author in one journal publication.

Journal Articles

[JA1] **Z. Rahimpour**, G. Verbič, and A. C. Chapman, "Can phase change materials in building insulation improve self-consumption of residential rooftop solar? An Australian case study", *Renewable Energy*, vol. 192, pp. 24-34, June, 2022.

Contribution made in this paper is incorporated in Chapter 7.

Contributors	Modelling	Analysing the results	Writing the paper
Z. Rahimpour	90%	85%	80%
Others	10%	15%	20%

[JA2] N. Morreau, G. Verbič, **Z. Rahimpour**, and A. C. Chapman "Optimal Scheduling of HVAC Systems in Buildings with Phase Change Materials using Deep Reinforcement Learning", submitted to *Engineering Applications of Artificial Intelligence.*

Contribution made in this paper is incorporated in Chapter 8.

Contributors	Modelling	Analysing the results	Writing the paper
Z. Rahimpour	90%	10%	10%
Others	10%	90%	90%

[JA3] **Z. Rahimpour**, G. Verbič, and A. C. Chapman, "Actor-Critic Learning for Optimal Building Energy Management with Phase Change Materials", *Electric Power Systems Research*, vol. 188, p. 106543, 2020.

Contribution made in this paper is incorporated in Chapter 6 and 8.

Contributors	Modelling	Analysing the results	Writing the paper
Z. Rahimpour	90%	85%	80%
Others	10%	15%	20%

[JA4] **Z. Rahimpour**, A. Faccani , D. Azuatalam, A. C. Chapman, and G. Verbič, "Using Thermal inertia of Buildings with Phase Change Material for Demand Response", *Energy Procedia*, vol. 121, pp. 102-109, July. 2017.

Contribution made in this paper is incorporated in Chapter 4.

Contributors	Modelling	Analysing the results	Writing the paper
Z. Rahimpour	80%	85%	80%
Others	20%	15%	20%

Conference Proceedings

[CP1] **Z. Rahimpour**, G. Verbič, and A. C. Chapman, "Computationally-Efficient Energy Management in Buildings with Phase Change Materials using Approximate Dynamic Programming", *2021 IEEE PowerTech*, virtual, June 2021.

Contribution made in this paper is incorporated in Chapters 5 and Chapter 7.

Contributors	Modelling	Analysing the results	Writing the paper
Z. Rahimpour	90%	85%	80%
Others	10%	15%	20%

[CP2] **Z. Rahimpour**, G. Verbič, and A. C. Chapman, "Energy Management of Buildings with Phase Change Materials Based on Dynamic Programming", *2019 IEEE PowerTech*, Milan, Italy, June 2019.

Contribution made in this paper is incorporated in Chapter 5.

Contributors	Modelling	Analysing the results	Writing the paper
Z. Rahimpour	90%	85%	80%
Others	10%	15%	20%

[CP3] **Z. Rahimpour**, G. Verbič, and A. C. Chapman, "Using Thermal Inertia of Buildings with Phase Change Material for Demand Response", *2018 Asia-Pacific Solar Research Conference (APSRC)*, Sydney, Australia, December 2018.

Contribution made in this paper is incorporated in Chapter 3.

Contributors	Modelling	Analysing the results	Writing the paper
J. Guerrero	90%	85%	80%
Others	10%	15%	20%

Table of contents

List of Figures

List of Tables

Nomenclature

Sets

$\mathcal{S}$ Set of state.

$\mathcal{X}$ Set of decision.

Variables

V_k^π Expected cost for following π from time step k.

s_k State at time step k.

x_k Decision at time step k.

$T_{\text{in},k}$ Indoor temperature at time step k.

p_k^+ Electricity demand (kW) at time step k.

p_k^- PV power generation (kW) at k.

P_k Electricity consumption at time step k.

$c_{\text{g},k}$ Electricity tariff at time step k.

R_k Reward at time step k.

C_k Cost at time step k.

$P_{\text{HVAC},k}$ Electricity used to run the HVAC system at time step k.

$P_{\text{load},k}$ Energy consumption of building load (excluding the HVAC consumption) at time step k.

Parameters

T_{out} Outdoor temperature.

T_{in} Indoor temperature.

T_{i} Surface temperature of slice i of the building element.

ΔT Thermal gradient.

R_{dw} Total thermal resistance of windows and door.

R_{out} External thermal resistance of the element.

R_{in} Internal thermal resistance of the element.

C_{e} Total thermal capacity of the element.

c Specific heat capacity of the building element.

k_e Thermal conductivity of the building element.

q Heat flux.

$\dot{q_{\mathrm{v}}}$ Heat gain inside the building element.

q_{i-} Heat flux between nodes $i-1$ and i.

q_{i+} Heat flux between nodes i and $i+1$.

Δx Width of each slice of the building element.

ΔQ Internal energy change per unit volume.

ρ Mass density of the material.

Q_{inf} Infilteration heat loss.

Q_{HVAC} HVAC power.

R_{e} Total thermal resistance of the element.

U Transmittance.

A_{e} Total element area.

r_{si} Inside surface thermal resistance of the element.

r_{so} Outside surface thermal resistance of the element.

d_{l} Thickness of the lth section in the element.

ρ_{l} Mass density of the lth section in the element.

λ_{l} Thermal conductivity of the lth section in the element

c_{l} Specific heat capacity of the lth section in the element.

β Thermal diffusivity.

α Accessibility factor of thermal resistance.

T_{e} Surface temperature of the element.

m_{a} Inside air mass of the building.

c_{a} Specific heat capacity of inside air of the building.

c_{pcm} Specific heat capacity of phase change material.

C_{PCM} Total heat capacity of phase change material.

T_{p} Melting point of phase change material.

K Total number of time steps.

k Time step.

λ A weighting factor applied to the electricity cost in the cost function.

T_s Desired room temperature.

H Enthalpy.

C Total heat capacity.

T Temperature.

π Policy.

π^* Optimal policy.

s^M Underlying mathematical model of the studied system.

$\bar{x}$ Mean value.

d_{ave} Average underlying electricity demand.

p^-_{ave} Average PV generation.

$d_{\mathrm{HVAC,ave}}$ Average HVAC consumption.

SC_{ave} Average PV self-consumption.

$\mathrm{SE}_{\bar{x}}$ Standard error of the mean value.

σ_x Standard deviation.

n Number of days in sample.

Abbreviations

ACH Air changes per hour.

ADP Approximate dynamic programming.

ANN Artificial neural network.

COP Coefficient of performance of the HVAC system.

DB Deadband control.

DDPG Deep deterministic policy gradient.

DERs Distributed energy resources.

DP Dynamic programming.

DR Demand response.

FiT Feed-in tariff.

HEMS Home energy management system.

HVAC Heating, ventilation and air conditioning.

MADP Multi time-scale approximate dynamic programming.

MDP Markov decision processe.

MSE Mean absolute error.

MT Melting temperature.

ODEs Ordinary differential equations.

PCM Phase change material.

PV Photovoltaic.

RL Reinforcement learning.

RMSE Root-mean-square error.

SC Self-consumption.

SOC State of charge.

ToU Time-of-use.

VI Value iteration.

Chapter 1

Introduction

Over the past decade, rising global warming concerns, finite fossil fuel reserves, and increasing energy demand has questioned the world's dependency on fossil fuels, encouraging the uptake of *distributed energy resources* (DERs). One of the worldwide fastest-growing DERs is rooftop solar photovoltaic (PV). By the end of 2020, the total cumulative installed capacity of PV systems reached well above 760 GW [5]. The major part of the PV installation ($37\,\%$) is for rooftop PV systems [6]. Australia is one of the countries with the highest solar energy potential. This puts Australia as a country with the highest uptake of rooftop solar PV, installed in nearly $30\,\%$ of homes [7].

However, integrating DERs into the existing electricity grid can jeopardise grid stability due to distributed form and variability nature of DERs. In contrast to a traditional energy system that increases generation to match demand, an alternative approach can manage demand in response to the variability of DER generation. This is known as *demand response* (DR). DR refers to a change in the energy consumption pattern of the electricity network end-users in response to electricity price signals or incentive payments from the network [8].

One of the major sectors that can be actively involved in DR is buildings. As per the international energy agency (IEA) report, the building sector consumes over one-third of stationary energy (gas and electricity), releasing nearly $28\,\%$ of global energy-related CO2 emissions in 2019 [9]. In Europe, space heating and cooling account for half of the EU's energy consumption, making it the biggest energy end-use sector ahead of both transport and electricity [10]. In the U.S., buildings utilise near $40\,\%$ of the energy and, more importantly, $76\,\%$ of the generated electricity [11]. The same trend can be seen in Australia. An important insight is the major energy consumption is in residential buildings, and the major energy drain is for *space heating and cooling*. Based on the Australian government's report in 2013, the HVAC system energy consumption accounts for approximately $40\,\%$ of total building energy usage in buildings in Australia [12].

Despite the high energy that is being consumed globally for heating and cooling of the buildings, the fact that this load is noncritical and can be shifted makes it one of the main DR resources that could provide significant flexibility to homeowners. To manage this flexibility in buildings with

rooftop PV system integration, *behind the meter energy storage* [1] has been introduced over recent years. Electrochemical battery energy storage systems (BESS) are commonly 'behind the meter'. A major benefit of the energy storage systems is to mitigate the variable nature of PV generation and to shift energy demand away from peak hours. However, access to efficient and cheap energy storage is still under investigation. One of the growing technologies of the storage system in Australia based on the ARENA report is batteries with Lithium-ion technology [13]. High-energy density, long term deployment, and low cost are the main reasons for the growth of Lithium-ion batteries in the electricity market. This battery, with minimum possible degradation each year ($1.5\,\%$), lasts 15 years. However, lithium-ion battery recycling is a barrier to consider these batteries as a sustainable solution. In more detail, the recycling process of these batteries are associated with safety issues of collection, transport, second-life re-manufacturing and materials recovery of the end-of-life batteries [14].

Against this background, an alternative energy resource available at almost no cost to all householders is the *thermal inertia* of the building. Based on the building's design and its construction materials, the building's envelope[2] can store the thermal energy that is known as thermal inertia of building. However, in countries such as Australia, a large number of residential buildings are lightweight buildings. These buildings dominate the building stock in Australia for two main reasons: (i) they are much cheaper and faster to construct than brick buildings; and (ii) they have a much better thermal performance in climates with insignificant diurnal temperature variation (i.e. tropical and sub-tropical climates in northern NSW and Queensland). Lightweight buildings also have a much lower carbon footprint compared to brick buildings. Their drawback is that they are much more susceptible to variations in external temperatures, which is particularly an issue in climates with high diurnal temperature variations, e.g. Victoria.

A new technology to improve the thermal inertia in the lightweight building is the integration of *phase change material* (PCM) into the building's envelope. PCM has a considerable amount of latent heat compared to high thermal inertia materials like brick. The latent heat of the type of PCM used in this work is almost 40 times of the brick with the same mass[3]. Storing or releasing this high latent heat during the phase-change (from solid to liquid or vice versa) could potentially provide the building with sufficient thermal energy storage to smooth indoor temperature fluctuations. In turn, this can result in fewer operating hours of the HVAC system while keeping the indoor temperature of the buildings in a comfort range for the occupants. This material is available in different forms. One common type that is easily integrated into new or retrofitted buildings is paraffin-based and available in a mat or encapsulated blocks. Compared to a battery energy storage system, this material has a long lifetime of 80 years and is much better for the environment as it is a plant-based material[3]. Using PCM as a storage system to implement DR in buildings is as follow; On a summer's day, an HVAC system can operate during shoulder and off-peak price periods to precool the building to a desired setpoint temperature, at or below the freezing point of the PCM. Then during peak pricing periods,

[1] Behind the meter refers to installed solar PV system, battery storage or any energy resources that are in home-users side to reduce or shift their electricity consumption.

[2] Building's envelope consist of walls, roof, floor and fenestration.

[3] https://phasechange.com

the HVAC system is turned off, and the indoor temperature of the building rises to a point where the PCM begins to melt. While the PCM melts, it absorbs heat from the building's interior, thereby maintaining a near-constant indoor temperature. To best exploit the energy storage potential of PCM, the aforementioned precooling (or preheating on a winter's day) by the HVAC system can be scheduled in a way to minimise the electricity cost while maintaining the indoor temperature of the building within the comfort range of the householders. In other words, this needs to be cast as an optimisation problem to automate the HVAC system's control in this setting.

To facilitate the active role of buildings in DR response and monitor and automation control of behind the meter resources, *home energy management systems* (HEMS) have drawn lots of attention in recent years. HEMS is a technology platform that consists of hardware and software that allows energy users to monitor, control, and manage electricity usage patterns during peak periods in response to electricity price signals. In practice, the developed algorithm will be embedded on a single-board computer (such as Raspberry Pi), which can be embedded in a smart meter or can be a stand-alone device. Despite technological advancement, still there is a limitation on the computational power and memory of these devices. Therefore, an optimisation algorithm needs to be computationally efficient while delivering an acceptable solution quality.

Within this context, the main focus of this work is a management system or software platform of the HEMS. More specifically, the main contribution of this thesis is to develop a computationally efficient optimisation method for scheduling a controllable device like an HVAC system in buildings with PCM and a rooftop PV system. The objective is to minimise the electricity cost and improve the *self-consumption* of the PV system while keeping the indoor temperature in the comfort range.

1.1 Research questions

For solving an optimisation problem in HEMS, different methods are applied in the literature [15–18]. Methods such as *linear programming (*LP) and *mixed-integer linear programming (*MILP) are widely used to solve HEMS problems mainly because of the simplicity associated with off-the-shelf solvers, such as CPLEX, Gurobi and MOSEK. The main drawback of these methods is that they optimise linear objective functions subject to linear constraints. However, one important feature of PCM makes these methods powerless for solving an optimisation problem in HEMS with PCM. PCM has a nonlinear nonconvex behaviour that converts the corresponding optimisation problem into a nonlinear nonconvex optimisation problem (explained in more detail in Chapters 2 and 3). Therefore, the desired feature of an optimisation method is to deal with the nonlinear characteristics of PCM. Solvers like Knitro can handle nonconvex nonlinear problems. However, the algorithm requires a good initial solution estimate, and usually, Knitro returns a locally optimal solution[4]. Applying MINLP formulation can capture the nonlinearity of the PCM through curve-fitting a nonconvex and nonlinear function to the data points of the PCM. However, this makes the problem computationally intractable while the existence of the solution is not guaranteed. Other methods that are extensively used in literature to solve HEMS

[4]https://documentation.aimms.com/platform/solvers/knitro.html

problem are heuristic methods like *particle swarm optimisation (*PSO) and *genetic algorithm (*GA) [17]. These algorithms search semi-randomly within a large population search space until they converge near a solution. In these types of methods, there is a risk that the solution ends up in a local optimum instead of the global optimum, which means the quality of the solution is uncertain.

The method to handle the nonlinearity of PCM is *dynamic programming* (DP). DP is the method to deal with problems that have a sequential structure. In the DP formulation, the problem is modelled as a *Markov decision process* (MDP) and solved by computing the expected future cost of following an optimal policy (in this problem specific, on/off combination of HVAC system) using backward induction. This algorithm is called *value iteration* of the system state variables (such as indoor temperature), and then an optimal policy can be extracted by selecting the state with the minimum value function using the Bellman optimality condition[5] [19]. DP is explained in more detail in chapters 2 and 5. However, in DP, computational time grows exponentially with the increase of a time horizon, the number of state variables or the number of controllable devices, the so-called *curse of dimensionality* [20, 21].

Given the limitations of existing approaches, we develop *computationally efficient multi-timescale approximate dynamic programming* (MADP) that performs well with a significant computational speed-up (up to 157,600 times over a 24 hours time horizon) compared to the direct application of DP. In more detail, we use a few abstraction techniques to reduce the algorithm's runtime while still the algorithm gives an acceptable solution quality. The first element of our method is a discretisation of the continuous state-space. As a second technique, we employ a multi-time scale MDP, in which decisions are made at different discrete timescales [22]. Specifically, rather than solving the original MDP as one monolithic problem, we solve several smaller MDPs that are connected successively together to form the original MDP. On top of this, to improve the computational performance further, we address the bottleneck of the state transition function, which is given by a thermal model of the building. Typically, in simulation, this is expressed as a system of *ordinary differential equations* (ODEs). We propose a *artificial neural network* (ANN) function approximation of the system of ODEs that maps the outdoor temperature, the indoor temperature of the building, and the HVAC heat flow, to the next time step's indoor temperature. This ANN can be trained offline, separate from the optimisation problem, and employed as a lookup table within our MADP method. Chapter 5 provides further details on the proposed method of MADP.

However, our proposed method (MADP) and all methods based on the principled DP method require a detailed thermal dynamic model of the building. In practice, the variety of building designs and construction types makes a plug-and-play implementation of the energy management algorithm challenging. Therefore, it is necessary to generalise the method to make it applicable to buildings with different design and construction types without affecting the algorithm's performance.

Moreover, another notable drawback of DP, in particular the VI algorithm, is that it cannot be directly applied to continuous domains such as control of the HVAC system.

These two drawbacks together lead us to consider *model-free reinforcement learning* (RL) methods,

[5] *"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision." [19].*

such as *deep deterministic policy gradient* (DDPG). The DDPG algorithm can deal directly with continuous variables. In comparison to DP, DDPG has no internal model of the system' state transitions and instead learns through direct interaction with its environment (which may be simulated offline). DDPG provides a significant computational speed-up over DP. In more detail, DDPG is a combination of deep learning for processing sensory input with an actor-critic for reinforcement learning [23]. In the actor-critic method, the policy structure, known as the actor, selects the action. The estimated value function is the critic that critiques the actor's action [24]. Neural network function approximators are used both for the actor and the critic in DDPG methods for online learning in large state and action spaces. DDPG for scheduling HVAC systems in buildings was initially proposed in [25]. They used DDPG to control the HVAC system's setpoint by minimising the overall cost of the HVAC system's electricity consumption while considering the thermal comfort of the home users. However, the quality of the solution that results from the DDPG algorithm was not benchmarked against well-proven algorithms like ADP that have full access to the model of the system.

In response to the existing gap in the literature, we developed the DDPG algorithm for solving an optimisation problem in HEMS with PCM, along with an extensive analysis of the DDPG algorithm's results against the outputs of the proposed method of MADP (Chapter 6).

In addition, the third literature gap is the co-optimisation of PCMs with other DERs such as rooftop PV, particularly in the context of maximising PV self-consumption. Therefore in Chapter 7 of this thesis, MADP is used as a computational engine to implement a techno-economic analysis of the viability of PCMs as an alternative to battery storage on a large number of residential buildings (210 dwellings in total). The benefits of using PCM are evaluated and quantified by calculating (i) electricity cost-saving (ii) an increase in PV self-consumption.

More specifically, this thesis aims to answer the following research questions:

[RQ1] What is the impact of different factors, including the geographical location, operational conditions of the HVAC system (duration of precooling and preheating, and setting points of the HVAC system), physical properties of PCM (melting point, thickness, location in the envelope), on PCM performance in particular electricity cost-saving and shift in the space heating and cooling demand, in the context of Australia?

[RQ2] What is the powerful and computationally efficient method to optimise the operation of controllable devices such as the HVAC system in HEMS consisting of PCM as a storage system and rooftop PV system as a distributed generation? In particular, the method needs to handle the nonlinearity of the corresponding optimisation method, deliver an acceptable solution quality and be executable on a smart meter with limited computational power and memory.

[RQ3] The bottleneck in the model-based method of DP to solve the optimisation problem in HEMS with PCM is to tackle the computational burden enforced by the thermal model of the building. What is the simplest model that can be used to be represented by low-order of ODEs (less computational cost) and at the same time serve as an acceptable representative model of a typical lightweight building in Australia? How can we validate the performance of the developed model?

[RQ4] How can we generalise the optimisation method in HEMS with PCM to be used as an embedded algorithm installed in the existing smart meters in buildings regardless of the buildings' design and construction types?

[RQ5] How far can we trust in model-free reinforcement methods compared to model-based methods such as DP?

[RQ6] To what extent can we use PCMs as an alternative to battery storage systems to reduce electricity costs and maximise PV self-consumption?

[RQ7] How can different preferences of householders (trade-off between electricity cost and householders comfortability), different electricity tariffs design, and the different scale of the PV system affect the HEMS performance in managing the cooling demand on hot summer days in Sydney?

1.2 Contributions and thesis outline

This thesis has three main contributions:

- A novel, powerful and computationally efficient MADP method for HEMS with PCM (Chapter 5). The proposed method can handle the nonlinearity of the corresponding optimisation problem.
- *Model-free reinforcement learning* (RL) methods, based on *deep deterministic policy gradient* (DDPG), that brings a practical implementation of energy management in smart buildings with PCM integration several steps closer to reality (Chapter 6).
- A large scale techno-economic analysis of the viability of PCMs as an alternative to battery storage to increase PV self-consumption and reduce electricity costs in HEMS with a layer of PCM as a storage system, a rooftop PV system and controllable device of the HVAC system (Chapter 7).

The outline of the thesis associated with the main contribution of each Chapter and the connection between chapters is illustrated in Figure 1.1. The chapters address the research questions that are previously raised in this Chapter.

The thesis is divided into eight more chapters: Chapter 2 provides a literature review on diverse topics, including the use of building thermal inertia for demand response, PCM application in residential buildings for demand response, HEMS in buildings with PCM, RL techniques to solve the corresponding optimisation problem, the available software for simulating thermal behaviour of the building, and the methods to represent the thermal dynamics of a building with the mathematical expressions. Moreover, at the end of Chapter 2, the thesis's contributions are discussed in more detail against the existing literature. In Chapter 3, a thermal model of a typical building in Australia, is developed in MATLAB and benchmarked against EnergyPlus software. Chapter 4 demonstrates the effectiveness of PCM as a DR resource in a building with a deadband HVAC control. A wide range of scenarios are considered

to investigate the impacts of geographical location, PCM melting point, duration of precooling and preheating, set points of the HVAC system, thickness and location of PCM, on the capability of the PCM in reducing or shifting the cooling and heating load. All simulations are performed using the EnergyPlus platform, examining three different types of residential buildings in five Australian cities: Sydney, Brisbane, Melbourne, Hobart and Perth. Chapter 5 discusses details of the MADP method. The method is implemented on a yearly case study in Sydney, using three defined metrics, the efficacy of the proposed method is evaluated. Chapter 6 provides details of the DDPG model. The simulation is run for 294 episodes that represent different weather conditions. And four defined metrics are used to benchmark DDPG performance against a model-based developed method of MADP. Chapter 7, using the MADP method, a large-scale techno-economic case study consisting of 210 dwellings in Australia across five cities of Sydney, Brisbane, Melbourne, Adelaide, and Perth is conducted to assess PCM integration's impact into solar rooftop PV systems. In particular, it addresses the electricity cost-savings and increases in PV self-consumption in the presence of PCM as a storage system. The analysis comprehensively covers different climatic conditions and the variability of end-user electricity demand and PV generation. Chapter 8, using the developed DDPG model, the usefulness of PCM is demonstrated for cooling the simulated building over a summer month in Sydney. The implementation includes 8 different scenarios, considering various weighting factors for trade-off between electricity cost and householders comfortability, different electricity tariffs design, and the different sizes of the PV system. And finally, Chapter 9 summarises the previous chapters' key findings and proposes insights on future work.

1.3 A potential demand response application for buildings with PCMs

In this thesis, PCM is used as a flexible energy resource rather than it is explicitly applied for demand response applications. However, in this section, we briefly explain how buildings with PCMs as a source of flexibility can be used for demand response.

With the operating envelopes (OE) approach, HEMS with PCMs and rooftop PV systems can be used for demand response. The distribution system operators (DSO) can compute PV export limits using day-ahead prediction of customer demand and PV generation. This export limits or imposed restriction on the DER is called an operating envelope. In order to prevent voltage stress adversely affecting the operation of the network, DER must stay within its operating limits. HEMS with PCMs use the OE and price signals computed by the system in order to optimises the scheduling of the HVAC system. The objective is to minimise the electricity cost, while maximising the self-consumption of the PV system, maintaining users' comfort thermal conditions and limiting PV generation to the operating envelopes submitted by the DSO.

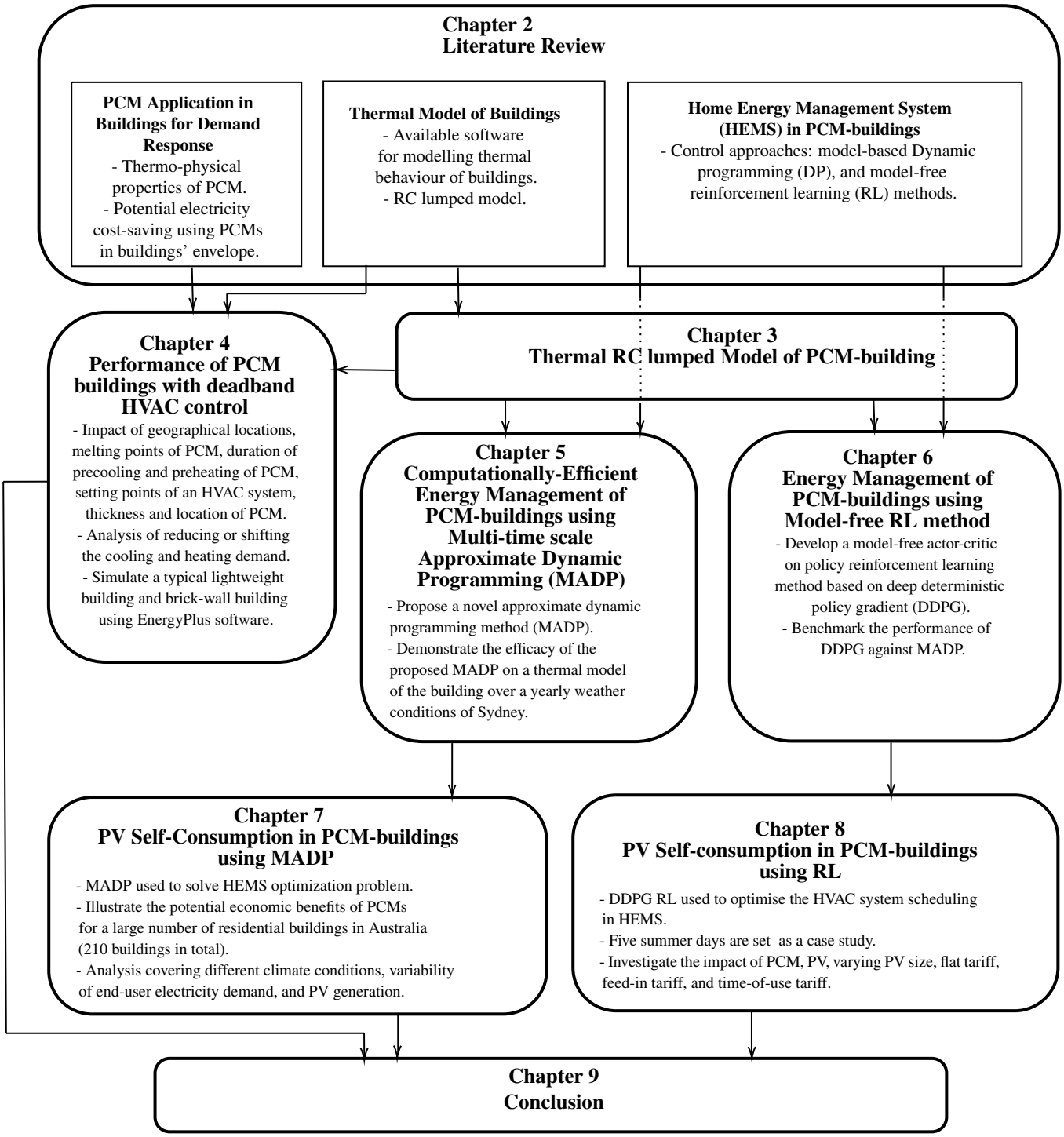


Figure 1.1: Thesis' contributions and connection between technical chapters in this thesis.

Chapter 2

Literature Review

As discussed in Chapter 1, using the thermal inertia of buildings for demand response (DR) has gained considerable attention over recent years. However, in some countries such as Australia, a common type of residential building has low thermal inertia that can be improved using promising phase change materials (PCMs). On the other hand, with the increasing uptake of distributed energy resources (DERs) and technological advancements in communication and metering devices, a home energy management system (HEMS) has become a platform that can be used to modify the residential energy consumption pattern to achieve more efficient use of DERs. In light of the above, current work aims to take a step towards the real-world deployment of PCM integration into HEMS.

As such, the first part of this section reviews the literature on potential energy-saving and shift in space heating and cooling demand that can be achieved using the thermal inertia of building as an energy storage system.

The second part starts with a description of the thermo-physical properties of PCM, different types of PCM, and a short description of the type of PCM used in this thesis. It follows by reviewing the literature on the potential energy saving obtained by integrating PCM into the building's envelope. The discussion continues by reviewing the literature on different approaches to investigate PCM's desired optimal performance. Finally, the section is completed with a review on the co-optimisation of PCMs with other DERs such as solar rooftop PV generation.

In the third part, first, the general concept of HEMS is discussed along with techniques that are used extensively in the literature to solve an optimisation problem in HEMS. Afterwards, we introduce HEMS in buildings with PCM. In particular, we detail what distinguishes optimisation problem in the HEMS with PCM from the other HEMS' optimisation problems. And finally, we introduce reinforcement learning (RL) methods to solve the nonlinear non-convex optimisation problem in the HEMS with PCM. The discussion expands to two different groups of RL, namely (i) model-based and (ii) model-free RL.

In the fourth part, the chapter continues by introducing different software tools that can be used to simulate the thermal performance of a building. This is followed by reviewing different methods that are used in the literature to model the thermal dynamics of a building.

Finally, the chapter concludes by discussing the literature gaps that paves the way for the remainder of the thesis to be elaborated.

2.1 Using thermal inertia of buildings for demand response

Thermal inertia is the inherent property of the building's envelope (walls, roof, floor, and fenestration). It is defined as the envelope's capacity to absorb, store or release thermal energy. The amount of the building thermal inertia determines how much it can delay or reduce the impact of the external weather conditions on the internal temperature of the building. Buildings with high thermal inertia can shift the energy consumption from on-peak hours to off-peak hours and save electricity costs for operating the HVAC system. Moreover, heat flow has low fluctuation in building with high thermal inertia, resulting in low heating and cooling device consumption. Therefore, it plays an essential role in the comfort level of the householders and their energy consumption.

Many studies show the impact of a building's thermal inertia on reducing the energy demand [2, 26–33]. Authors in [26] analysed the effect of thermal inertia on total energy consumption. They simulated two different fibro house and brick veneer house in Sydney. In the fibro house, they used fibro as an external layer of the building and raised timber for the floor. In a brick veneer house, they replaced the fibro with brick veneer and used an insulated concrete slab on the ground. They found that using high thermal inertia construction such as brick veneer instead of construction with low thermal inertia such as fibro can reduce the energy consumption in winter by $35\,\%$ and in summer by $36\,\%$. As discussed, the brick veneer house with the concrete floor has high thermal inertia that acts as a storage system. During summer, the brick veneer house envelope absorbs heat during the day and releases it during the night. Moreover, in winter, the envelope stores heat during the day and discharge it to the house interior during nighttime. They verified their simulation results with field measurements of the actual buildings energy consumption.

Moreover, authors in [31] used measurements on two identical detached single-family buildings to demonstrate the high thermal inertia performance on a summer month in central-western Poland. The only difference between the two buildings was the construction of the external and internal walls. One was lightweight against the other building with masonry construction that was built from cellular concrete. They considered a deviation from $28\,^{\circ}\mathrm{C}$ as a discomfort measure for the house users. In the lightweight house, the discomfort was measured 18.6 days compared to 8 hours in heavy structure building. The summer month was hot with 14 days temperature exceeding $30\,^{\circ}\mathrm{C}$, and for 8 days, the temperature reached above $32\,^{\circ}\mathrm{C}$ during the day. The average nighttime temperature was above $18\,^{\circ}\mathrm{C}$. Results show that using high thermal mass yields $67\,\%$ and $75\,\%$ reduction in energy consumption in the summer month when temperature thresholds of $25\,^{\circ}\mathrm{C}$, and $26\,^{\circ}\mathrm{C}$ are considered, respectively.

Similarly, the authors in [2] also demonstrated the impact of the high thermal inertia construction by comparing different building construction types. In more detail, they compared the thermal performance of four types of wall construction: cavity brick, brick veneer, reverse brick veneer, and lightweight. They used simple building modules with the same floor and roof structure for all four case

studies. The case studies are tested for the climate of Newcastle in Australia. They used the AccuRtae energy rating tool for numerical calculations. AccuRate assigns a star rating to a building based on seasonal and annual heating and cooling required to meet predetermined thermal conditions for each season and room type. They conclude that thermal inertia has a significant impact on annual space heating and cooling consumption. Building with reverse brick veneer wall was the most energy-efficient design for this design in this climate. In contrast, the lightweight wall module had the highest energy consumption, $39\,\%$ more than the reverse brick veneer wall module. Energy consumption in cavity and brick veneer was almost $21\,\%$ more than reverse brick veneer module.

Moreover, work in [2] highlights the impact of the building envelope composition. Placing the building material with high thermal inertia close to the interior of the building (like reverse brick veneer) results in lower energy consumption than putting the envelope layer with the highest thermal inertia towards the building exterior. In addition to [2], some works such as [33] also demonstrate that placing the layer of the wall with the highest thermal inertia in close vicinity to the building interior that is insulated from outside conditions results in higher energy saving rather than placing it as an external layer of the building's envelope. In more detail, in [33], the authors investigated the performance of a solid concrete block and hollow concrete blocks in two cases of putting high thermal inertia material (cement plaster-moulded polystyrene) close to the interior or alternatively placing it near to the exterior of the building.

However, having high thermal inertia in the building is not always beneficial for reducing energy consumption and improving householders' thermal conditions [28, 32]. Authors in [32] investigated the impact of the thermal dynamics of the building envelope elements (roof, walls, and floor) on the heating and cooling demand for a retrofitted office building case study. They performed a parametric simulation and calibrated the model with the data collected during the building performance monitoring process. To investigate the impact of the different refurbishment solutions in the simulations, they considered the elements of the building such as external walls, rooftop, and internal floor with three construction types of light, medium-heavy, and heavy construction. These three types of constructions are different in the amount of the building effective thermal capacitance. They found that using medium-heavy construction in the climate zone of the Southern European and Mediterranean areas, results in the highest reduction in cooling demand compared to heavy and light constructions. In more detail, using medium-heavy construction reduces the cooling energy consumption by $27\,\%$, compared to $18\,\%$, and $12\,\%$, utilising heavy construction, and light construction, respectively. This highlights the importance of capturing the dynamic thermal performance of the building when optimising the energy consumption in buildings. Another work is [28] that the authors simulated a one-dimensional model consisting of an external lining, internal air volume and internal lining. The interior lining thickness is considered more than the external lining, which means the high thermal inertia is close to the building's interior. The authors conducted a few case study simulations. They demonstrated that adding thermal mass to heavy construction can change the energy consumption pattern. However, in the case of using intermittent heating (using the heating system during the weekend), having high thermal inertia is a disadvantage in terms of energy saving in cold climates. If the indoor temperature swing is tight (¡$0.5\,^{\circ}\mathrm{C}$), having

high thermal inertia cannot reduce the peak demand. By contrast, if the indoor temperature is allowed to decrease, particularly during peak hours, the stored energy in high thermal inertia envelopes can be released, thereby reducing energy consumption. In more detail, they concluded that the impact of a high thermal inertia envelope depends on several factors, such as the comfort range of the indoor temperature, the frequency of heating use, the electricity tariff, free heating from the sun and the climate. They also noted that heavy construction can prolong the building's precooling and preheating, causing discomfort to home owners. This agrees with the authors finding in [29]. They claimed that high thermal mass can be effective for a hot climate with a large diurnal temperature. However, high thermal inertia increases energy consumption in cold climates.

Besides the effects of the envelope construction type and position of the highest thermal mass in the envelope, the study in [30] demonstrated the benefits of precooling the envelope of a building. The study was conducted in a medium-sized Santa Rosa, California, governmental office building. In particular, they reported that by precooling the building during off-peak hours, the electricity demand during afternoon peak times could be reduced significantly by using the storage capacity of the thermal mass in a commercial building. They used a building control system for recording the HVAC system performance. In the building, they normally used to set the HVAC system to the constant point of $22\,^{\circ}\mathrm{C}$ throughout the occupied time and shut down the HVAC system after 5pm (baseline scenario). They applied two precooling strategies: (i) precool the building from 5am to 2pm, with the HVAC system setpoint of $21\,^{\circ}\mathrm{C}$, and from 2pm to 5pm, with the HVAC system set to $25.5\,^{\circ}\mathrm{C}$, the HVAC system was shut down after 5pm, (ii) in the second strategy, called extended precooling, the HVAC was turned on at midnight (12am), set to $20\,^{\circ}\mathrm{C}$, and ran until 5am. From 5am to 2pm, they changed the setpoint to $21\,^{\circ}\mathrm{C}$, and from 2pm to 5pm, the setpoint of the HVAC system was set to $25.5\,^{\circ}\mathrm{C}$. They modified the HVAC system's temperature setting to precool the building to the lowest point of the comfort range till 2pm. After 2pm, the temperature was allowed to rise to the highest point within the defined comfort zone ($25.5\,^{\circ}\mathrm{C}$). They defined three types of cold days ($22\,^{\circ}\mathrm{C}$ - $25.5\,^{\circ}\mathrm{C}$), warm days ($25.5\,^{\circ}\mathrm{C}$ - $29\,^{\circ}\mathrm{C}$) and hot days ($29\,^{\circ}\mathrm{C}$ - $32\,^{\circ}\mathrm{C}$) based on the peak outside air temperature. They compared two strategies for the warm days and concluded that there is not a significant difference in the power consumption between the two scenarios. However, comparing two scenarios to baseline scenarios demonstrates that using the precooling strategy could reduce the HVAC system consumption by $80\,\%$ - $100\,\%$ during the peak hours of 2pm-5pm. Their work showed the importance of precooling to reduce the cooling demand during peak hours. However, their study lacks sufficient data compare two precooling strategies, particularly for hot days. In other words, the impact of precooling duration and setpoint of the HVAC system on the energy consumption of the building was not thoroughly investigated.

2.2 Integration of PCM into buildings envelope for demand response

This section discussion is broken into the topics:

- Phase change materials (PCM), including thermo-physical characteristics, various types of PCM, and type of PCM used in this thesis (Section 2.2.1).
- Potential energy saving in PCM-integrated buildings (Section 2.2.2).
- Optimal thermal performance of PCM (Section 2.2.3).
- Co-optimisation of PCMs with distributed energy resources (Section 2.2.4).

2.2.1 Thermo-physical characteristics of PCM

From a thermodynamic point of view, the PCM absorbs (melting) or releases energy as heat (freezing) due to enthalpy changes when the temperature varies within a certain range. This heat is called *latent heat*; latent heat is the main useful property of PCM. The high amount of latent heat in PCM makes this material supreme thermal storage.

Applications of PCM in residential and commercial buildings began several decades ago. In 1949, Telkes investigated the use of PCM in a solar heater for space heating. Since then, the number of studies on applying PCM in buildings for thermal comfort and energy saving has increased extensively [34].

Based on recent studies, for making the selection between different types of PCM, a few criteria are identified: (i) PCM melting temperature range (ii) specific latent heat of the material, (iii) chemical stability, (iv) supercooling, and (v) thermal conductivity. PCM has a wide range of applications such as industrial heat recovery, healthcare, buildings, and aerospace, to name a few. The initial decision in selecting the PCM type is choosing the PCM melting range, which is based on the PCM application. In this work, we are interested in the PCM type that can be integrated into a building's envelope. For this type of application, PCM melting is in the range of $15\,^{\circ}\mathrm{C}$-$32\,^{\circ}\mathrm{C}$, which is close to the human comfort range [35]. Another important criterion is latent heat. Needless to say, a PCM with higher latent heat is more desirable in thermal storage applications. However, this feature is beneficial only when the other three characteristics of the PCM have values within an acceptable range. To begin, consider the chemical stability of the PCM. Since this material undergoes numerous thermal cycles, the chemical stability of the material prevents reduction of the latent heat, change in melting point range and risk of PCM leakage from its capsule [36]. It is essential to understand the chemical stability of the PCM as it defines its available duration. Another essential characteristic is the supercooling ability of the PCM. Suppose the liquid PCM is rapidly cooled down. In that case, the temperature of the PCM can decrease faster than the solidification process, resulting in liquid PCM with a temperature lower than its freezing point. This causes disruptions to the cycling process of the PCM. Therefore it is crucial to choose a PCM with low supercooling potential. The last criterion is the thermal conductivity of the PCM.

The thermal conductivity is critical as it allows complete penetration of the heatwave in and out of the material and increases the thermal efficiency of the material. In addition to physical and chemical criteria, choosing PCM that has a low environmental impact, is cost-effective, and commercially available are other important factors to be considered in selecting the type of PCM [35]

A common type of PCM used in buildings' envelopes are paraffin types, fatty acids and polyethylene glycol, all classed as organic types of PCM. This PCM type has high latent heat and thermal conductivity, good chemical stability, and limited supercooling [3].

Another important aspect of PCM application is how PCM can be integrated into buildings construction. PCM can be incorporated into a building's envelope in four ways [37, 38]:

1. Direct incorporation: in this method, PCM in liquid or powder state is added to construction material such as gypsum, concrete or plaster. This method is simple; however, PCM leakage and the incompatibility of the final product with other construction materials are potential issues of using this method.

2. Immersion: the building material such as gypsum or concrete are immersed in melted PCM. Therefore, PCM is absorbed into the internal pores of these materials. Same as direct incorporation, this method has two issues of PCM leakage and construction incompatibility.

3. Encapsulation: in this method, PCM is encapsulated in a container. Therefore the leakage problem of PCM can be avoided. Moreover, encapsulation of PCM enhances the compatibility of the PCM with the building structure. Encapsulation can be done using tubes, spheres, or panels as containers, which is called microencapsulation. Nowadays, with advancements in technology, PCM particles can be encapsulated in a thin sealed unique polymeric film. Using this method, the leakage issue can be avoided. Moreover, integrating the PCM panel has less effect on other construction types and designs. However, the main drawback of using a thin sealed panel is poor thermal conductivity and solidification of materials at the edges of the panel [38].

4. Shape-stabilised method: in this method, PCM is dispersed in another high dense supporting material to become a stable composite material. This method provides better thermal conductivity, large specific heat and the ability to keep the shape in phase change cycles of PCM. This method is the most reliable method as it makes the PCM service life longer. In more detail, PCM melting/solidification cycles are repeated in high performance without critical degradation. However, this method is expensive to implement.

The type of PCM used in this work is BioPCM (TM) material. Using the encapsulation method, PCM is encapsulated in small discrete blocks joined together in the form of a mat. The mat shape of BioPCM provides easy installation of this material, particularly for retrofitted buildings as the PCM mat can be spread on the ceiling's plasterboard[1].

[1] http://phasechange.com

2.2.2 Potential energy saving in PCM-integrated buildings

The effectiveness of PCM to reduce peak heating and cooling demand and improve indoor thermal comfort is highlighted in most of the recent literature [34, 39–43]. In [39], the researchers built two identical test rooms. One room was constructed of gypsum board, and the second one with PCM-impregnated plasterboard. Both experimental and numerical results indicated that the indoor temperature fluctuation of the test room with PCM was $4\,^{\circ}\mathrm{C}$ less than the room without PCM. As expected, they demonstrated that lower indoor temperature fluctuation results in energy saving on a typical summer day. In particular, using PCM reduced the energy consumption of the air-conditioning with a setpoint of $22\,^{\circ}\mathrm{C}$, by $35\,\%$. In [34] the authors conducted experiments on several house size cubicles to explore the impact of including PCM in concrete, conventional brick and alveolar brick. Also, a Trombe wall was added to the concrete cubicles to investigate the free cooling impact, and in the brick cubicles, a heat pump was fitted to simulate an HVAC system. By using PCM, the temperature oscillation of the concrete blocks was reduced by about $4\,^{\circ}\mathrm{C}$ and also the peak temperature was shifted about two hours to later hours of the day. In the brick cubicles, including PCM, led to $17\,\%$ saving in cooling energy consumption on a summer day. On a winter day, PCM acted like insulation and kept the inside temperature of the block warmer than the brick cubicle without PCM.

Another work that demonstrated the impact of PCM as an energy storage system for electricity DR is [40]. They used two identically designed offices built at Tamaki Campus of the University of Auckland, New Zealand. In the office with PCM, instead of ordinary gypsum board, PCM impregnated gypsum board is used for the interior layer of all walls and ceiling. The PCM used in this work is RT-20 with a melting range of $18\,^{\circ}\mathrm{C}$-$22\,^{\circ}\mathrm{C}$. Both of the offices are equipped with a heating system. The heating system operates to maintain the temperature in winter in the desired comfort range of $18\,^{\circ}\mathrm{C}$-$22\,^{\circ}\mathrm{C}$ that is matched with the melting range of PCM. They observed that over 12 days of winter, office with PCM consumed $31\,\%$ of the total energy consumed in non-PCM office to keep the indoor temperature in the same range. They witnessed that when the ambient temperature is higher during the day and lower during the night ($13\,^{\circ}\mathrm{C}$, and $21\,^{\circ}\mathrm{C}$), the energy-saving is higher. Wide ambient temperature variation allows exploiting the storage capacity of PCM efficiently as PCM undergoes both melting and solidifying cycles. Moreover, they could observe 150 minute load shifting on a typical winter day.

In addition to [40], another two studies also reported the peak load shifting benefits of PCM. The authors of [41] used a simulation model to analyse the ability of PCM to shift air conditioner load away from peak hours. The wall and ceiling of the simulated building were modelled by three layers of glass fibre insulation, PCM, and glass fibre insulation. Outdoor and indoor temperature variations was modelled with time-varying boundary conditions. The model predicted a $26\,\%$ reduction in peak cooling load by including PCM over spring and summer seasons. The results also showed a significant load shift, but the results were theoretical, and the outcomes were not verified with experimental results. In [42], the authors conducted experiments to demonstrate the peak load shifting that results from applying PCM. They built two identical test huts, one with gypsum board, and the second with PCM impregnated gypsum board that consisted of $26\,\%$-wt PCM. Tests were taken over a typical January

week in New Zealand, and the performance of the two huts, without any active cooling, was compared. They observed that PCM-hut's temperature at the daytime rose at a lower rate than hut without PCM. However, at night, the interior temperature of PCM-hut was higher than in the hut without PCM. The weekly-averaged indoor temperature of the PCM hut varied $5.1\,^{\circ}\mathrm{C}$, while the corresponding temperature variation in the hut without PCM was $10.6\,^{\circ}\mathrm{C}$. In winter, without any heating system, the PCM's temperature did not reach to transition range. Therefore there was a major difference between the performance of the two huts. Following this, they used an electrical heater that was scheduled to be turned on from 1:00 am to 7:00 am (the off-peak period). After turning off the heater in both test huts, the indoor temperature of the hut without PCM dramatically decreased. In contrast, the temperature profile in the PCM hut was in the acceptable range since the indoor temperature fell below the PCM freezing point, and the stored heat from the heater was released into the room's interior. This shows the capability of PCM in terms of peak load shifting, which will bring a notable reduction in the electricity cost.

Authors in [43] investigated the impact of PCM on reducing the heat stress risk during heatwaves in Melbourne. They have simulated a single-story house without air conditioning as a case study. The simulation results benchmarked against a refurbished house with the installation of commercially available Bio-PCM. Their simulation results revealed that using PCM in severe heatwaves can reduce the discomfort period by $65\,\%$.

PCM performance is complex, and it depends on several factors including the PCM properties, the thickness, the location of the PCM in the envelope, the thermal conductivity, the climate, the parameters of the building, and the active or passive application [38]. Numerous works exist that demonstrate PCM's energy-saving capacity while investigating how to make it perform at its best [44–56]. In [44], the authors studied the impact of PCM on the cooling demand of an office building in the whole summer season. The office was refurbished by adding PCM plaster to the building's envelope. The experiment was accompanied by simulation, and five Mediterranean climates of Ankara (Turkey), Athens (Greece), Naples (Italy), Marseille (France) and Seville (Spain) were taken into account. Integrating PCM did not return the same saving throughout the season. In Ankara, by including PCM (with a melting point of $29\,^{\circ}\mathrm{C}$), cooling demand was reduced by around $7.2\,\%$, while in Seville and Naples, it resulted in less than $3\,\%$ savings in cooling consumption. This shows that the performance of PCM depends on the weather condition. The other indicator that the authors used was comfort hours. They calculated the number of hours that the indoor temperature was in the defined comfort temperature range of $24\,^{\circ}\mathrm{C}$ to $28\,^{\circ}\mathrm{C}$ and utilised this as a comfort measure. They concluded that using PCM with a different melting point of $26\,^{\circ}\mathrm{C}$ and $29\,^{\circ}\mathrm{C}$ increased comfort hours from $11.2\,\%$ to $21.9\,\%\%$ in Athens and from $32.9\,\%$ to $51\,\%$ in Marseille. These studies show that energy saving is achievable by using PCM in buildings. More importantly, it shows that the performance of PCM is highly dependent on the weather conditions that PCM is used in. Another study that demonstrates the importance of weather conditions on PCM performance is [45]; the authors investigated the optimal melting point of PCM for five cities in Australia. They concluded that the optimal melting point of PCM for specific weather conditions is approximately equal to the average outdoor air dry-bulb temperature. This agrees with the finding of

the authors in [46].

In some literature [47, 48], authors used a PCM energy storage efficiency as a metric to evaluate the PCM performance. In [47], the PCM energy efficiency factor is defined as a ratio of decrease in discomfort hours to the amount of PCM used in the PCM-impregnated gypsum. They concluded that PCM energy storage efficiency is highly dependent on the outdoor temperature, and in geographical locations where temperature oscillations were far from the melting point of PCM, phase transitions occurred less frequently. Consequently, PCM efficiency is low. Similarly, authors in [49] demonstrated that using PCM is not beneficial in hot and humid climates with insufficient diurnal temperature variation as PCM cannot go through the full cycle of solidification and melting.

In addition to the climate impact of PCM performance, literature investigated the impact of natural ventilation on the thermal behaviour of PCM. In some studies like [48], authors demonstrated night ventilation impact on optimal performance of PCM. They explored the thermal comfort during summertime in single zone lightweight buildings considering six cities of Kazakhstan. PCM with different melting points of $26\,^{\circ}\mathrm{C}$, $28\,^{\circ}\mathrm{C}$, $30\,^{\circ}\mathrm{C}$, and $32\,^{\circ}\mathrm{C}$ were considered. PCM that gives the minimum total discomfort hours was selected as an optimal melting point of PCM. Based on the results, the optimal performance of PCM achieved in all cities using PCM with a melting point of $26\,^{\circ}\mathrm{C}$, in combination with night ventilation. For the type of PCM used in their study, the phase change of PCM with the melting of $26\,^{\circ}\mathrm{C}$ was between $24\,^{\circ}\mathrm{C}$ to $28\,^{\circ}\mathrm{C}$. Natural night ventilation is beneficial in this type of application since the night temperature is well below $24\,^{\circ}\mathrm{C}$ in the summer. During summer, the average outdoor temperature in all cities was below $24\,^{\circ}\mathrm{C}$; thus, the night ventilation makes the PCM solidify and ready to use during the summer day. Moreover, they used the PCM storage efficiency metric, defined as a ratio of the stored thermal energy to the PCM latent heat. They demonstrated that even using the optimal melting point of PCM for their simulations, the PCM storage efficiency is $39\,\%$. In addition to this work, authors in [50] also showed the importance of nigh ventilation in PCM-integrated buildings through simulating a typical PCM-building thermal behaviour for ten cities in Western China. Compared to buildings without PCM and night ventilation, using PCM with nigh ventilation reduces the discomfort hours by a minimum $11\,\%$ to a maximum $16\,\%$ depending on the season. Using night ventilation increases PCM effectiveness in reducing discomfort in hot seasons. This is because, during hot seasons, the temperatures during the night do not fall below or near the freezing point of the PCM. As a result, PCM cannot solidify at night. In contrast, using night ventilation in PCM buildings located in cities with severe cold weather conditions increases the discomfort hours in transition seasons compared to PCM buildings without night ventilation.

However, night ventilation, and the general passive application[2] of PCM, cannot always maintain the indoor temperature in the desired comfort range [49]. In particular, if the average indoor temperature is not within householders comfort range, passive application of PCM is not always effective in reducing the average indoor temperature, particularly on hot summer days. Where the PCM application is limited, the air conditioning system can be used to improve PCM performance [51].

[2]Passive application refers to the use of PCM where the phase change occurs without the aid of mechanical devices such as an HVAC system.

Another factor that impacts the performance of the PCM in reducing energy consumption is the orientation of the PCM-building. Authors in [53] demonstrated that the application of PCM in west-facing walls results in $2.9\,\%$ annual energy saving in cooling consumption compared to $2.2\,\%$, $1.9\,\%$, and $1.4\,\%$ for using PCM on east, south, and north-facing wall, respectively. The reason is that Hong Kong is located in the equatorial zone of the northern hemisphere. Therefore, the west-facing and the east-facing walls receive more solar radiation in summer. As a result, PCM absorbs the heat from solar radiation to solidify and make the interior of the building cool.

In addition to the factors previously mentioned, the literature also studies how the thickness of the PCM, the building's insulation, and the location of the PCM in the envelope affect its performance. In [52], the authors demonstrated the performance of PCM-enhanced buildings under Danish conditions. They witnessed that placing PCM in the interior side of the building results in more energy saving compared to placing the PCM in the middle or external side of the wall. They also investigated the impact of different thicknesses of PCM on energy saving. Based on their results, increasing PCM thickness from 5mm to 40mm results in a higher reduction in energy consumption; however, the reduction rate is much lower if the PCM thickness increases beyond 40mm. They reported integration of PCM in a typical one-story house and an apartment results in $7\,\%$ and $15\,\%$ annual energy saving, respectively. Moreover, they observed the impact of the studied building's insulation on the PCM performance. They concluded that using PCM in buildings with more insulation results in higher energy savings. In more detail, they increased the insulation of the building by changing the U-values ($\mathrm{W/m^2K}$) of the external walls (0.93 to 0.18), roof (0.43 to 0.12), floor (0.43 to 0.10), windows (3.35 to 1.5), and door (3.35 to 1.5). Increasing insulation coupled with PCM application increased energy saving from $18\,\%$ to $34\,\%$. This is against the author's findings in [53] that showed integration of PCM into the uninsulated brick wall in Tunisian Mediterranean climate results in a higher reduction in energy consumption ($12.21\,\%$) than the brick wall with PCM and expanded polystyrene insulation layer. Authors in [54] also demonstrated the impact of the PCM location and thickness on it's performance. In particular, they simulated the performance of a PCM-enhanced prefabricated building in five different climate regions in China. They found that the thickness of PCM that gives the highest energy saving depends on the climate. For areas with severe cold, cold, mild, hot summer and cold winter, hot summer and warm winter, the optimal thickness for PCM are 10mm, 10mm, 20mm, 30mm, and 30mm, respectively. Their results showed that placing PCM on the interior side of the envelope in all five different climate conditions results in more savings on annual energy consumption than placing it on the exterior side of the building's envelope. Moreover, they observed that energy-saving using PCM in the prefabricated building is low in the area with cold weather conditions, while it is high in the hot and mild regions. Using PCM in all five different climate conditions resulted in energy savings ranging from $17.7\,\%$ to $77\,\%$. In contrast, authors in [55] concluded the optimal location for the PCM layer is the exterior layer. The reason could be that they did not consider the whole year's energy saving like in [52, 54]. In more detail, they set up an experiment to study the optimal performance of PCM in three locations of the inside, middle, or outside in a brick wall. They ran the test for 60 days, the summer season. Moreover, they used different melting points of PCM for different locations of PCM

that made the judgment difficult. Another study that explored the impact of PCM thickness and its location on energy saving in the buildings is [56]. In more detail, the authors simulated a residential building thermal performance to evaluate the feasibility of PCM application in energy saving in the savanna climate zone considering eight cities. Their result demonstrated that PCM applications in eight selected cities could reduce the cooling consumption by a minimum $16.6\,\%$ to a maximum $68.6\,\%$. For all cities with an increase of the PCM thickness, the energy-saving was increased. However, they witnessed that the energy consumption reduction rate was decreased with the increase of thickness. Their other interesting finding was that for a constant PCM volume, the thinner layer of PCM with a greater surface area increases energy saving. Additionally, they also observed that when the PCM layer accounts for a substantial thickness of the entire wall, its location (external, middle, internal) has a very small impact on energy savings.

2.2.3 Optimal thermal performance of PCM

In the literature (Section 2.2.2), we observed that to exploit PCM's full potential as an energy storage system, precooling or preheating (depending on the season) of PCM with mechanical devices such as the HVAC system (active application of PCM) is more effective than passive application of PCM. However, the efficiency of PCM in active use depends on control and implementation strategies [57]. In more detail, optimising the PCM performance needs to be defined as an optimal control problem. Moreover, it is critical how we define the optimal PCM performance. For example, the optimal PCM melting point that achieves the highest resiliency of the building to extreme weather conditions yields only around $60\,\%$ of energy-saving compared to optimising the PCM's performance for energy-saving [58].

The literature on optimal control of buildings with PCMs is scarce. Most of the existing research uses simulation-based optimisation [59], with the problem variables typically including PCM properties, building's envelope properties, or the HVAC system operating conditions. The optimal solution is found iteratively by optimising the objective function with different values of the input variables so that in each iteration, the solution moves closer to the optimum. Because the set of possible input variables is limited, the optimal solution is not optimal globally. The other drawback is that simultaneously considering all the variables that affect the objective function is either infeasible or time-consuming, or requires a lot of trial and error. For example, the authors in [57] used price-based control to switch on/off the underfloor heating system in two identical test huts. They demonstrated that PCM results in peak load shifting and electricity saving. However, they only considered the electricity price as a signal to operate the heating system without considering other factors that contribute to optimal control of the heating system, such as the optimal performance of the PCM.

The other widely-used optimisation method in PCM buildings is inverse problem-based simulation. In this method, reverse engineering is used to find an optimal solution. The desired results are considered an optimisation objective, and variables of the problem are adjusted to achieve the said objective. For instance, the authors in [60] applied inverse problem-based optimisation to determine the optimal thermo-physical properties of a PCM-concrete brick. The distribution of the specific heat

of the envelope structure with temperature is used as the optimisation variable. They continuously adjust the distribution of specific heat with temperature using the sequential quadratic programming (SQP) method until the optimisation objective is achieved. In addition, they also compared this solution with particle swarm optimisation and a genetic algorithm. They concluded that SQP gives a better solution than two other methods.

The main challenge in optimising the PCM's performance is the highly nonlinear specific heat capacity characteristics equation (3.21), which throws up a couple of challenges when using Newton-based methods like SQP. First, equation (3.21) is nonsmooth at the melting point, requiring computationally expensive evaluation of derivatives (either through black-box simulations and/or via finite-differencing) [61]. Second, the resulting optimisation problem is highly nonconvex, so choosing a good starting point to prevent the algorithm from getting trapped in a local optimum is not trivial [62].

2.2.4 Co-optimisation of PCMs with distributed energy resources

To the best of our knowledge, [63] is the only paper investigating the use of PCMs in building insulation in conjunction with other distributed energy resources (DERs). Specifically, the authors in [63] investigated energy-saving and electricity demand shifting that can be achieved using insulation boards with PCM in a lightweight building equipped with a PV system and a battery storage system. In more detail, they simulated a $2.5\,\mathrm{m} \times 2.5\,\mathrm{m} \times 2.5\,\mathrm{m}$ cubic chamber equipped with an air conditioning system. They simulated the thermal behaviour of the building in EnergyPlus software and validated the simulation results with an experiment. Considering the building with a PV system without PCM as a baseline case, they observed a $47\,\%$ reduction in peak cooling load and an hour shift in the cooling demand in the summer. The battery charging and discharging is determined by a simple heuristic to maximise PV self-consumption. In more detail, the HVAC system is supplied by PV first, and when the PV generation exceeds the HVAC consumption, the excess energy is stored in the battery. When the PV generation is insufficient, the energy is taken from the battery, and when the battery is at the low discharge range, the power is taken from the grid. The authors tested the performance of the test chamber with different battery PV system sizes. The main drawback of [63] is the lack of a principled optimisation to achieve the optimal HVAC performance in a more realistic setting with and underlying electricity consumption and time-varying electricity prices. Also, [63] only considered a single summer day, which leaves the question of the PCM performance throughout the year.

2.3 Home energy management system in PCM-integrated buildings

In this section, we first describe the concept of a HEMS. This is followed by a review of techniques extensively used in the literature to solve the optimisation problem of HEMS. Next, we highlight the challenge of solving an optimisation problem in HEMS with PCM. And finally we review the reinforcement learning methods used in this thesis to develop the novel model-based (Chapter 5) , and

model-free approaches (Chapter 6) to solve an optimisation problem of HEMS in PCM-integrated buildings.

2.3.1 Home energy management system

With the advent of advanced communication and metering devices and massive growth of DERs, a HEMS can be used to reduce residential energy consumption and to modify the usage pattern of home devices in order to achieve more efficient use of distribution networks and renewable electricity generation. In more detail, a HEMS is a device that monitors, controls, and manages the operations of DERs to modify electricity usage patterns during peak periods in response to electricity price signals. It achieves this by means of a 2-way exchange of information with the grid through communication technologies and advanced control methods [64–66].

2.3.2 Home energy management system with PCM

In the HEMS optimization problem, the objective is to schedule the controllable devices to reduce energy cost while maintaining comfort thermal conditions for the home users. The objective is subject to constraints like energy balance, comfort boundaries and limitations of the controllable devices. What makes the problem of the PCM-integrated buildings distinct from other HEMS formulations is the nonlinear behaviour of the PCM, which results in a nonconvex model.

For solving an optimisation problem in HEMS, different methods are applied in the literature [15–18]. One of the widely used methods to solve HEMS problems are *linear programming (*LP) and *mixed-integer linear programming (*MILP). These methods are easy to implement using commercial solvers such as CPLEX, Gurobi and MOSEK or XPRESS. The main drawback of these methods is that they optimise linear objective functions subject to linear constraints. Therefore, in solving the nonlinear nonconvex optimisation problem of HEMS with PCM, these methods are powerless, and the solution quality cannot be guaranteed [15]. However, the optimisation problem in HEMS with PCM can be formulated as a mixed-integer nonlinear programming (MINLP) problem and solved with different commercial solvers such as BONMIN or KNITRO. However, since the HEMS with PCM problem is a nonconvex problem, these solvers return a locally optimal solution. Moreover, for large scale optimisation problems, using the MINLP method is computationally intractable [18]. A number of heuristic methods have been extensively applied in literature to solve the HEMS problem, such as *particle swarm optimisation (*PSO) and *genetic algorithm (*GA). The algorithms search in a large population search space semi-randomly until they converge to a solution. With these methods, the solution may end up in a local optimum instead of the global optimum, which makes the quality of the solution doubtful [17].

Against this background, a method that can deal with the nonlinear non-convex optimisation problem of PCM-integrated buildings is *dynamic programming* (DP). To solve a problem with DP, the problem has to be formulated as a Markov decision process (MDP). The drawbacks of using the DP approach are it is computationally expensive, and to solve the MDP using DP, a dynamic of the

underlying system is required. In response to these issues, *model-free reinforcement learning* (RL) can solve the MDP where the system model is unknown.

Given this, the next section is devoted to reviewing the DP in more detail, followed by introducing *model-free reinforcement learning* (RL) that provides the foundation for our novel contributions in this thesis.

2.3.3 Dynamic Programming

The method for dealing with the nonlinear characteristic of PCM is *dynamic programming* (DP) [67]. DP is known for solving planning and sequential decision problems, and recently it has been applied for solving an optimisation problem in HEMS [16, 68].

To solve a problem with DP, the problem has to be formulated as a *Markov decision process* (MDP). However, the requirement for this is the problem has to have a sequentially-separable structure or has the *Markov property*[3].

For our specific HEMS problem, the MDP consists of accumulated instantaneous costs or rewards over a decision horizon. To measure the quality of a policy, there are two ways of (i) calculating state value function, which is known as *value function*, or (ii) state-action value function, which is known as *Q-function*. The value function is formed by adding the expected future cost of following a *policy* (in this problem specific on/off sequence of HVAC system) given state transition probabilities. Importantly, the objective is to find the minimum value function over the time horizon of the problem. In doing so, the *value iteration* (VI) algorithm is employed in DP, which computes the minimum value function in a backward fashion using the Bellman optimality condition[4] [20]. With this, an optimal policy can be derived by tracing back the states with minimum value functions over the time horizon of the problem.

However, DP has a limitation that is known as the *curse of dimensionality*, which means VI becomes computationally intractable when the time horizon of the problem, number of state variables, or number of controllable devices grows [20]. More specifically, solving this problem with a pure dynamic programming algorithm entails a huge computational burden that makes impossible its application in the current context of DR.

In response, the Artificial intelligence (AI) literature contains methods, namely *state-space approximation*, that leads to approximate dynamic programming (ADP). However, despite state-space approximation in ADP, it still cannot be applied to a large scale optimisation problem due to its computational cost [69].

The other drawback of DP is that it is a model-based method. In more detail, to apply DP to solve the HEMS problem in PCM-building, an explicit thermal dynamic of PCM-integrated buildings is required as part of optimisation problem formulation. As a result, *model-free reinforcement learning* methods have garnered considerable attention in recent years.

[3] A process has a *Markov property* if the future state depends only on the current state, not on the sequence of events that preceded it.

[4] An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

2.3.4 Model-free Reinforcement learning

One of the important machine learning algorithms that have recently attracted growing research interest is reinforcement learning (RL). RL has been demonstrated to be a powerful tool in solving realistic decision-making problems such as intelligent robots, game competitions, and industrial manufacturing, to name a few [70, 71]. In addition, we have witnessed an increase in RL application in energy management in building over recent years [25, 72–74].

RL consists of an agent and a learning environment. An agent is a decision-maker or learner. Anything outside the agent forms the environment. The agent takes action depending on the state of the environment and is rewarded for taking desired actions. The agent tries to learn actions that maximise the total amount of reward over time [75].

The learning environment is often modelled as a Markov decision process (MDP). MDP consists of state, action, and transition functions. As mentioned in Section 2.3.3 where the state transition functions in MDP are known, DP can be used as a solution technique. In contrast, where the agent does not have access to the explicit model of the environment, model-free RL can be used to solve the MDP. In the model-free RL, the agent learns the optimal policy with a trial and error interaction with an environment.

An evident advantage of using model-free RL is that users can avoid the tedious work of developing a detailed building model and its verification. A model-free RL method that is used predominantly due to its simplicity is *Q-learning* [74, 76]. The Q-learning is an off-policy RL method that values of each state-action pair is stored in a lookup table (Q-table). In Q-learning, the agent tries to maximise the reward; therefore, it learns the best action taken from each state.

The tabular format of the Q-learning makes it ineffective for a problem with a large state-space. However, with the recent advances in RL, a combination of RL with deep neural networks can elevate this issue to an acceptable level. In more detail, the Q-table is replaced with deep neural network function approximators that estimate action-value functions [71]. This method is known as Deep Q-learning (DQN). However, applying DQN on continuous control problems requires discretisation of the state-action space, which reduces the control precision and performance quality in continuous control problems such as the HVAC control. Consequently, the optimal policy can be a suboptimal solution rather than an optimal one.

Against this backdrop, work in [23] presented a model-free actor-critic on-policy reinforcement learning method based on deep deterministic policy gradient (DDPG) that can learn policies in continuous high dimensional action-space. The model-free DDPG method obviates the need for an exact mathematical model of a building. Another advantage of the DDPG is that it can apply to continuous control problems as opposed to the MADP method. In the actor-critic method, there are two separate memory structures, namely actor and critic. The actor takes action like on/off of the HVAC system based on observation from the environment that can be a sensory input like the indoor temperature of a building. The critic observes both state and the reward obtained based on actors' action. As its name implies, critic critiques the action that an actor takes. In more detail, the critic uses a loss function (difference between estimated and expected reward) to evaluate the action taken

by actor. Afterward, the actor updates policy parameters, in the direction suggested by the critic. A detailed explanation of the model-free DDPG method is given in Chapter 6.

DDPG for scheduling the HVAC systems in buildings was initially proposed in [25]. They used DDPG for controlling the setpoint of the HVAC system by minimising the overall cost of the HVAC system's electricity consumption while considering the thermal comfort of the home users. However, the quality of the solution that results from the DDPG algorithm was not benchmarked against well-proven algorithms like ADP that have full access to the model of the system.

2.4 Thermal model of a building

To study the thermal inertia of the building envelope, we require a model of the dynamic or transient thermal behaviour of the building's components. Therefore in the first part of this section, we introduce software tools that are widely used for simulating the thermal performance of buildings. In the second part, we present a short description of the method that is applied in this thesis to extract a mathematical model of a building with PCM.

2.4.1 Available software for thermal modelling of a building

In recent years, several software tools have been developed for simulation studies of the thermal behaviour of buildings. However, only a few of these tools have been validated by evaluating their performance in building energy simulation against the ANSI-ASHRAE Standard 140[5] [77]. EnergyPlus, TRNSYS and ESP-r are among the tools that comply with the standard.

EnergyPlus is an open-source whole building simulation software that was developed by the U.S. Department of Energy. Compared to TRNSYS and ESP-r, EnergyPlus is used widely in the literature [78]. EnergyPlus software uses the conduction transfer function (CTF) to estimate the heat transfer in buildings [79, 80]. The CTF is essentially the lumped parameter method expressed as transfer functions using the state-space method. CTFs are efficient in calculating surface heat fluxes since they eliminate the need to know temperatures and fluxes within the surface. However, the CTF series gradually becomes unstable with a decrease in the time step. This problem is overcome in EnergyPlus by using a master history with interpolation method. In this method, surface temperature and heat flux histories at the intermediate instance of time are obtained by interpolation. One drawback of CTF is that it's a constant coefficient (being a time-invariant method). Hence it cannot be used for modelling temperature-dependent thermal properties like PCM.

To model PCM, partial differential equations with a moving boundary need to be formulated. To solve the governing equations, either numerical or analytical methods can be used. In EnergyPlus, a one-dimensional conduction finite difference (CondFD) is a method applied to solve this problem, which saves considerable time and computational cost. The CondFD algorithm is validated through some on-site experiments of PCM's performance and also through the standard evaluation of building

[5]ANSI-ASHRAE Standard 140 specifies test procedures for assessing the technical capabilities and ranges of applicability of computer programs that simulate whole buildings' thermal performance [77].

software published by ASHRAE (the American Society of Heating, Refrigerating and Air-conditioning Engineers, Inc.). To use this algorithm, there is an option to select either the Crank-Nicholson or fully implicit methods: The fully implicit method is used in this work. A detailed explanation of the CTF and CondFD is given in [81].

The drawback of using building simulation software such as EnergyPlus is that users can only access input and output parameters. This limits the use of this software in applications that requires explicit mathematical equations that govern the thermal performance of the building. As one of the main contributions of this dissertation, we develop a computationally-efficient optimisation method to solve the HEMS problem in PCM-enhanced buildings. Formulating the optimisation problem necessitates access to the underlying differential equations that govern the building thermal dynamics. Therefore, in the next part, we briefly describe the RC lumped model as a method we use to simulate the thermal behaviour of the building.

2.4.2 Thermal RC lumped model of a building

The modelling approach used in this research is a lumped Resistance-Capacitance method (known as RC model). Lumped models were introduced for the first time by Laret in 1980 [80, 82]. This work was followed by [83], on which the heating system was included in the model, and the theoretical results were compared to the outcomes of field monitoring of the building. Laret concluded that there was an excellent agreement between simulation and experimental outputs. Moreover, he highlighted that this type of modelling is simple, sufficiently accurate and computationally efficient.

Some authors, instead of modelling individual construction elements with an RC lumped model, used lumped method to model the whole building thermal performance [84–90]. In this approach, all elements of buildings with a significant thermal capacitance are lumped and simulated as a single equivalent capacitance. On top of that, the indoor air zone and internal mass (such as furniture) are modelled as an additional capacitance. As a result, a second-order RC lumped model that consists of two capacitances can simulate the thermal behaviour of a building [90].

The main feature to model in building with PCM is the PCM model. Due to the phase change of this material, the type of model to describe the thermodynamic of PCM is known as Stefan problem that has moving boundary conditions. In the problem formulation, the heat transfer equations are written separately for liquid and solid phases of PCM. The Stefan condition enforces the heat balance at the solid-liquid interface [41]. A few numerical methods are used to solve this problem in the literature [41, 91–93]. What is common in different used methods in literature is the first PCM is modelled as a fixed space grid. Afterwards, either the enthalpy method or capacity method is employed to capture the high variation of the latent heat during the phase-change of PCM. Then, a numerical method is used to solve the resulting one-dimensional diffusion equation. Authors in [91], and [93], applied a linear solver of iterative the Gauss-Seidel method. Works in [41] and [92] used the finite difference method and the finite element method, respectively. The accuracy of all these methods depends on the discretisation space and time step. The smaller time step and discretisation result in

a higher quality of the solution, however, in the cost of increasing the computation time. Therefore, all of these methods are computationally expensive. In particular, where we deal with a large-scale optimisation method in buildings with PCM that require a thermal dynamics of PCM as part of the optimisation formulation.

In contrast, model PCM as an additional capacitance in a simplified RC model of a building can result in a computationally efficient numerical solution. Besides low computational cost, and the simplicity of the model, as demonstrated by authors in [94], the model outputs can have an acceptable level of accuracy compared to a theoretical model that is solved using the finite element method.

2.5 Summary and literature gaps

This chapter presented the necessary background material to our work. The chapter started with a review of the literature to demonstrate the potential of thermal inertia in building to be utilised as thermal energy storage for energy saving in buildings. Next, the discussion continued with an introduction to PCM as a proven technology to improve the thermal inertia of buildings. This technology motivates the use of thermal inertia of building for demand response in countries like Australia that a large number of residential buildings are made from low thermal inertia material. On this, Section 2.2 starts with a short description of the PCM thermo-physical properties and different types of PCM, particularly the types of PCM that can be applied in buildings. The section ends with a brief description of the BioPCM that is used in this dissertation. Following this section, the review moved on to illustrate the promising performance of PCM integration into the building envelope, with the focus on its feasibility to reduce or shift the space heating and cooling demand. In Section 2.2.3, the thermal performance of PCM was investigated in more detail. In particular, it was shown that PCM thermal performance is complex. It is affected by PCM thermo-physical properties (such as thermal resistance, PCM melting temperature, thermal conductivity, latent heat), geometrical parameters (such as PCM location, PCM thickness), and the parameters of the building that PCM is applied (such as the HVAC system setting point). Moreover, we highlighted the methods used in PCM literature to explore the optimal performance of PCM, such as simulation-based optimisation or optimisation method such as SQP. However, as discussed, the former approach is either infeasible or time-consuming to consider simultaneously all the variables that affect the objective function. Moreover, methods such as SQP cannot always yield a globally optimal solution as its performance depends on the choice of an algorithmic starting point.

As of yet, none of the works in the literature formulated the effort to seek the optimal performance of PCM in building as a HEMS optimisation problem. Therefore, next, we described a HEMS that is used as a platform for defining energy management optimisation problems in this work. The discussion was extended further by a brief introduction of widely used approaches in the HEMS literature to solve the optimisation problem in HEMS. However, we highlighted that the nonlinear nonconvex feature of PCM's specific heat equation makes the corresponding problem challenging to solve using the existing method in the literature. Therefore, in Section 2.3.3, the literature focus was on describing DP as a principled method to solve the nonlinear nonconvex optimisation problems. Moreover, the

shortcomings of DP, including (i) curse of dimensionality and (ii) it requires an explicit model of system dynamics (in this work; explicit mathematical representation of building thermal performance), were brought out for attention. Next, model-free RL methods are reviewed as techniques that can be used in practical instances where the access to the thermal model of building due to high variability in building design and construction is infeasible. However, using the model-free RL method, in particular, DDPG, does not eliminate the need for a thermal model of the building. Still, the thermal dynamics of the building is required to use as a learning environment for the model-free RL method.

Therefore, Section 2.4 starts with introducing widely used software for building simulation. Using widely used software such as EnergyPlus provided users with a validated tool against ASHRAE standards to simulate the building's thermal dynamics. However, the exact mathematical formulation of a thermal building is not accessible for the users. This is considered as a drawback where the mathematical equations that govern the thermal behaviour of the building are required to incorporate explicit dynamic formulation of a building to an optimisation problem. Therefore, the section is extended further by introducing RC lumped approach as a method that can be used for modelling the thermal behaviour of a building with an acceptable accuracy against a validated building simulation software such as EnergyPlus. In more detail, the RC approach can be used to capture the thermal dynamics of the building as a set of ODEs.

In analysing this literature, the identified gaps from the literature review above can be summarised as follows:

- The literature lacks an extensive analysis of all factors that influence PCM performance as thermal energy storage, such as geographical location, duration of precooling and preheating of PCM with the HVAC system, and the HVAC system setting point, to name a few.
- To our knowledge, no work has been done on classifying an optimal performance of PCM-buildings as a HEMS problem and sought a computational method that can deal with a nonlinear nonconvex feature of PCM.
- Literature on using the principled DP method to solve an optimisation method in HEMS is scarce due to its well-known drawback of the curse of dimensionality. In particular, limited work attempts to speed up the DP algorithm for energy management applications.
- In light of the soaring uptake of DERs, considering PCM as an alternative to battery energy storage in a HEMS setting is not investigated thoroughly considering real-world variability of electricity demand, solar generation and weather conditions.
- With recent advances in technology, HEMS can become an inevitable part of energy management in buildings. However, current literature lacks investigating a PCM as a part of HEMS considering the real-world application. In particular, the existing literature gap is (i) the absence of a computationally-efficient optimisation method that can be applied for continuous control problems such as the HVAC scheduling and (ii) in energy management of any building regardless of its design and construction type.

- There are limited works on using the DDPG method for energy management applications in buildings. Prior to our work in [95], only [25] applied DDPG for scheduling the HVAC system in a building. However, in [25], the quality of the DDPG solution compared to the DP method that has full access to the model of the system is not investigated.

Given this background, the coming chapters will fill the literature gaps mentioned above. Chapter 3 provides a necessary step towards the main contributions of this thesis. In more detail, in Chapter 3, a simplified thermal model of a typical building will be developed to propose a computationally-efficient technique (MADP) for solving the HEMS optimisation problem in building with PCM (chapters 5, 7). Moreover, the developed thermal model of the building will be used as a learning environment for developing a model-free DDPG algorithm in chapters 6 and 8. Technical chapters 4, 5, and 6, 7, and 8 contain the key contributions of this dissertation. In each chapter, the literature gaps within the scope of the chapter and the contributions are discussed. The reader is referred back to Figure 1.1 for the connections between subsequent chapters and the literature review developed in the present chapter.

Chapter 3

Thermal RC Lumped Model of PCM-building

The first step towards solving an optimisation problem in a home energy management system (HEMS) with phase change material (PCM) is to model the thermal behaviour of the building. In particular, an explicit representation of the relevant mathematical equations is required.

Several software tools have been developed for simulation studies of buildings' thermal behaviour in recent years, however users can only access inputs and outputs. This limits the use of the software, particularly preventing them from being employed in conjunction with optimisation algorithms that require mathematical expressions of the buildings' thermal dynamic. We note that there are some interfaces developed to make the connection between the building thermal model software and programming language. However, using these interfaces to integrate the thermal model as a black box into the optimisation methods based on DP is not computationally efficient (this is explained in more detail in Chapter 5). In other words, to increase the computational efficiency of the optimisation algorithms, the model needs to be representable and straightforward with a low order of ordinary differential equations (ODEs) while sufficiently accurate to be validated by the widely used existing software such as EnergyPlus. In this work, the RC lumped approach is used for simulating the thermal behaviour of the building. As discussed in Chapter 2, RC lumped model is the most popular for buildings because of its explicit physical meanings and low computational complexity. Furthermore, the model requires to be representative of a lightweight building in Australia. The type of construction material and the fenestration were chosen based on the common practice in lightweight dwellings in Australia.

Adding PCM in this model adds another level of complexity due to the phase change (from solid to liquid or vice versa) of this material. However, for this work, PCM is simply modelled as temperature-varying heat capacity. Moreover, we consider an assumption that makes the modelling of PCM building simple to avoid the extra computational burden that increasing ODEs order can bring to the optimisation algorithm. In more detail, we assume that the PCM layer is placed in the vicinity of the highest thermal inertia of the building. Therefore we can lump these two major thermal capacitances by adding them together.

In summary, the contributions of this chapter are as follows:

- Present a simple model of a common light weight residential building in Australia, presented with a second-order ODEs that offers a good compromise between simplicity and computational expense of the applied optimisation algorithm in HEMS.
- Including PCM in the developed thermal model without adding computational complexity
- Using construction materials of the model as per the current practice of a lightweight building in Australia.
- Validate the developed model by benchmarking against the identical model in EnergyPlus software.

This chapter starts by explaining the heat transfer equation that governs the heat transfer in the building's element. This will be used as a starting point for the derivation of the RC lumped model of a building. Afterwards, we present an RC lumped model of a typical residential building. The building construction and materials are selected to represent a typical lightweight building in Australia. Next, we introduce the type of PCM used in this work along with its formulation. The RC lumped model of the PCM building is built on the RC lumped model of a lightweight building by adding extra nonlinear thermal capacitance for simulating PCM in the building's envelope. And finally, the RC lumped model of both lightweight building and building with PCM are built in MATLAB and benchmarked against identical models in EnergyPlus. RMSE is used as a measure to show the discrepancy between the two models.

The thermal model that is developed in this chapter will be used in Chapters, 5 and 7, to represent the system dynamics and in Chapters 6 and 8 as the DDPG learning environment.

3.1 Modelling of heat transfer in buildings

The applied modelling method in this work is the RC lumped model approach [80, 82, 83]. To explain the RC lumped method in more detail, we start with the heat equation. The heat equation that governs heat transfer through a homogeneous material is devised from Fourier's law and energy conservation. Fourier's law is given as follows:

$$q = -k_e \Delta T \tag{3.1}$$

where q is heat flux ($\mathrm{W/m^2}$), k_e is thermal conductivity ($\mathrm{WK/m}$) and ΔT is the thermal gradient. With the assumption of work done is zero, internal energy change per unit volume in the material (ΔQ) is proportional to the temperature change as given by:

$$\Delta Q = c\rho \Delta T \tag{3.2}$$

where ρ is the mass density of the material ($\mathrm{kg/m^3}$), and c is the specific heat capacity. Combining two equations of (3.1) and (3.2), give us:

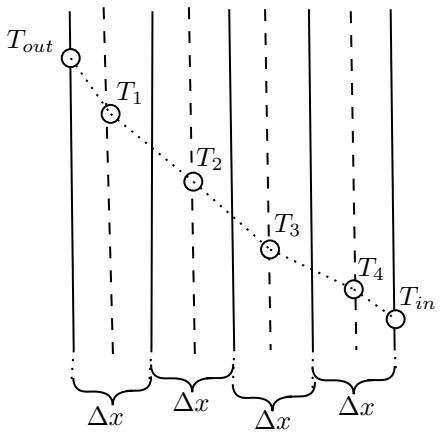


Figure 3.1: Finite difference method.

$$\rho c \frac{\partial T}{\partial t} - \Delta \cdot (k_e \Delta T) = \dot{q}_{\mathrm{v}} \tag{3.3}$$

where k_e is thermal conductivity ($\mathrm{W/mK}$) and $\dot{q}_{\mathrm{v}}$ is the heat gain inside the volume. Assuming that no internal heat gain or sink exists inside the building's element, $\dot{q}_{\mathrm{v}}$ is zero. Therefore, the equation (3.3) can be written as:

$$\frac{\partial T(x,y,z,t)}{\partial t} - \beta \Delta^2 T(x,y,z,t) = 0 \tag{3.4}$$

where β is thermal diffusivity defined as:

$$\beta = \frac{k_e}{c\rho} \tag{3.5}$$

Heat equation (equation (3.4)), can be simplified further by a few assumptions:

- The heat transfer needs to be modelled in only one direction, that is the direction across the building's element thickness (typically assumed x-direction). In other words, heat transfer in two directions of y- and z-directions can be ignored.
- Heat transfer in the material is isotropic that is the same in all directions.
- The properties of the material are independent of temperature variations.

To solve the heat equation, the most popular numerical technique is the so-called finite difference method. In this method, each element of the building is divided into a finite number of slices. The temperature across each slice is considered constant, and the temperature of a node located in the central plane of the slices is used to present the state of each slice. Moreover, the temperature variation between adjoining nodes is assumed to be linear. For clear explanation, Fig. 3.1 illustrates a building element that is divided into four slices.

Let represent the temperature of slices $i-1$, i, $i+1$, by T_{i-1}, T_i and T_{i+1} respectively. Assuming the width of each slice Δx, the temperature gradient is given by:

$$\frac{T_{i+1} - T_i}{\Delta x} \quad and \quad \frac{T_i - T_{i-1}}{\Delta x} \tag{3.6}$$

Because the rate of conduction is proportional to the temperature gradient, we can approximate the heat flux between nodes $i-1$ and i as q_{i-}, and between nodes i and $i+1$ as q_{i+}, therefore we have:

$$q_{i-} = -k_e \frac{T_i - T_{i-1}}{\Delta x} \tag{3.7}$$

$$q_{i+} = -k_e \frac{T_{i+1} - T_i}{\Delta x} \tag{3.8}$$

Thus, the heat balance of slice is given by:

$$\rho c \Delta x \frac{\Delta T_i}{\Delta t} = -k_e \frac{T_i - T_{i-1}}{\Delta x} + k \frac{T_{i+1} - T_i}{\Delta x} \tag{3.9}$$

Rearranging

$$\frac{\Delta T_i}{\Delta t} = \beta \frac{T_{i-1} - 2T_i + T_{i+1}}{\Delta x^2} \tag{3.10}$$

Considering

$$\frac{\partial T}{\partial t} \approx \frac{\Delta T}{\Delta t} \tag{3.11}$$

and

$$\frac{\partial^2 T}{\partial x^2} \approx \frac{T(x + \Delta x) - 2t(x) + T(x - \Delta x)}{\Delta x^2} \tag{3.12}$$

The heat equation can be described as:

$$\frac{\partial T}{\partial t} = \beta \frac{\partial^2 T}{\partial x^2} \tag{3.13}$$

By assuming uniform temperature across each building's element, each element of the building can be represented a single dimensionless lump [80].

3.2 Thermal model of a lightweight building

As explained in Section 3.1, for energy management application, each element of a building can be presented as an RC lumped model. Due to the close resemblance of the thermal energy balance to Ohm's law, electric resistance and capacitance can be considered analogous to thermal resistance and capacitance of the building. Therefore, each element of the building is simulated as an RC electric circuit. In this work, for simplicity, all elements of the roof, walls and floor are lumped together as a united 2R1C model (two lumped resistances and one lumped capacitance), as shown in Fig. 3.2. In more detail, the parallel R_{dw} represents the fenestration of the building (such as windows and doors) and is the sum of the thermal resistances of the door and windows. The infiltration heat loss, the HVAC system power that enters the building are represented by Q_{inf}, and Q_{HVAC} respectively. To simulate the indoor air, $m_{\mathrm{a}} c_{\mathrm{a}}$ is considered in the model, while the C_{e} is thermal inertia (capacity) of the envelope that we aim to improve and use as a storage system. In this model, the total thermal resistance of the building's element is divided into two resistance of R_{in} and R_{out} that are named inner

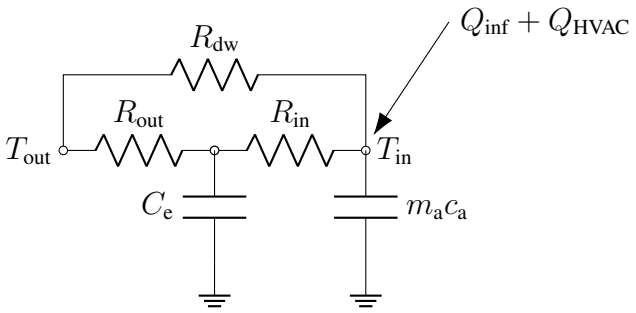


Figure 3.2: 2RC lumped model of lightweight building

and outer resistances, respectively. And finally, T_{in} and T_{out} represent indoor and outdoor temperatures, respectively.

The studied building is a simple, one-zone a rectangular prism with $8\text{m} \times 6\text{m} \times 2.7\text{m}$ dimension and a total floor area of $48\,\text{m}^2$. To reflect a real lightweight building in Australia, the building elements type and arrangement are selected based on a common lightweight building in Australia that are presented in [2]. Moreover, the material properties are extracted from ASHRAE handbook [1]. Therefore, the building (from outside to inside) is made up of three layers of rendered fibro-cement, a timber stud wall containing insulation batts, and a plaster board [2]. The properties of the building's materials are shown in Table. 3.1, and the fenestration is detailed in Table 3.2.

Table 3.1: Building elements composition and its material properties [1, 2].

Element	d (m)	k_e (W/mK)	ρ (kg/m^3)	c (J/kg K)
Rendered fibro-cement	0.005	0.25	1150	840
Timber studwall with insul. batts	0.09	0.15	650	1200
Plaster board	0.01	0.25	950	840

Table 3.2: Fenestration details [3].

Element	**Description**	U (W/m^{2}K)	Area (m^2)
Windows	Single glazing with Aluminium frame	7.01	7.8
Door	Wooden slab with wooden frame	2.61	2.1

Computing inner (R_{in}) and outer resistances (R_{out}) of Fig. 3.2, necessitates calculation of total thermal resistance (R_{e}) and total thermal capacitance (C_{e}) of the buildings'element based on the given (3.14) and (3.15):

$$R_{\text{e}} = \left(r_{\text{si}} + r_{\text{so}} + \sum_{l=1}^{n} \frac{d_{\text{l}}}{k_{\text{l}}} \right) / A_{\text{e}} \tag{3.14}$$

$$C_{\mathrm{e}} = A_{\mathrm{e}} \sum_{l=1}^{n} (d_{\mathrm{l}}.\rho_{\mathrm{l}}.c_{\mathrm{l}}) \tag{3.15}$$

In an international reference book for buildings [1], r_{si} and r_{so} have values of 0.13 $\mathrm{m^2\,K/W}$ and 0.04 $\mathrm{m^2\,K/W}$; the same values are used here. Refer to (3.16) and (3.17), to obtain the values of R_{in} and R_{out}, where total resistance of the element is multiplied by a factor (α) [80]. This factor is called the accessibility factor and varies among the available literature, and this causes ambiguity for the researchers in this area. In [96], the author performed an extensive investigation by frequency response analysis of the building and conducting experiments on a set of houses and office rooms. As an outcome, they estimated factors for different elements of building with various boundary conditions. Based on their results, α=0.01 is used in this model.

$$R_{\mathrm{in}} = \alpha R_{\mathrm{e}} \tag{3.16}$$

$$R_{\mathrm{out}} = (1-\alpha) R_{\mathrm{e}} \tag{3.17}$$

Finally, the thermal time-derivation equations (dot above variables shows time-derivation of that variable) of the model are given in (3.18), (3.19) and (3.20):

$$\dot{T}_{\mathrm{e}} = \frac{1}{C_{\mathrm{e}} + C_{\mathrm{PCM}}} \left(\frac{T_{\mathrm{in}} - T_{\mathrm{e}}}{R_{\mathrm{in}}} + \frac{T_{\mathrm{out}} - T_{\mathrm{e}}}{R_{\mathrm{out}}} \right), \tag{3.18}$$

$$\dot{T}_{\mathrm{in}} = \frac{1}{m_{\mathrm{a}} c_{\mathrm{a}}} \left(\frac{T_{\mathrm{out}} - T_{\mathrm{in}}}{R_{\mathrm{dw}}} + \frac{T_{\mathrm{e}} - T_{\mathrm{in}}}{R_{\mathrm{in}}} + \dot{Q}_{\mathrm{HVAC}} + \dot{Q}_{\mathrm{inf}} \right), \tag{3.19}$$

where:

$$\dot{Q}_{\mathrm{inf}} = \frac{m_{\mathrm{a}} c_{\mathrm{a}} (T_{\mathrm{out}} - T_{\mathrm{in}}) ACH}{3600} \tag{3.20}$$

In this work, based on [1], the assumed value for ACH [1] is considered as 0.3 h. Note that (3.18), (3.19) and (3.20) can not be solved by fixed-step solvers since as shown in [97], the outputs become unstable. For solving these differential equations in MATLAB, solver ode23s with variable time step is adopted in this work. This solver is based on the Runge-Kutta method, which uses a variable, continuously adjusted time step.

In the next section, PCM is added to the current model to achieve a model that we aim to use in the optimisation problem.

3.3 Including PCM in thermal model

Including PCM in the thermal model of a building is not a straightforward task because of the solid-liquid phase transition. In this section, the simple way of modelling PCM is presented.

For a typical PCM, the specific heat capacity variation by temperature is as shown in Fig. 3.3 [98]. The temperature range of 22 $^{\circ}\mathrm{C}$ to 28.5$^{\circ}\mathrm{C}$ is where the phase changing occurs. At 27.6$^{\circ}\mathrm{C}$ (the melting

[1] ACH is an abbreviation of Air Changes per Hour which is a measure of changes in air volume for a specific space.

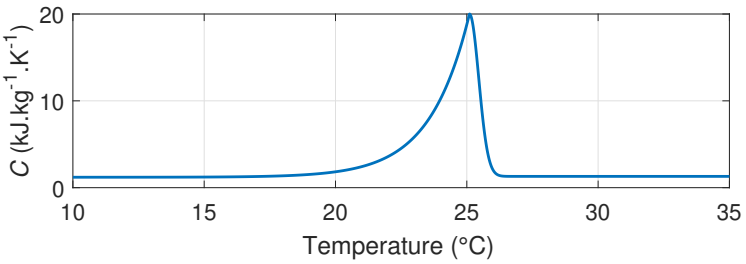


Figure 3.3: Specific heat capacity plot of a typical PCM

Table 3.3: Honeycomb PCM properties [3].

PCM type	d (m)	k_e (W/mK)	ρ (kg/m^3)	c (J/kg K)
Honeycomb PCM	0.03	2.8	545	varies

point), the specific heat capacity has its highest value. This nonlinear curve of specific heat capacity of PCM is what we referred to in the previous chapters as the nonlinear behaviour of PCM [3, 98].

The type of PCM that is utilized in this research is honeycomb PCM that is studied in [3]. The reason for using this type of PCM is that it is the closest product to commercial BioPCM[2] and more importantly the mathematical equations of PCM's specific heat capacity are provided in [3]. Therefore it can be easily plugged into the thermal model equations. The honeycomb PCM consists of a honeycomb matrix that is enclosed in aluminium sheaths. The formulas of specific heat capacity are given in (3.21a) and (3.21b).

$$c_{\text{pcm}} = 1200 + 18800e^{-\left(\frac{T_{\text{p}}-T}{1.5}\right)} \quad \text{if} \quad T < T_{\text{p}}, \tag{3.21a}$$

$$c_{\text{pcm}} = 1300 + 18700e^{-4(T_{\text{p}}-T)^2} \quad \text{if} \quad T \geq T_{\text{p}}, \tag{3.21b}$$

where T_{p} is the melting point of the PCM. Due to the discontinuous form of (3.21a) and (3.21b), they can not be directly substituted continuous equations of in (3.18), (3.19) and (3.20).

To overcome this, the equations are curve fitted with a continuous polynomial function and the resulted function is applied as c_{pcm} or specific heat capacity of PCM in the equations. Specifically, the total heat capacitance of utilized PCM (C_{PCM}) in the building can be calculated from the properties of PCM that are shown in Table. 3.3.

Note that the PCM layer is placed underneath the plasterboard layer. As a result, the thermal resistance between the PCM layer and plasterboard is very small. This means that the two layers with high thermal inertia (PCM layer and timber studwall) can be lumped together as a single capacitance. Therefore, the PCM's surface temperature and the building's indoor temperature are approximately the same. The thermal model of the building with PCM is based on the electric circuit shown in Fig. 3.4. The only difference between Fig. 3.4 and Fig. 3.2 is that C_{PCM} is added to the thermal capacitance of the building.

All equations (3.18) to (3.20) are applicable to the PCM-building, and only (3.20) is changed. In more detail, C_e is replaced by $C_e + C_{PCM}$ to form (3.22):

[2]http://phasechange.com.au

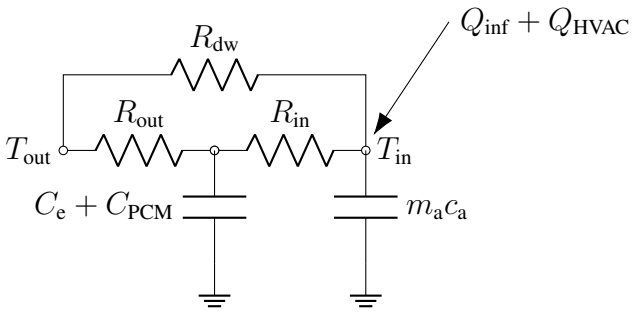


Figure 3.4: 2RC lumped model of PCM-building

$$\dot{T}_{\text{e}} = \frac{1}{C_{\text{e}} + C_{\text{PCM}}} \left(\frac{T_{\text{in}} - T_{\text{e}}}{R_{\text{in}}} + \frac{T_{\text{out}} - T_{\text{e}}}{R_{\text{out}}} \right) \tag{3.22}$$

In the next section, the validity of the described model is checked by benchmarking against EnergyPlus sofware.

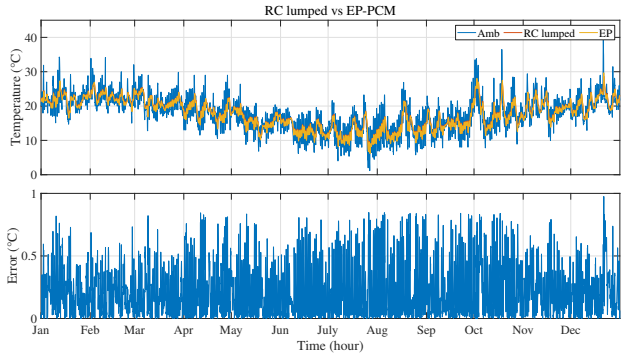
3.4 Benchmarking developed thermal model in MATLAB against EnergyPlus

To check the validity of the proposed thermal model, an identical model is built in EnergyPlus software. The simulations are run for outdoor air dry-bulb temperature of Sydney over a year, summer month, and winter month time horizons in both NOPCM and PCM-included buildings. Fig. 3.5, presents the comparison between the two models. In more detail, the top curves in each subfigure of Fig. 3.5 compare the indoor temperature that results from two models. Moreover, the outdoor temperature is also included in the plots of Fig. 3.5, to show the performance of the models in response to the outdoor temperature variations. Root-mean-square error (RMSE) is used as a measure to show the difference between the thermal model in MATLAB versus the model in EnergyPlus (the bottom curve of each subfigure in Fig. 3.5). We can witness that the maximum value of RMSE in both NOPCM and PCM, over the year time horizon, is almost less than 0.8 $^{\circ}\text{C}$ which is acceptable regarding the model uncertainties and also human sensitivity.

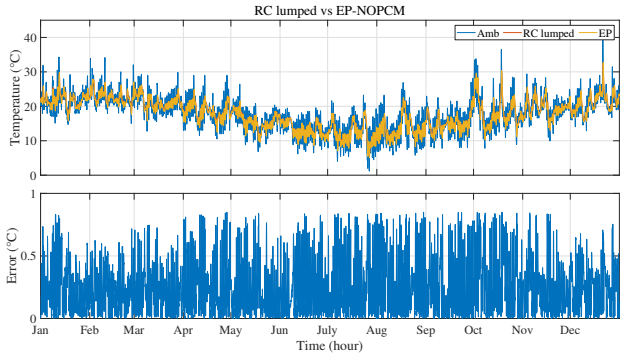
Another observation that is more pronounced in monthly simulations of Fig. 3.5 is that the indoor temperatures that result from the two models follow each other very closely. It shows the acceptable approximation of the developed RC lumped model compared to the model that is built in EnergyPlus.

3.5 Summary

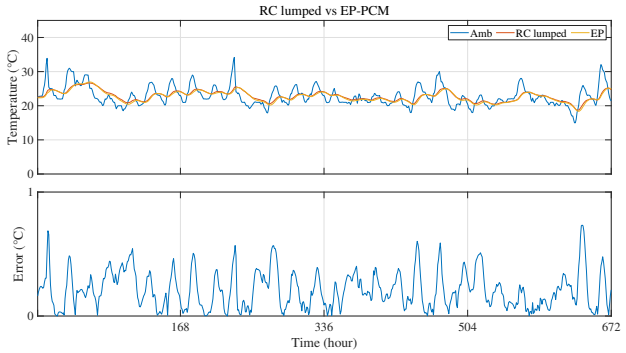
In this chapter, the thermal model of a building is developed using an electric circuit analogy or an RC lumped model. In this method, each building element is modelled with a combination of thermal resistance and thermal capacitance. Treating temperatures as potentials draw an analogy with an RC electric circuit. To reflect the reality, the materials and configuration of a common lightweight building are adopted in this work.



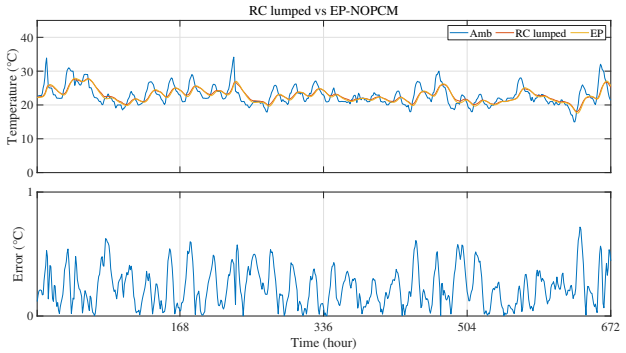
(a) Over a typical year in Sydney



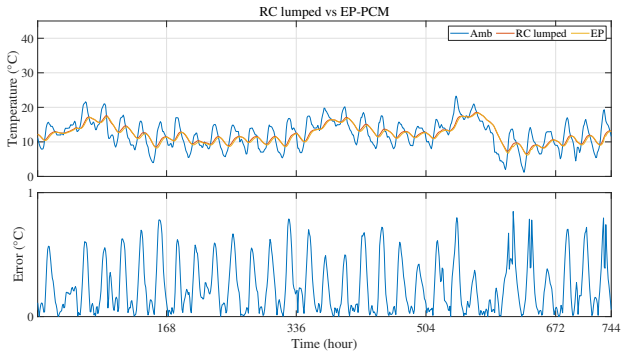
(b) Over a typical year in Sydney



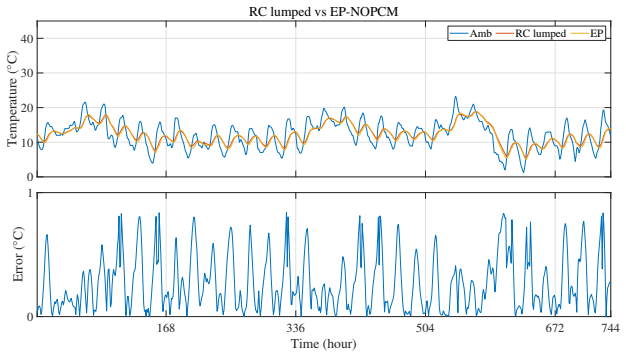
(c) Over a typical summer month in Sydney



(d) Over a typical summer month in Sydney



(e) Over a typical winter month in Sydney



(f) Over a typical winter month in Sydney

Figure 3.5: Benchmarking RC lumped model against EnergyPlus

To represent a thermal model with low-order ODEs; equations (3.18), (3.19) and (3.20), a few assumptions are made in this work. First, all elements of the building's envelope, such as roof, walls, and floor, are considered with the same materials. Second, all elements of the envelope are lumped as a single 2R1C circuit. Third, for further simplification, the PCM is located underneath the plasterboard layer of the envelope. Therefore two major heat capacitances of the model can be unified as a single nonlinear heat capacitance.

These simplifications are implemented mainly to reduce the computational task of the mathematical model of the building that results in less computational burden of the optimisation algorithm of MADP (Chapter 5, 7). Moreover, as it is utilised as a learning environment in the model-free RL method, a low-order representation of the thermal model will reduce the training process time (Chapters 6, and 8).

However, this simplified model of the building is benchmarked against a widely used EnenergyPlus software. To achieve that, an identical model is built in EnergyPlus software, and the indoor temperatures

of both the PCM building and the NOPCM building are monitored. The results show indoor temperature traces in both models match well. To quantify the discrepancy between the models, RMSE is used as a metric. The results of the RMSE for a yearly simulation, using outdoor air dry-bulb temperature of Sydney, shows that the maximum error is less than 0.8 $^{\circ}\mathrm{C}$ that is acceptable.

This chapter is based on [99], which I was the first author.

Chapter 4

Performance of PCM buildings with deadband HVAC control

Using phase change material (PCM) in lightweight buildings is highly beneficial in reducing the HVAC system demand or shifting the demand to off-peak hours. However, achieving an effective PCM performance is complex, and it is affected by several factors, including the climate, building parameters, and the properties of the PCM employed [38]. In this chapter, a wide range of scenarios are considered to investigate the impacts of geographical location, PCM melting point, duration of precooling and preheating, set points of the HVAC system, thickness and location of PCM in the buildings envelope, on the capability of the PCM in reducing or shifting the cooling and heating load. The study is conducted in the context of Australia, including five Australian cities: Brisbane, Sydney, Melbourne, Hobart and Perth. To our knowledge, this work is the first attempt to comprehensively analyse the impact of all PCM performance indices covering different climate conditions in Australia. Unlike in Chapter 3, we are not interested in an explicit mathematical representation of the building. In the context of this chapter, the optimal performance of PCM is simulation-based. In more detail, the optimal solution is selected based on analysing the outputs of limited case studies simulations, using three defined metrics. The HVAC system is controlled with a simple deadband relay. This is in contrast to Chapters 5, 6, 7, and 8, where the optimal performance of PCM or optimal scheduling of the HVAC system is decided based on an optimisation problem solution in a home energy management system (HEMS) with PCM. So, in this chapter, we use a more detailed EnergyPlus thermal model.

In summary, the contributions of this chapter are as follows:

- Explore the optimal performance of PCM in buildings in Australia by investigating the impact of different factors such as: climate conditions, PCM melting point, duration of PCM precooling and preheating with the HVAC system, HVAC system setting point, thickness of PCM layer, and PCM location in the envelope.

- We select the optimal performance of PCM, using three defined metrics of (i) annual HVAC demand, (ii) discomfort level, and (iii) shifting time of the HVAC demand.

- To compare the annual HVAC system consumption in the optimal PCM-integrated building with the HVAC demand in brick-wall and lightweight buildings, we simulate three types of (i) lightweight, (ii) brick-wall, and (iii) PCM-integrated buildings in EnergyPlus. The dimension and thermal resistance of the envelope in these three buildings are the same to study these three models' thermal storage capacity (thermal capacitance).

This chapter starts with presenting the details of each: lightweight, brick-wall, and PCM-integrated building. Next, we describe the assumptions that are made for this study. Afterwards, we discuss the results of the case studies simulations. And finally, the chapter concludes with a summary.

4.1 Implementation

In this section, the detailed models of the residential buildings considered are introduced, followed by some simplifying assumptions.

4.1.1 Three different buildings thermal model

In this work, three different building types are taken into account: (i) lightweight, (ii) brick-wall, and (iii) PCM-enhanced lightweight; they are modelled in the EnergyPlus software. The physical dimensions of the buildings are kept unchanged in all building types. The only difference between a lightweight building and a brick-wall building is the wall composition. In the PCM-integrated building, PCM was added to three different locations in the building: the roof, wall and the floor of the lightweight building. In changing the thickness of the PCM, the thickness of the gypsum board was adjusted accordingly so as to result in the same wall thermal resistance. Hence, in all three building types, the thermal resistance of the wall is same. The building is north facing with a total floor area of $150\,\mathrm{m}^2$. The house is modelled as one zone, a simple rectangular prism of $15\mathrm{m} \times 10\mathrm{m} \times 2.7\mathrm{m}$ dimension. The details of the building's fenestration is presented in Table 4.1.

Table 4.1: Fenestration details [3].

Element	**Description**	U $(\mathrm{W/m^2K})$	Area (m^2)
Window on North wall	Single glazing with Aluminium frame	7.01	6.50
Windows on East wall	Single glazing with Aluminium frame	7.01	9.56
Door on North wall	Wooden slab with wooden frame	2.61	1.78

The wall composition and materials of the lightweight building, brick-wall building and PCM-integrated building are shown in Table 4.2, Table 4.3 and Table 4.4 respectively. The PCM considered in this chapter is BioPCM with the properties detailed in Table 4.4. Other material properties such as thermal conductivity, density and specific heat which are required inputs in EnergyPlus are extracted from ASHRAE handbook [1].

Table 4.2: Wall composition and its material properties in lightweight buildings [1].

Wall material from outside to inside	λ ($\mathrm{W/mK}$)	ρ ($\mathrm{kg/m^3}$)	**c** ($\mathrm{J/kg\,K}$)
Cement board	0.25	1150	840
Glass-fiber batts	0.043	14	840
Gypsum board	0.16	640	1150

Table 4.3: Wall composition and its material properties in brick-wall buildings [1].

Wall material from outside to inside	λ ($\mathrm{W/mK}$)	ρ ($\mathrm{kg/m^3}$)	**c** ($\mathrm{J/kg\,K}$)
Brick (fired clay) layer	0.84	1760	800
Glass-fiber batts	0.043	14	840
Brick (fired clay) layer	0.84	1760	800

Table 4.4: Wall composition and its material properties in PCM-integrated buildings [1, 4].

Wall material from outside to inside	λ ($\mathrm{W/mK}$)	ρ ($\mathrm{kg/m^3}$)	**c** ($\mathrm{J/kg\,K}$)
Cement board	0.25	1150	840
Glass-fiber batts	0.043	14	840
BioPCM	0.2	860	1620
Gypsum board	0.16	640	1150

4.1.2 Assumptions

Some simplifying assumptions made in this chapter are given below:

- BioPCM is available in the form of a mat consisting of plastic blocks in which the PCM is encapsulated. In the simulations, we assumed PCM as a continuous layer.
- The same dates were selected for 'typical summer day', 'typical winter day', 'typical summer week'and 'typical winter week' for all five cities. Typical summer day: 1st February and typical winter day: 12th July. Typical summer week: 26th January to 1st February and typical winter week: 12th July to 18th July.
- Two working durations are assumed for the HVAC system in the summer season. The first working period starts from 10pm to 8am of the next day and the second working duration is from 8am to 6pm. For winter, only the working duration of 8am to 6pm is considered. Precooling and preheating of buildings are scheduled to be in the HVAC working duration.
- According to ASHRAE standard [1], the comfort level for winter is considered $20\,^{\circ}\mathrm{C}$-$23\,^{\circ}\mathrm{C}$ and for summer $23\,^{\circ}\mathrm{C}$-$26\,^{\circ}\mathrm{C}$. In this research, discomfort level is calculated based on the percentage of the hours in a year that the temperature is less than $20\,^{\circ}\mathrm{C}$ or higher than $26\,^{\circ}\mathrm{C}$.
- In our simulation, a complete mixing model for room air is selected, so the room temperature is uniform without any thermal stratification.
- The HVAC system that used in this chapter is a Packaged Terminal Heat Pump (PTHP). In the simulation, the HVAC setting is mostly at $20\,^{\circ}\mathrm{C}$ for heating in winter and $23\,^{\circ}\mathrm{C}$ for cooling in summer; otherwise its setting is stated.

Table 4.5: Demand shifting for $0.01\,\mathrm{m}$ PCM-integrated building with different melting point of PCM with various schedule of HVAC system (HVAC setting: $23\,^{\circ}\mathrm{C}$ in Hobart).

PCM melting point	**21 PCM**	**23 PCM**	**25 PCM**	**27 PCM**	**29 PCM**
Summer day HVAC (10pm to 8am)	51 min	25 min	No shift	No shift	5 min
Summer day HVAC (8am to 6pm)	No shift	No shift	No shift	No shift	6 min
Winter day HVAC (8am to 6pm)	57 min	36 min	33 min	30 min	27 min

- The weather data for the simulation is downloaded from the EnergyPlus software website[1]. This data is representative of a typical meteorological year. The average outdoor air dry bulb temperatures are $19.8\,^{\circ}\mathrm{C}$, $18.3\,^{\circ}\mathrm{C}$ $15\,^{\circ}\mathrm{C}$, $12.5\,^{\circ}\mathrm{C}$, and $18\,^{\circ}\mathrm{C}$, for Brisbane, Sydney, Melbourne, Hobart and Perth, respectively.

4.2 Results and discussion

As mentioned in Chapter 2, the melting point of a PCM is highly dependent on the geographical location of the building. In this chapter, three criteria were used for the selection of applicable melting temperature for each city: (i) annual HVAC demand, (ii) discomfort level, and (iii) shifting time of the HVAC demand. Based on the simulation results of the three mentioned criteria, the selected melting points of PCM for the five cities are: PCM25 for Brisbane, PCM23 for Sydney, PCM23 for Melbourne, PCM21 for Hobart and PCM25 for Perth. It is noticed that there is a correlation between the chosen PCM melting point and the average outdoor air dry bulb temperature ($^{\circ}\mathrm{C}$) of the city where PCM is used. PCM with lower melting points are more suitable for the cities with lower average temperatures. After selection of the PCM type (melting point) for each city, different scenarios are implemented for each of the five cities and the five major findings are given below: Firstly, the results show that in all cities, precooling from 10pm to 8am is more effective than precooling from 8am to 6pm in shifting the demand. In other words, changing the working hours of the HVAC system from 10pm-8am to 8am-6pm will result in a lower or no shift in the demand. The simulation results for Hobart summarised in Table 4.5 shows that the highest shift in both cooling and heating demands occur when PCM21 is applied. Precooling the PCM from 10pm-8am in the summer season in Hobart, shifts the cooling demand by 51 minutes.

If the internal temperature of the buildings is plotted for summer and winter day in Hobart (as shown in Fig. 4.1), it visualises the findings illustrated in Table 4.1. Fig. 4.1a shows that precooling the building to $23\,^{\circ}\mathrm{C}$ from 10pm to 8am (PCM21) causes a delay in exceeding the comfort level of $26\,^{\circ}\mathrm{C}$, when compared to the case without PCM. This is an approximately 51 minutes shift in cooling demand. But using PCM21 for precooling from 8am-6pm did not cause any shift in the demand (see Fig. 4.1b). Fig. 4.1c shows that for Hobart, using PCM21 yields the highest shift in the demand compared to other PCM types (as presented in Table 4.5).

Secondly, the effect of varying the setting point of the HVAC system was examined in this work. Results show that changing the cooling system setting from $23\,^{\circ}\mathrm{C}$ to $22\,^{\circ}\mathrm{C}$, increased the demand shift

[1] https://energyplus.net/weather

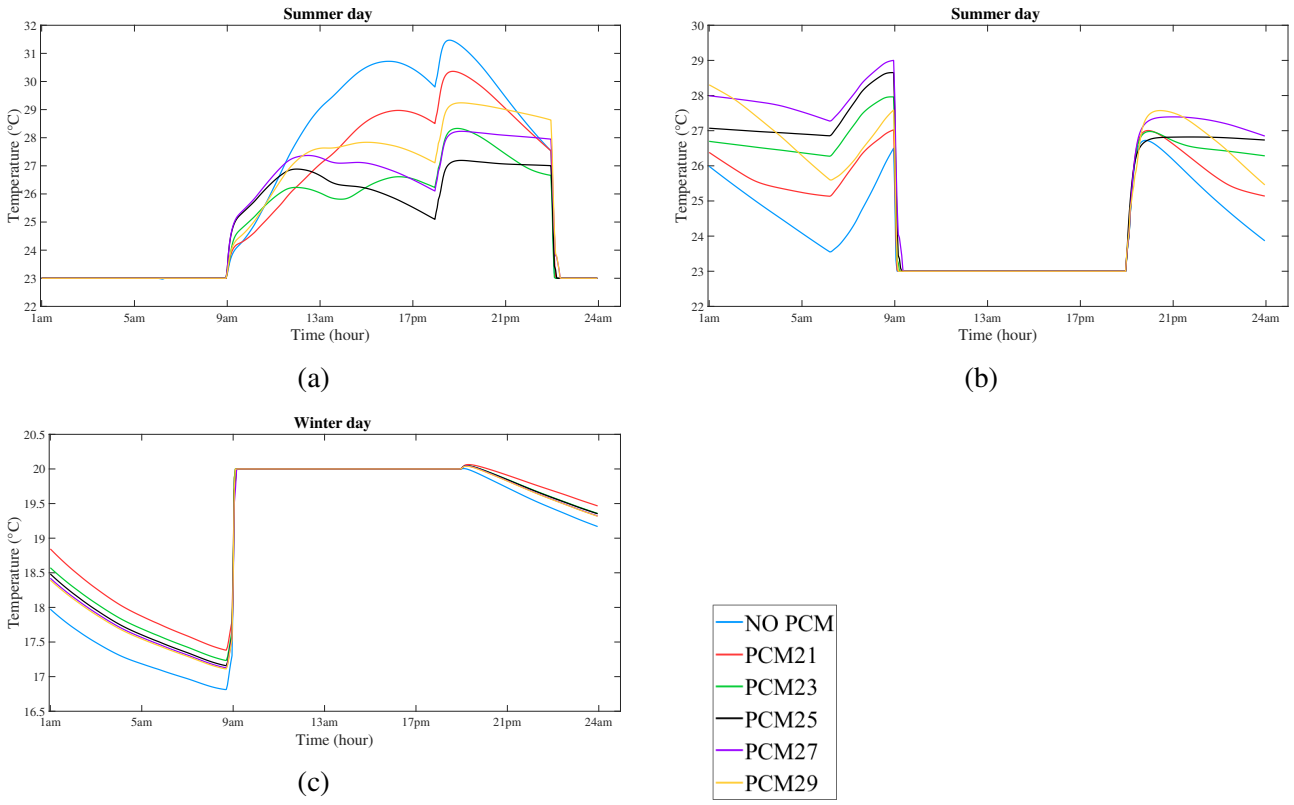


Figure 4.1: (a) Internal temperature in the typical summer day of Hobart using different PCM types (with HVAC working duration: 10pm to 8am). (b) Internal temperature in the typical summer day of Hobart using different PCM types (with HVAC working duration: 8am to 6pm). (c) Internal temperature in the typical winter day of Hobart using different types of PCM (with HVAC working duration: 8am to 6pm).

time for all cities except Brisbane while increasing the setting point of the cooling system to $24\,^{\circ}\mathrm{C}$ reduced the demand shift time in all cities. The maximum shift in demand observed in all cities with the adoption of PCM is: 9 minutes for Brisbane with an HVAC setting point of $23\,^{\circ}\mathrm{C}$; 3 minutes, 103 minutes and 60 minutes respectively for Sydney, Hobart and Melbourne with HVAC setting point of $22\,^{\circ}\mathrm{C}$. In Perth no shift in the demand could be seen.

Thirdly, the inclusion of the PCM in the building resulted in the reduction in annual and weekly HVAC demand when compared to the building without PCM. Fig. 4.2a and Fig. 4.3a illustrates this finding for Hobart city. Also, three different thicknesses (of $0.01\,\mathrm{m}$, $0.015\,\mathrm{m}$ and $0.02\,\mathrm{m}$) PCM were analysed in the simulations; the results show that applying $0.02\,\mathrm{m}$ thick PCM will result in the highest reduction in annual and weekly heating and cooling demand of the building for all cities (Fig. 4.2b and Fig. 4.3b illustrate the results for Hobart.). Integration of $0.02\,\mathrm{m}$ thickness of PCM layer, results in $6.9\,\%$, $6.8\,\%$, $16.7\,\%$, $21.8\,\%$, and $12.3\,\%$ reduction in the annual HVAC demand consumption, compared to building without PCM. However, this is with the assumption that the HVAC system is operating 24 hours a day for a whole year to keep the temperature in the comfort range of $20\,^{\circ}\mathrm{C}$-$23\,^{\circ}\mathrm{C}$ for winter and for summer $23\,^{\circ}\mathrm{C}$-$26\,^{\circ}\mathrm{C}$.

Among all cities, increasing the thickness of PCM from $0.01\,\mathrm{m}$ to $0.015\,\mathrm{m}$ reduced the annual HVAC system demand by a maximum of $2.1\,\%$ and a minimum of $0.6\,\%$; and increasing thickness of PCM from $0.015\,\mathrm{m}$ to $0.02\,\mathrm{m}$ reduced the annual HVAC system demand by a maximum of $2.4\,\%$ and a

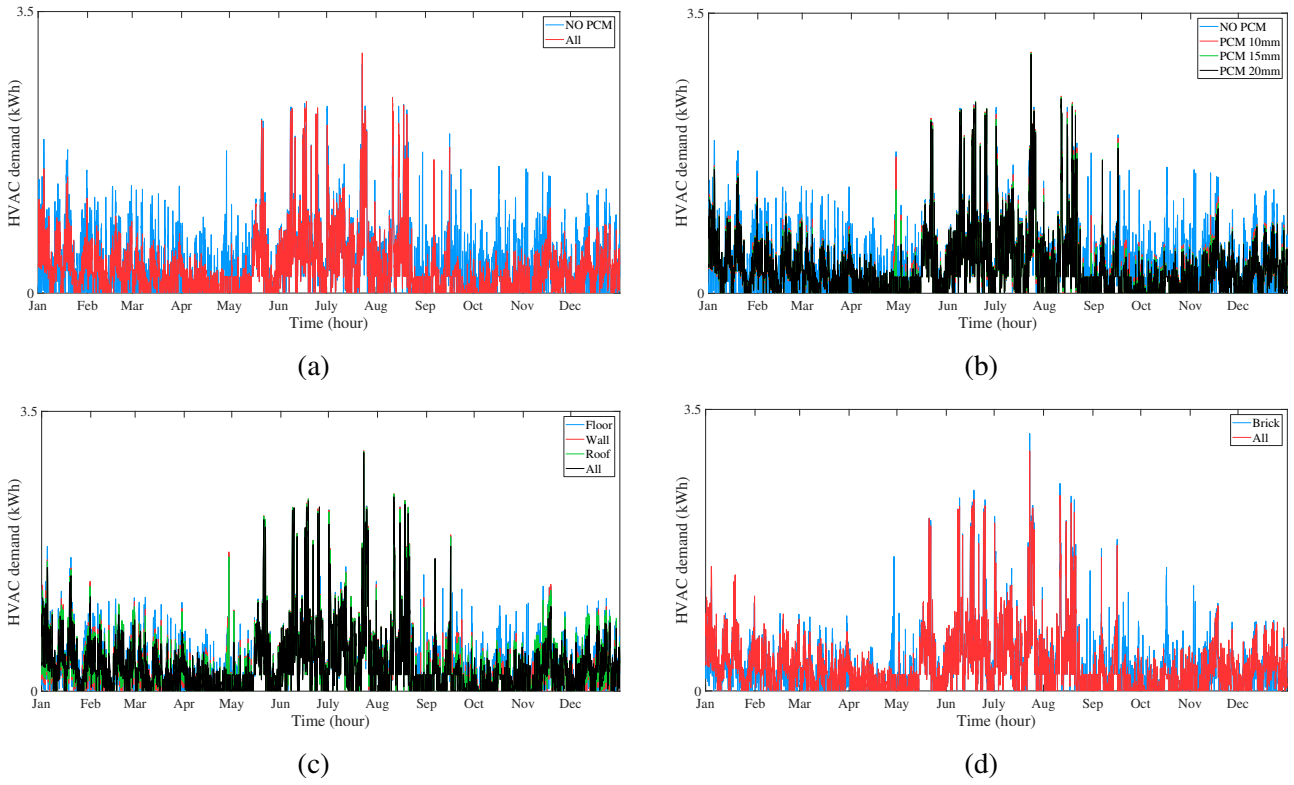


Figure 4.2: (a) Annual HVAC demand (kWh) of building with $0.01\,\mathrm{m}$ thick PCM compared to building without PCM in Hobart. (b) Annual HVAC demand (kWh) of PCM-integrated building with PCM of different thicknesses compared to building without PCM in Hobart. (c) Annual HVAC demand (kWh) of PCM-integrated building with $0.02\,\mathrm{m}$ thick PCM in three different locations (roof, walls and floor) compared to building without PCM in Hobart. (d) Annual HVAC demand (kWh) of PCM-integrated building with $0.02\,\mathrm{m}$ thick PCM in all three locations compared to brick-wall building in Hobart.

minimum of $0.4\,\%$. Finally, changing the PCM thickness from $0.01\,\mathrm{m}$ to $0.02\,\mathrm{m}$ decreases the annual HVAC demand by a maximum of $4.5\,\%$ and a minimum of $0.95\,\%$.

Fourthly, the effects of using PCM at three different locations (roof, walls and floor) were examined separately and also with PCM in all three locations. For the highest reduction in annual and weekly energy demand, it is best to use the $0.02\,\mathrm{m}$ thick PCM in all three locations due to an increase in the overall thermal capacity of the envelope. When there is an economic constraint, using PCM in the roof is an optimal way of integrating PCM into the building. The simulations results show that by integration of PCM in the roof of the building, we can achieve a minimum $66\,\%$ (in Melbourne), up to maximum $80.7\,\%$ (in Hobart) of the reduction in the annual HVAC consumption that we can obtain by applying PCM in the whole envelope (roof, walls and floor). The second efficient option is installing the PCM in all the walls. This gives almost the same amount of reduction in heating and cooling demand with that of the roof. These findings are shown in Fig. 4.2c and Fig. 4.3c.

And finally, we compared the results of the optimal PCM-integrated building (in this case the building with $0.02\,\mathrm{m}$ PCM in all roof, walls and floor) to that of the brick-wall building, for all cities. The inclusion of PCM resulted in an annual HVAC demand reduction of $7.3\,\%$ and $1.4\,\%$ in Hobart and Melbourne, respectively. However, in Brisbane, Sydney, and Perth, the brick-wall building has about $6.6\,\%$, $9.5\,\%$, and $6.3\,\%$ less HVAC demand (annual and weekly) compared to the PCM-enhanced

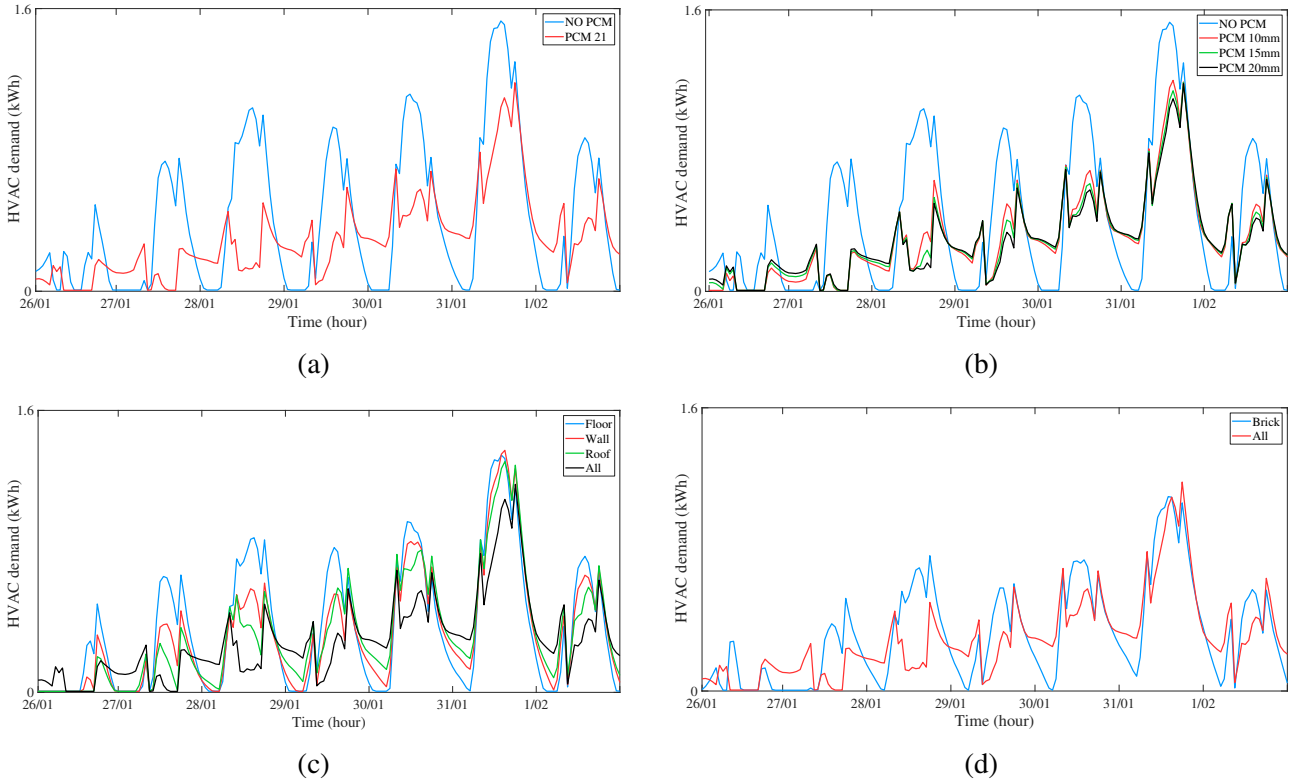


Figure 4.3: (a) HVAC demand (kWh) of building with $0.01\,\mathrm{m}$ thick PCM compared to buildings without PCM for a typical summer week (26/01 to 1/02) in Hobart. (b) HVAC demand (kWh) of PCM-integrated buildings with PCM of different thicknesses compared to building without PCM for a typical summer week (26/01 to 1/02) in Hobart. (c) HVAC demand (kWh) of PCM-integrated building with $0.02\,\mathrm{m}$ thick PCM in three different locations (roof, walls and floor) compared to building without PCM for typical summer week (26/01 to 1/02) in Hobart. (d) HVAC demand (kWh) of PCM-integrated building with $0.02\,\mathrm{m}$ thick PCM in all three locations compared to brick-wall building for typical summer week (26/01 to 1/02) in Hobart.

building. The reason is in Hobart and Melbourne, the outdoor air dry bulb temperatures have an average higher daily fluctuations during the year compared to the other cities. Therefore, during the year, there are higher number of temperatures which are above the PCM melting point or below the PCM freezing point in these cities means the PCM had more efficient cycles of melting and freezing which resulted in less HVAC system demand. Fig. 4.2d and Fig. 4.3d shows the comparison between optimal PCM-integrated buildings with brick wall building for Hobart.

In all, simulation results confirm that the PCM-building is operating more efficiently in Hobart and Melbourne compared to the other three cities. Using the precooling strategy in a typical summer day resulted in a maximum shifting of demand by 103 minutes and 60 minutes in Hobart and Melbourne respectively. And in a typical winter day, it resulted in a demand shifting of 57 minutes and 54 minutes for Hobart and Melbourne respectively. With a HEMS in place, it utilizes these shifts and schedules the HVAC system to move demand from peak periods to off-peak periods. In terms of HVAC demand reduction, the PCM-integrated building yielded $21.8\,\%$ and $16.7\,\%$ reduction in Hobart and Melbourne, respectively, assuming that the HVAC system is operating 24 hours in a year with a setting point of $20\,^{\circ}\mathrm{C}$ for winter and $23\,^{\circ}\mathrm{C}$ for summer.

4.3 Summary

This chapter has comprehensively investigated the impact of different factors of geographical location, PCM melting point, duration of precooling and preheating, setting points of the HVAC system, thickness and location of PCM on PCM performance. A few case studies are defined in Australia considering five cities of Brisbane, Sydney, Melbourne, Hobart and Perth. Three metrics of (i) annual HVAC demand, (ii) discomfort level, and (iii) shifting time of the HVAC demand are used for the selection of optimal PCM performance among different defined case studies. Moreover, the optimal configuration of the PCM building is compared with the brick-wall building. In more detail, we consider three types of buildings: (i) lightweight, (ii) brick-wall, and (iii) PCM-integrated lightweight buildings. To reflect the true impact of PCM, compared to high inertia material like brick and lightweight envelope, we consider the same physical dimension for all three types of building. Moreover, we adjust the thickness of the envelope compositions so that all three models have the same total thermal resistance for their envelope.

The summary of key findings are as follows:

- We realised that cities with lower average outdoor air dry bulb temperatures have lower optimal PCM melting point temperatures. Based on our simulations, the selected melting points of PCM for the five cities are PCM25 for Brisbane, PCM23 for Sydney, PCM23 for Melbourne, PCM21 for Hobart and PCM25 for Perth.

- The results show that the time of precooling and preheating of PCM impacts shifting the demand. In particular, we visualised that precooling of the building in the summer season from 10pm to 8am, is more effective than precooling from 8am to 6pm. Moreover, the impact of the setpoint of the HVAC system on the HVAC demand shift is also examined. We observe that for Brisbane, using the HVAC system with a setpoint of $23\,^{\circ}\mathrm{C}$ is more effective in shifting the demand. However, for the cities of Sydney, Melbourne and Hobart, setting the HVAC system setpoint to $22\,^{\circ}\mathrm{C}$ is more effective in shifting the demand. In Perth, we have not seen any shift in the demand. The maximum recorded demand shift in the simulations was: 9 minutes, 3 minutes, 60 minutes and 103 minutes in the cities of Brisbane, Sydney, Melbourne and Hobart, respectively.

- Moreover, we analysed the impact of PCM layer thickness in reducing the heating and cooling demand. PCM with $0.02\,\mathrm{m}$ thickness is more beneficial in reducing annual HVAC demand compared to $0.015\,\mathrm{m}$, and $0.01\,\mathrm{m}$. The integration of a $0.02\,\mathrm{m}$ thick PCM in the roof, wall and floor of the building compared to the lightweight building, resulting in a reduction in annual HVAC system demand as high as $21.8\,\%$ in Melbourne and as low as $6.8\,\%$ reduction in Sydney.

- Moreover, we witnessed that locating the PCM layer at all different envelope locations such as roof, walls and floor building results in the highest reduction in annual HVAC consumption. However, when there is an economic constraint, integrating PCM into the roof is more effective in reducing the HVAC demand than placing it in the walls or floor. Based on our simulations,

integrating PCM only to the roof, we can achieve $65.9\,\%$ to $80.7\,\%$of reduction in the annual HVAC consumption that we can achieve by integrating the PCM in all locations of the envelope.

- Based on our simulations, the adoption of PCM technology in the buildings in contrast to high thermal inertia constructions is more beneficial in the cities of Melbourne and Hobart (an annual HVAC demand reduction of $7.3\,\%$ and $1.4\,\%$in Hobart and Melbourne, respectively). However, in other cities, the annual HVAC consumption in the brick-wall building is on average $7.5\,\%$ less than the PCM-integrated building.

The findings in this chapter show potential of PCM to use as a storage system for demand response. However, capturing the optimal performance of PCM involves many factors. Selection of PCM performance optimal based on all factors is a time-consuming task despite defining a limited number of case studies. This motivates us to cast the problem as an optimisation problem in the HEMS that is discussed in more detail in the following chapters.

Many of the methods and results discussed in this chapter are published in [45].

This chapter is based on [45], which I was the first author.

Chapter 5

Computationally-Efficient Energy Management of PCM-buildings using Multi-time scale Approximate Dynamic Programming (MADP)

In Chapter 2, we discussed that one of the demand response (DR) resources is using thermal inertia of building. In this thesis, our focus is on lightweight dwellings that dominate the residential building stock in Australia. This type of buildings has low thermal inertia and correspondingly high HVAC costs. However, they are the most suitable for applying *phase change materials* (PCM). However, as demonstrated in Chapter 4, the optimal performance of PCM as an energy storage system depends on many factors. Moreover, as discussed in Chapter 2, passive application of PCM can not always unleash the full PCM potential. Therefore, we use PCM in conjunction with active mechanical devices, which is the HVAC system. As demonstrated in Chapter 4, although determining PCM melting point, its thickness, and PCM location in the envelope is important in the initial stage of PCM application. However, it is critical to exploit PCM storage capacity efficiently during its long lifetime (almost 80 years).

In more detail, to use PCM as a DR resource, PCM needs to be precooled or preheated (depending on the season) by the HVAC system during periods of shoulder or off-peak energy prices. This task can be cast as an optimal HVAC scheduling problem, to minimise electricity cost while maintaining the indoor temperature within a desired comfort range.

In the existing literature, this type of optimization problem is classified as a *home energy management system* (HEMS) problem [15, 100, 101]. In spite of the ample literature on the use of PCM for improving the thermal performance of buildings [3, 34, 36, 42, 44, 45, 102], there is a palpable lack of understanding on how to integrate PCM into HEMS, where the non-linear nature of its energy storage can be exploited using suitable scheduling methods.

To bridge this gap, and in contrast to much of the literature on HEMS [15, 100, 101], we consider

HEMS that consists of an HVAC system as a controllable device and a PCM layer as an energy storage system. To date, most HEMS optimization problems are solved using *linear programming* (LP) and *mixed-integer linear programming* (MILP). However, these methods cannot be used to solve nonlinear optimization problems, which phase-change characteristics impart. Other methods widely used to solve the HEM problems are heuristic methods, such as *particle swarm optimization* (PSO) and *genetic algorithms* (GA). The downside of using these methods is that the solution may end up in a local optimum instead of the global optimum. Therefore, the solution quality is questionable [103]. More importantly, PSO and GA are black-box optimization routines. In our specific problem, they rely on the huge computational task of solving the initial value problems associated with the ordinary differential equations that govern the building's thermal behavior. In this sense, they provide no benefit over using principled optimization methods like *dynamic programming* (DP) [20, 100].

The state-of-the-art algorithm for dealing with the nonlinear characteristic of PCM is *dynamic programming* (DP) [67]. A problem that has a sequentially-separable structure, or has the *Markov property*[1], can be formulated as a *Markov decision process* (MDP) to be solved by DP. In more detail, for our specific HEMS problem, the MDP consists of accumulated instantaneous costs or rewards over a decision horizon. The main operator in DP is a *state-action value function (Q-function)*, which is formed by adding the expected future cost of following a *policy* (in this problem, specific on/off sequence of HVAC system) given state transition probabilities. Importantly, the objective is to find the minimum value function over the time horizon of the problem. In doing so, the *value iteration* (VI) algorithm is employed in DP, which computes the minimum value function in a backward fashion using the Bellman optimality condition[2] [20]. Using this, an optimal policy can be extracted by tracing back the states with minimum state-action value functions over the time horizon of the problem.

However, DP has a limitation that is well-known as the *curse of dimensionality*, which means VI becomes computationally intractable when the time horizon of the problem, number of state variables or number of controllable devices grows. More specifically, solving this problem with a pure dynamic programming algorithm entails a huge computational burden that makes impossible its application in the current context of DR.

Given the limitations of existing approaches, as one of the major contributions of this thesis, we develop *computationally-efficient multi-timescale approximate dynamic programming* (MADP) to optimise the schedule of an electrical HVAC system in HEMS with PCM. As a first step towards our computationally efficient method, we discretise the continuous state-space of the problem. However, because we require a relatively fine-grained discretisation, we are left with a very large state-space. To reduce the computation further, we employ a multi-time scale MDP, in which decisions are made at different discrete timescales [22]. Specifically, rather than solving the original MDP as one monolithic problem, we solve several smaller MDPs that are connected successively together to form the original MDP. And finally, to improve the computational performance, we address the bottleneck of the state

[1] A process has a *Markov property* if the future state depends only on the current state, not on the sequence of events that preceded it.

[2] An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

transition function ((3.18), (3.19) and (3.20)), given by a thermal model of the building. To avoid timely online solving of ordinary differential equations (ODEs), we trained an *artificial neural network* (ANN) to use as an approximation of the system of ODEs. The ANN maps the outdoor temperature, the indoor temperature of the building, and the HVAC heat flow, to the next time step's indoor temperature. The ANN is trained offline, therefore significantly reducing the runtime of the VI algorithm. The results demonstrate that our proposed method of MADP speeds up the computational time of the VI algorithm up to 157,600 times (over a day time horizon with an hour time step) compared to the direct application of DP while maintain the acceptable quality of the solution.

Beyond this, this chapter advances state of the art in the following ways:

1. We derive a novel and computationally-efficient optimisation method for nonlinear online scheduling of HVAC systems in HEMS with PCM. The method exploits several techniques of (i) state-space approximate (ii) multi-time scale MDP, and (iii) ANN function approximation from artificial intelligence in one framework.

2. We demonstrate the performance of the method on an HVAC scheduling problem over a year time horizon using Sydney climate conditions.

3. Each step of the proposed methodology is evaluated individually by defining appropriate metrics or mathematical proof.

4. The proposed method gives substantial computational speed-ups as high as 157,600 times faster than DP.

5. The method described in this chapter can be implemented on current smart meters and IoT gateway devices, such as those built on Raspberry Pi boards.

This chapter progresses as follows: first, the optimisation problem of the PCM-building is described. Then using the ODEs (3.18), (3.19), and (3.20) of the RC lumped model that is developed in Chapter 3, we formulate the problem as an MDP and explain how VI is used to solve it. Next, as the main technical contributions of this chapter, the proposed MADP is derived. Afterwards, the method is used for optimal scheduling of the HVAC system in the typical PCM-building in Sydney over a year time horizon. Next, using the simulations results and three defined metrics, the proposed method of MADP is evaluated. And finally, this chapter concludes with a summary of the chapter.

5.1 Markov decision process in PCM buildings

In this section, we formulate the HVAC-PCM optimization problem as an MDP, using differential equations (3.18), (3.19) and (3.20), of the thermal model as transition functions. Then we show how dynamic programming, specifically the VI algorithm, can be used solve the optimization problem, as a precursor to our ADP method.

An MDP comprises a *state-space*, $(s \in \mathcal{S})$, a *decision-space*, $(x \in \mathcal{X})$, *transition functions* and *contribution functions*. Let $k = \{1, \ldots, K\}$ denote a time step of one hour. A state variable, $s_k \in \mathcal{S}$, contains the information that is necessary and sufficient to make the decisions and compute costs, rewards and transitions. The decision variable, $x_k \in \mathcal{X}$, is an action that results in a transition from one state to another in a sequence over the decision horizon. Finally, random effects are in general, used to represent chance exogenous information, such as weather conditions or inhabitants' behavioural patterns [15]. However, for simplicity and because we focus on the non-linear characteristics of PCM, in this work, the problem is treated as deterministic. Thus, the form of the HVAC-PCM MDP is given by

$$\begin{aligned} \min_{\pi} \mathbb{E} \left\{ \sum_{k=0}^{K} C_k(s_k, x_k = \pi(s_k)) \right\} \\ \text{s.t.} \quad \text{thermal comfort constraints, and} \\ \text{thermal energy balance constraints,} \end{aligned} \tag{5.1}$$

where $\pi : \mathcal{S} \rightarrow \mathcal{X}$ is a policy, i.e. a sequence of actions taken to move from each state to the next state over the whole time horizon. In this work, a policy is a sequence of on/off status of the HVAC system over a defined time horizon.

The contribution function is $C_k(s_k, x_k)$, which is the cost incurred at a given time step k that accumulates over time [15]. For our specific optimization problem, the cost consists of the electricity cost and the discomfort cost:

$$C_k(s_k, x_k) = \lambda c_{\mathrm{g},k} p_k^+ + (1 - \lambda) \left(|T_{\mathrm{room},k} - T_{\mathrm{s}}| \right). \tag{5.2}$$

To balance the two cost components, the contribution function includes a weighting factor λ, applied to the electricity cost, with $(1 - \lambda)$ applied to the penalty for deviating from the desired HVAC setpoint T_{s}. We assume a reverse-cycle HVAC system able to operate both in a heating and cooling mode. The setpoint T_{s} for the two modes is assumed $20\,^{\circ}\mathrm{C}$ and $23\,^{\circ}\mathrm{C}$, respectively. The electricity cost of the HVAC system is the electricity time-of-use tariff[3], $(c_{\mathrm{g},k})$, multiplied by the energy used to run the HVAC system, p_k^+. The following equation shows how p_k^+ relates to $\dot{Q}_{\mathrm{HVAC}}$ in (3.19):

$$p_k^+ = \frac{\dot{Q}_{\mathrm{HVAC}}}{COP} \tag{5.3}$$

where COP stands for the *coefficient of performance*, which indicates the efficacy of the HVAC system in moving thermal energy in and out of the building[4]. In this work, *COP* value is mentioned when it is required.

To describe the transition functions in (5.1), let $s_{k+1} = s^M\left(s_k, x_k\right)$, define the state transition from time step k to the next time step, $k + 1$, where s^M is the underlying mathematical model of the studied

[3]EnergyAustralia Residential Energy Price Fact Sheet for Ausgrid Distribution Zone (https://www.energyaustralia.com.au, release date: 30-Jul-2018).

[4]Note that the *COP* is defined as the ratio between the thermal energy removed and the electrical energy used for that. Therefore, this is not *efficiency*, so the *COP* can take values greater than 1 [1].

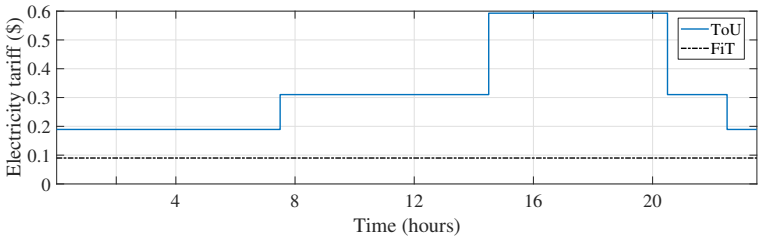


Figure 5.1: The electricity time-of-use tariff and the feed-in-tariffs.

system [15]. In this problem, the system model is the thermal model of the building, so the MDP transition functions are given by equations (3.18), (3.19) and (3.20).

Cost function (5.2) only considers the short-term cost that results from the decision that is taken at each time step. Building on this, the problem is solved by computing the *Q-function*: $Q^{\pi}(s_k, a_k)$, which is the expected future discounted reward of taking action a_k in state s_k and then following the policy, π, in subsequent states. The Q-function is given by:

$$Q^{\pi}(s_k, a_k) = \sum_{s' \in \mathcal{S}} \Pr(s'|s_k, a_k) \left[C_k(s_k, a_k, s') + Q^{\pi}(s', \pi(s')) \right], \tag{5.4}$$

where $\Pr(s'|s_k, a_k)$ is the probability of making a transition from state s' to s if we take action a. However, because the system model s^M is a deterministic function, we have that:

$$\Pr(s'|s_k, a_k) = \begin{cases} 1 & \text{if } s^M(s_k, a_k) = s', \\ 0 & \text{otherwise,} \end{cases}$$

The expression in (5.4) is a recursive reformulation of the objective function. Thus, in general, *Bellman's optimality condition* states that the optimal Q-function is given by:

$$Q_k{}^{\pi^*}(s_k, a_k) = \max_{a_k \in \mathcal{A}_k} \left(C_k(s_k, a_k) + \mathbb{E}\left[Q^{\pi^*}_{k+1}(s', \pi^*(s')|s_k, a_k \right] \right), \tag{5.5}$$

where π^* is an optimal policy. To find π^*, we need to solve (5.5) for each state.

The value iteration algorithm computes (5.5) for each state by backward induction, which means it starts from states at time step K (i.e. a terminal state in the decision horizon) and moves backwards to the first time step ($k = 1$). To start from time step K, we initialize the expected state-action value functions ($Q^{\pi}(s', \pi(s'))$ in (5.4)) for all of the terminal states. For this specific problem, the state is the indoor temperature of the building. The expected value functions of states within the comfort temperature range are set to a lower value, while states with a value out of the acceptable range are given high expected values such that they will not be selected in an optimal policy. Then, for each time step $k = K-1, K-2, \ldots, 2, 1$, adding the instantaneous cost (5.2) to the minimum expected value function of the subsequent state gives the optimal value function (5.5) of that state. Once this process reaches $k = 1$, the computed value functions for all time step are optimal. Then, by tracing a minimum value-function path forward for a given time horizon, the optimal policy is extracted [15].

Although in the VI algorithm we only consider one state-variable of indoor temperature and use a high-performance computer cluster, the runtime of the algorithm is very long; specifically, it takes

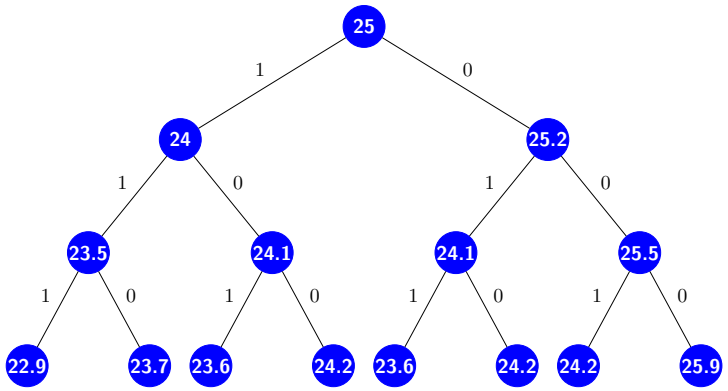


Figure 5.2: Curse of dimensionality: increase of time horizon from the initial state of $25\,^{\circ}\mathrm{C}$ to three time steps. Action 0, and 1 on the edges shows the on/off status of the HVAC system.

almost nine days for a 24 hour time horizon with hourly time steps. Fig. 5.2 illustrates the state and action space for one state variable (indoor temperature) over three time steps starting from the initial indoor temperature of $25\,^{\circ}\mathrm{C}$. The on/off action of the HVAC system is shown as $0/1$ on the edges. We witness an exponential growth of states and action space, which is known as the *curse of dimensionality*. For this specific problem, the computational bottleneck of the VI algorithm is solving the MDP transition functions. In more detail, the algorithm solves the differential equations (3.18), (3.19) and (3.20) at each time step for each on/off status of the HVAC system over a defined time horizon. Given this shortcoming, we propose a method to overcome the computational burden of the VI algorithm.

5.2 Methodology

In this section, we describe our methodology in three steps, namely: a state-space approximation, multi-timescale MDP, and neural network transition function approximators.

5.2.1 State-space approximation

To begin, we describe the state-space approximation as a ground model to reduce the computational burden of the HVAC-PCM MDP. For clarity, we call the MDP that uses equations (3.18), (3.19) and (3.20), without any change, the *exact model*; the state and state-space used in the exact model are called *exact state* and *exact state-space*, respectively. The corresponding terms in the approximated methodology are called, respectively, the *approximate model*, *approximated state* and *approximated state-space*.

For the approximated model, at each time step, the outputs of the transition functions (equations (3.18), (3.19) and (3.20)) are rounded to the nearest multiple of 0.1. In this work, we assume the comfort temperature range is between $20\,^{\circ}\mathrm{C}$ and $26\,^{\circ}\mathrm{C}$. Since we highly penalize the states (temperatures) that are out of comfort range, any state in the desired state-space has a value between $20\,^{\circ}\mathrm{C}$ and $26\,^{\circ}\mathrm{C}$ with a $0.1\,^{\circ}\mathrm{C}$ discretization.

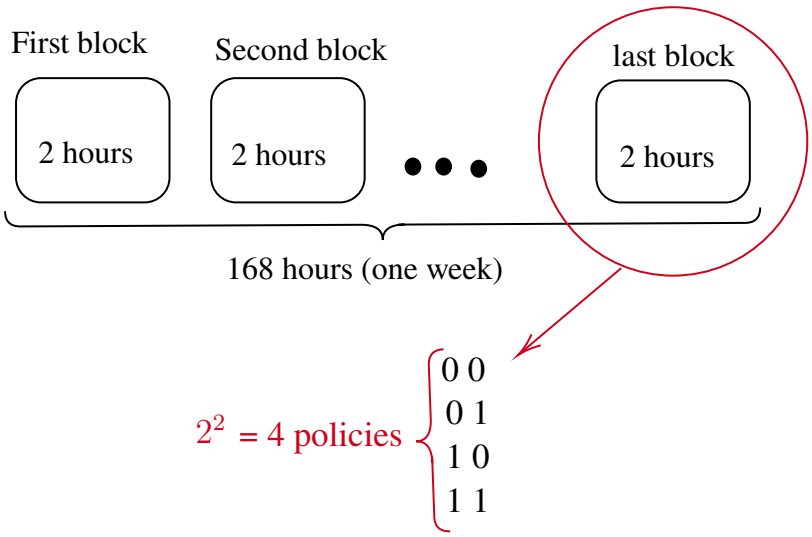


Figure 5.3: Multi-timescale Markov decision processes.

In contrast, in the exact model, a state can have any value in the continuous interval between 20 °C and 26 °C. In practice, only reachable states in the exact state-space are evaluated by VI. Even so, computing over the approximated state-space compared to the reachable elements of the exact state-space returns a huge reduction in computation time of the VI algorithm because a greatly reduced number of state transitions is computed. However, how this approximation affects the quality of the solution needs to be examined. Therefore, in the first part of Section 5.3, we assess the state-space approximation performance. Building on this, the approximate MDP model described is used as a ground MDP for the multi-timescale MDP and neural network transition function approximators discussed next.

5.2.2 Multi-timescale Markov decision processes

Applying a multi-timescale abstraction significantly reduces the computational burden of the energy management optimization problem. Building on the approximated state-space introduced above, we divide the time horizon of the problem into *blocks*, each consisting of two time steps (hours), as illustrated in Fig. 5.3. We can formulate each block as a separate MDP; therefore, we solve a few successive block-MDPs using VI to find the optimal policy over the whole time horizon. We present this schematically in, assuming For reference, we denote the multi-timescale method Algorithm 1 (Alg 1).

In more detail, VI is first applied to the last block-MDP (Lines 1-11 of the Alg 1 pseudocode). We set the corresponding value functions in the last time step to zero for the states that have a value within the desired comfort range, and assign a high value for the states with the values out of the comfort range (Lines 5-9). To exploit the advantage of the approximated state-space, we run the VI algorithm for 61 step discretization (20 to 26 in 0.1 intervals in Lines 1 and 13) for each block-MDP except the first one. We save all the optimal value functions that correspond to each of the 61 initial points in look-up tables (Lines 10 and 23). Before running VI on the remaining block-MDPs, we update the initial value functions by replacing the corresponding value function of the current state, by finding

Algorithm 1 : Multi-timescale algorithm (ALG 1)

L: length of each block
T_1: lower bound of desired temperature range
T_2: upper bound of desired temperature range
d: discretization step
T_0: fix initial temperature

▷ Value iteration (VI) of the last block (Blk_M)

1: **for** $T_1, ..., T_2$ with d step discretization **do**
2: **for** all 2^{L} combinations of the action-space **do**
3: calculate states using equations equations (3.18), (3.19) and (3.20)
4: **end for**
5: **if** $T_1 \leq s_{\mathrm{Blk_{M,colL}}} \leq T_2$ **then**
6: Initialize $v_{Blk_{M,colL}}$ to zero vector
7: **else**
8: Initialize $v_{\mathrm{Blk_{M,colL}}}$ to infinity vector.
9: **end if**
10: Execute the VI and store final value function in $v_{\mathrm{final,Blk_M}}$.
11: **end for**

▷ Value iteration of Blocks 2 to $M-1$ (Blk_2 to Blk_{M-1}).

12: **for** $M-1 : ... : 2$ **do**
13: **for** $T_1, ..., T_2$ with d step discretization **do**
14: **for** all 2^{L} combinations of the action-space **do**
15: calculate states using equations equations (3.18), (3.19) and (3.20)
16: **end for**
17: **if** $T_1 \leq s_{\mathrm{Blk_{M-1,colL}}} \leq T_2$ **then**
18: for any state for which $s_{\mathrm{Blk_{M-1,colL}}} = s_{\mathrm{Blk_{M,col1}}}$,
19: Initialize $v_{\mathrm{Blk_{M-1,colL}}}$ with the corresponding $v_{\mathrm{final,Blk_M}}$.
20: **else**
21: Initialize $v_{\mathrm{Blk_{M-1,colL}}}$ with infinity vector
22: **end if**
23: Execute the VI and store final value function in $v_{\mathrm{final,Blk_{M-1}}}$.
24: **end for**
25: **end for**

▷ Value iteration of the first block (Blk_1).

26: Set T_0.
27: **for** all 2^{L} combinations of the action-space **do**
28: calculate states using equations (3.18), (3.19) and (3.20)
29: **end for**
30: **if** $T_1 \leq s_{\mathrm{Blk_{1,colL}}} \leq T_2$ **then**
31: for any state for which $s_{\mathrm{Blk_{1,colL}}} = s_{\mathrm{Blk_{2,col1}}}$,
32: Initialize $v_{\mathrm{Blk_{1,colL}}}$ with the corresponding $v_{\mathrm{final,Blk_2}}$.
33: **else**
34: Initialize $v_{\mathrm{Blk_{1,colL}}}$ with infinity vector
35: **end if**
36: Execute the VI and store final value function in $v_{\mathrm{final,total}}$.

the initial states that have the same value as the current state and replacing the corresponding value functions as the initial value functions of the current block-MDP (Lines 18-19 and Lines 31-32). This process repeats by backward recursion until the first block-MDP. To find a solution for the optimization problem over a defined time horizon, we need to fix either the initial or the final temperature. In this work, we provide the ALG 1 with a fixed initial temperature (Line 26). Therefore, we have only one VI to run for the first block-MDP.

Comparing the results of ALG 1 with a one-block MDP model shows that both methods converge to the exact same solution. This corroborates with Sutton's result that n-block MDPs act exactly the same as the corresponding one-block MDP [22]. Our simulations, with hourly time steps over a year, show that using ALG 1 reduces the computational burden of finding the optimal policy by a factor of approximately 13,100.

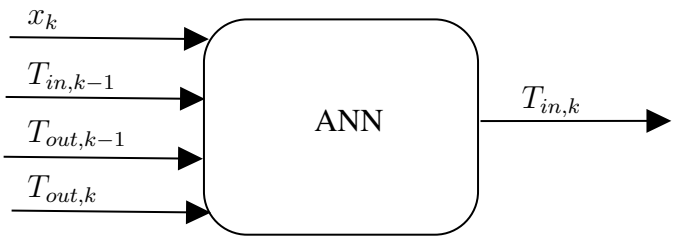


Figure 5.4: The ANN transition function approximator. The ANN has four inputs, including HVAC system status (on/off), x_k, the indoor temperature at $k-1$, $T_{\text{in},k-1}$, the outdoor temperature at k, $T_{\text{out},k}$, the outdoor temperature at $k-1$, $T_{\text{out},k-1}$, while the output is the indoor temperature at k, $T_{\text{in},k}$.

5.2.3 Artificial neural network function approximators

Building on ALG 1, we now describe the ANN function approximator that is used to further improve the algorithm. As mentioned in Section 5.1, the main computational burden in ALG 1, is solving the transition function ODEs (equations (3.18), (3.19) and (3.20)) at each time step. To speed up the algorithm, we train an ANN to be used instead of ODEs in each time step. We used a simple two-layer feed-forward network. The hidden layer has 10 neurons, and the output layer with only one neuron. In the ANN structure, we have three inputs: previous time step outdoor temperature, current time step outdoor temperature and previous time step indoor temperature and one output which is the current time step indoor temperature. We separately trained ANN for ODEs when HVAC is on or off.

To this end, we implement our methodology and call it Algorithm 2 (ALG 2). In the next section, we evaluate the quality of the solutions resulting from each state-space approximation, multi-timescale, and ANN function approximator.

5.3 Evaluation and discussion

We begin by quantifying the loss in solution quality from using the state-space approximation. We then examine the quality of the policies computed using, respectively, ALG 1 and ALG 2, and compare their computational performance. We define three measures of the solution quality, which we use in Section 5.3.1. These are: (i) *mean-absolute error* (MAE) of the approximated state-space against the exact state-space (i.e. the temperature error), (ii) MAE of the final value function resulting from applying VI on the approximated state-space versus the final value function resulting from applying VI on the exact state-space without temperature discretization, and (iii) normalized *calibration error* between the optimal policy using DP on the approximated state-space compared to the optimal policy using DP on the exact state-space[5]. In Section 5.3.2, similar measures are used to compare the results of ALG 1 and ALG 2. The simulations of the Section 5.3.2 were run in MATLAB using a computing platform with an Intel 2.7 GHz i7-7500U CPU, 64-bit operating system and 16 GB RAM. In contrast, for Section 5.3.1, we used a high-performance computer cluster due to the excessive computational burden.

[5]Here we use the term *calibration* in the statistical sense, to measure the fit of the approximate DP method to exact DP.

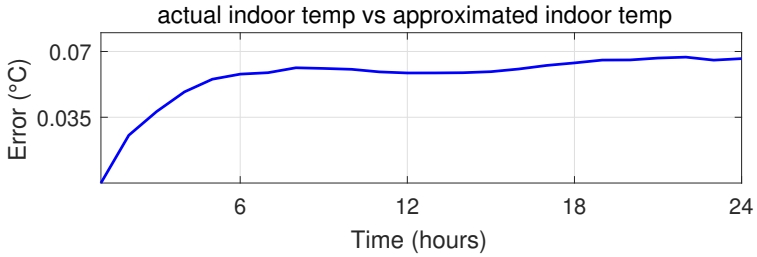


Figure 5.5: Root-mean-square error (RMSE) of indoor temperature of approximated state-space versus indoor temperature of actual state-space.

5.3.1 Evaluation of state-space approximation

As mentioned in the previous section, the runtime of the exact VI for a time horizon of 24 hours is nine days. Thus, we evaluate the performance of state-space approximation by calculating the performance metrics for only one typical summer day in Sydney. To evaluate the approximation with the three criteria above, a sufficient population is required. Therefore, we generated both the approximated and the exact state-space for the 61 initial points between $20\,^{\circ}\mathrm{C}$ and $26\,^{\circ}\mathrm{C}$ with a $0.1\,^{\circ}\mathrm{C}$ discretization.

First, we calculate the MAE of the approximated state-space, which represents the indoor temperature versus the actual state-space over a time horizon of 24 hours. Fig. 5.5 illustrates the RMSE of approximated state space, which represents indoor temperature versus the actual state space over a time horizon of 24 hours. The maximum error is approximately $0.06\,^{\circ}\mathrm{C}$, which is small and acceptable. Second, as mentioned in Section 5.1, the final value function (5.2) consists of two parts: the electricity cost and the discomfort cost. For each, the MAE is calculated separately. The results show that the MAE of the electricity cost for $\lambda = 0.95$ is 3.7 cents, relative to an average value of 44.2 cents. For the discomfort cost, the MAE is about $1.07\,^{\circ}\mathrm{C}$ relative to an average of $22.3\,^{\circ}\mathrm{C}$. These results indicate that the error of the final value function is acceptable.

Third, the approximation is further verified by calculating the calibration error between the optimal policy of the exact and approximated models over a 24-hour horizon. In more detail, the calibration error for each starting point is the difference in the total number of on-cycles in the equivalent optimal policy divided by a total number of time steps in the defined time horizon (24), and the results are averaged over the 61 starting points. Comparing the two cases shows that the difference in the total number of on-cycles in the optimal policies is three. Given this, the calibration error is $0.2\,\%$, which is acceptable.

Overall, we can witness a good trade-off between the superior runtime[6] (almost speed-up of up to 3 times) of VI algorithm on approximated state-space and a better accuracy of applying it on continuous state-space, which permits us to move on to ALG 1 and ALG 2.

Table 5.1: Evaluation of the proposed method of MADP.

HVAC system with deadband relay	Average no. of HVAC on cycles		1859.0
	Average cum.	Elec. cost ($)	1159.1
		Discomfort (°C)	27172.0

			$\lambda = 0.95$	$\lambda = 0.05$
ALG 1	Average no. of HVAC on cycles		952.5	1397.1
	Average cum.	Elect. cost ($)	565.1	952.1
		Discomfort (°C)	6516.3	5356.0

			$\lambda = 0.95$	$\lambda = 0.05$
ALG 2 vs ALG 1	MAE of indoor temp. of optimal policy (°C/hrs)		0.23	0.30
	MAE of the cum.	Elect. cost error (%)	0.08	0.12
		Discomfort error (%)	0.16	0.84
	Normalised calibration error (%)		0.01	0.52

5.3.2 Evaluation of Algorithm 1 and Algorithm 2

We now evaluate the proposed methodology, ALG 2, and the approach using only the multiple timescales abstraction, ALG 1, in order to ascertain the benefits of the multi-timescale and ANN function approximator. Moreover, to get a better sense of the benefits of using optimal HVAC scheduling in PCM-buildings, the algorithms are compared to a simple deadband relay for controlling the HVAC system. Furthermore, to demonstrate the ability of the two algorithms to capture the customer preferences in terms of electricity cost and comfort, the algorithms are run for two different weighting factors: $\lambda = 0.95$ (more weight on the electricity cost) and $\lambda = 0.05$ (more weight on the thermal discomfort). The results for yearly weather conditions are summarized in Table 6.1.

Deadband policy vs. ALG 1

To begin, we consider a conventional HVAC system operating with a deadband control. We simulate an identical home with an identical HVAC system. The only difference is that the HVAC system is controlled using a deadband controller as opposed to optimal scheduling used in ALG 1 and ALG 2. The deadband range is set between $20\,^{\circ}\mathrm{C}$ and $22\,^{\circ}\mathrm{C}$ for heating, and between $22\,^{\circ}\mathrm{C}$ and $26\,^{\circ}\mathrm{C}$ for cooling. The simulations are run for 61 initial points. We record the average number of HVAC operating hours and the average cumulative electricity cost and compare the performance of the HVAC system with a deadband relay against ALG 1. The results show considerable benefits from using optimization over simple deadband control. Specifically, using ALG 1 with a weighting factor of $\lambda = 0.95$ reduces the number of HVAC operating hours by $48.8\,\%$. Importantly, both cost function components see improvements: the electricity cost decreases by $51.2\,\%$, while the number of discomfort hours also decreases by $76\,\%$. Putting more weight on the thermal discomfort ($\lambda = 0.05$) *reduces* the discomfort

[6]Note that because of high computational burden, this comparison is based on simulation runtime for 24hrs time horizon with hourly time steps.

by $80.3\,\%$ at the expense of a lower reduction in the electricity cost (now $17.9\,\%$), which goes to show that the weighting factor λ has to be carefully tuned for optimal performance with respect to customers' preferences. The results confirm the superiority of ALG 1 over the deadband control.

ALG 1 vs. ALG 2

Finally, we evaluate ALG 2 with respect to ALG 1, using as a comparison metrics: (i) MAE of the indoor temperature; (ii) MAE of the cumulative cost expressed as a percentage deviation from ALG 1; and, (iii) the normalized calibration error. The results for the first metric show that the maximum MAE of the indoor temperature over a typical year for $\lambda = 0.95$ and $\lambda = 0.05$, is $0.2\,^{\circ}$C/hrs and $0.3\,^{\circ}$C/hrs, respectively, which is compared to average temperature of optimal policy using ALG 1 ($21.2\,^{\circ}$C/hrs and $21.1\,^{\circ}$C/hrs for $\lambda = 0.95$ and $\lambda = 0.05$, respectively), is highly acceptable. For the second metric, using neural network approximators in ALG 2, results in acceptable solution quality compared to ALG 1, since MAE of electricity cost and discomfort with both weightings ($\lambda = 0.95$ and $\lambda = 0.05$) is at most $0.84\,\%$. The third metric, the maximum calibration error, is $0.01\,\%$ and $0.52\,\%$ for $\lambda = 0.95$ and $\lambda = 0.05$, respectively. This implies that the optimal policies from ALG 2 and ALG 1 have a very similar number of on-cycles.

Overall, the results of these three metrics show that the performance of ALG 2 is comparable compared to ALG 1. We have also compared the runtime of both algorithms to be able to quantify the computational cost saving resulting from the use of ANN and multi-timescale abstractions. We can observe a speedup of up to 157,600 times compared to the direct application of DP (this is calculated for a day time horizon with an hour time steps). The use of ANN in ALG 2 results in a speedup of up to 12 times, from 6 hours to half an hour, compared to ALG 1.

5.4 Summary

In this chapter, we addressed the challenge of solving an optimal scheduling problem of the HVAC system in HEMS with PCM. The corresponding optimisation problem is a nonconvex nonlinear problem. The current approaches such as LP, MILP, MINLP, PSO, and GA are neither powerful to handle the nonlinearity or provide a global solution. Moreover, the existing method of DP is computationally expensive.

Against this, we developed MADP method to deal with the computational burden of a nonlinear, nonconvex HVAC scheduling problem. The objective of the optimisation problem is to minimise the electricity cost while maintaining the indoor temperature of the building within the comfort range of the end-users. The developed thermal model of the PCM-building in Chapter 3 is used to formulate the optimisation problem. The MADP method involves a state-space approximation, multi-time scale MDPs, and ANN function approximation. We demonstrated the efficacy of the MADP method, by optimising the HVAC schedule over a year time horizon using weather conditions of Sydney. The results demonstrate the superior computational performance of the proposed scheduling algorithm compared to a direct application of DP while maintaining acceptable solution quality.

Applying the developed computationally-efficient method of MADP, solving large scale nonlinear optimisation problems such as an HVAC scheduling problem in a large stock of PCM-integrated Buildings is feasible. This is explained in more detail in Chapter 7.

This chapter is based on [104], and [69], which I was the first author.

Chapter 6

Energy Management of PCM-buildings using Model-free RL method

In Chapter 5, we tackled the nonlinear nonconvex problem of home energy management system (HEMS) with phase change material (PCM) by proposing an efficient version of *approximate dynamic programming* (ADP), namely MADP. Specifically, we demonstrated that approximating the function of the thermal model ordinary differential equations (ODEs) by artificial neural network (ANN) and, on top of that, solving multi-time scale *Markov decision processes* (MDPs) rather than single large MDP significantly reduces the computational time without sacrificing the quality of the solution. However, MADP requires a detailed thermal dynamic model of the building, which is an unachievable requirement due to the wide variety in buildings design and construction. In practice, the developed algorithm will be embedded in a smart meter, which makes plug-and-play functionality paramount. Therefore, it is necessary to generalise the method to make it applicable to buildings with different design and construction types without affecting the algorithm's performance. Another notable drawback of MADP is that it cannot be directly applied to continuous domains.

Against this background, this chapter uses an actor-critic model-free on-policy reinforcement learning method based on deep deterministic policy gradient (DDPG) for energy management in buildings with PCM. In particular, DDPG has garnered considerable attention in solving classic problems such as CartPole swing-up, Atari video games and car driving, to name a few [23]. A DDPG-based algorithm is particularly suited for solving high-dimensional complex problems where sensory inputs are available. In more detail, DDPG is a combination of deep learning for processing sensory input with an actor-critic for reinforcement learning [23]. To explain, in the actor-critic method, the policy structure, known as the actor, selects the action. The estimated value function is the critic that critiques the actor's action [24]. Neural network function approximators are used both for the actor and the critic in DDPG methods for online learning in large state and action spaces. To make the learning process efficient, the agent stores the current state, action, reward and next state in a finite-sized cache which is called *replay buffer* in each interaction with an environment. To make the learning of the neural networks stable, target networks are designed for both the actor and the critic

using soft target updates. The weights of these target networks are updated by having them slowly track the learned networks [23]. Moreover, to efficiently explore the physical control problem, the *Ornstein-Uhlenbeck* (O-U) process is added as a noise process to the actor policy [105].

The idea of using DDPG to schedule HVAC systems in buildings was first proposed in [25]. The DDPG was used in this project for controlling the setpoint of the HVAC system to reduce the overall cost of the heating and cooling system's electricity consumption while ensuring the comfort of the home's inhabitants. The quality of the DDPG solution, however, was not assessed against MADP, which has full access to the system's model.

In contrast to the existing work, this chapter conducts an extensive analysis of the results of the DDPG algorithm against the outputs of the MADP method, proposed in Chapter 5. In summary, the contributions of this chapter are as follows:

1. Developed DDPG algorithm for HEMS in buildings with PCM.
2. Run simulation for 294 episodes that represent different weather conditions.
3. Benchmark the DDPG results against MADP outputs by defining four metrics of (i) the difference in the average discomfort level over 294 episodes, (ii) the difference in the cumulative discomfort level, (iii) the difference between the cumulative number of duty cycles of the HVAC system, and (iv) the difference in the cumulative electricity cost.

This chapter progresses as follows: The following section introduces the reward function that is used instead of the cost function to formulate the MDP. This is just to ease the comparison of the MADP results with DDPG outputs. Next, Section 6.2 covers the necessary theoretic background of the DDPG algorithm, which is the main contribution of this chapter. In Section 6.3, we adopt DDPG on our specific optimisation problem and benchmark the results against the ADP-based approach (MADP). This chapter concludes with a summary in Section 6.4.

6.1 Reward function

To use the MADP method as a benchmark for the proposed model-free optimisation approach, we used the same formulation of the MADP ((5.1)), and the Q-function ((5.5)) that is presented in Chapter 5 of this thesis. However, the cost function of (5.2), to ease the comparison, is converted to the reward function that is given by:

$$R_k(s_k, a_k) = -\lambda c_{\mathrm{g},k} p_k^+ + (\lambda - 1)\left(|T_{\mathrm{in},k} - T_{\mathrm{s}}|\right). \tag{6.1}$$

where (6.1) is the reward function, which the agent gains when taking action a_k in state s_k. The weighting factors are given by λ and $(\lambda - 1)$ are that are applied to the electricity cost of HVAC system and the penalty for deviating away from the desired setpoint of the HVAC system, which is T_{s}. All other variables and parameters are same as (5.2). The reward function includes two parts to balance both electricity cost and comfortability. We refer to the former as the cost part of the reward function

Algorithm 3 : DDPG algorithm [23].

Randomly initialise critic network $Q\left(s, a|\theta^{Q}\right)$ and actor network $\mu\left(s|\theta^{\mu}\right)$ with weights θ^{Q} and θ^{μ}.
Initialise target network $Q'\left(s, a|\theta^{Q'}\right)$ and $\mu'\left(s|\theta^{\mu'}\right)$ with weights $\theta^{\mathcal{Q}'} \leftarrow \theta^{\mathcal{Q}}$, $\theta^{\mu'} \leftarrow \theta^{\mu}$.
Initialise replay buffer B.
for $episode = 1, ..., M$ **do**
    Initialise random process $\mathcal{N}$ for action exploration.
    Receive initial observation state s_1.
    for $k = 1, ..., T$ **do**
        Select action $a_k = \mu\left(s_k|\theta^{\mu}\right) + \mathcal{N}_k$ according to the current policy and exploration noise.
        Execute action a_k and observe reward r_k and observe new state s_{k+1}.
        Store transition (s_k, a_k, r_k, s_{k+1}) in B.
        Sample random mini-batches of N transitions (s_i, a_i, r_i, s_{i+1}) from B.
        Set $y_i = r_i + \gamma Q'\left(s_{i+1}, \mu'\left(s_{i+1}|\theta^{\mu'}\right)|\theta^{Q'}\right)$.
        Update critic by minimising the loss using (6.4).
        Update the actor policy using the sampled policy gradient using (6.5).
        Update the target networks using (6.6a), and (6.6b).
    end for
end for

and the latter as discomfort part. As we mentioned in Section 5.1, the value of T_{s} is different in the heating or cooling mode of the HVAC system. For the simulations of this chapter, we assume the HVAC system works only in heating mode, and T_{s} is assumed $20\,^{\circ}\mathrm{C}$. Moreover, the value of λ in this work is assumed 0.95, which means that reducing the electricity cost is preferred to minimising the discomfort level.

In the next section, we will develop DDPG method as an alternative to model-based approaches. Nonetheless, the MADP method proposed in Chapter 5 can be used as a benchmark for alternative approaches. Thus, even if MADP's performance is not implementable in real systems, it does help us quantify and validate the performance of practical alternatives like DDPG.

6.2 Deep deterministic policy gradient algorithm (DDPG)

In this section, we describe an actor-critic model-free on-policy reinforcement learning method based on the *deep deterministic policy gradient* (DDPG) algorithm to overcome the difficulties of using MADP. As discussed before, the MADP method has two major shortcomings that may limit its deployment in real-world settings. First, MADP can only handle problems with discrete action and state-spaces, which may lead to an exponential growth of action and state spaces for problems with a large state-space. Second, the MADP requires a model of the system under control, which may include the sorts of detailed thermal dynamics outlined in Chapter 3. Both of these shortcomings result in significant increases in the computational requirements of MADP methods in complicated settings.

In contrast, the DDPG algorithm can deal directly with continuous control variables, and is model-free. In particular, in this chapter, the control variable of the HVAC system (the action vector) has a continuous form and is defined as a value $a \in [0, 1]$. DDPG is the combination of deep learning methods from artificial intelligence (AI) with the actor-critic method of reinforcement learning (RL).

The DDPG agent is an actor-critic RL agent that aims to find an optimal policy to maximise its long-term reward. To do so, the agent is trained through observations of state and action sequences and rewards that are received by interaction with the learning environment. Each actor and critic has a deep neural network structure to model the sensory inputs, like the indoor temperature of the building.

The DDPG algorithm, introduced in [23], is described in Algorithm 3. The critic and actor networks are denoted by $Q(s, a|\theta^Q)$ and $\mu(s|\theta^\mu)$, respectively. The actor network maps the states to actions, while the critic network calculates the expected reward by following the policy from the current state pair using the action recommended by the actor. After defining a deep neural network for each actor and critic, the algorithm randomly initialises the weights of each critic and actor networks (Line 1). A copy of each critic and actor network is created as $Q'(s, a|\theta^{Q'})$ and $\mu'(s|\theta^{\mu'})$, respectively to improve the training stability. These networks are called *target networks*. The algorithm randomly initialises the weights of the target networks (Line 2).

To make the training of the actor and critic network robust, a finite-sized cache is defined as a *replay buffer*. In other words, instead of using a transition sample that is collected immediately at each decision epoch, the sample is stored in the replay buffer, and then mini-batches of transitions are sampled randomly from the replay buffer to train the actor and critic networks. To speed up the computation, the length of the buffer replay is limited, and as it hits the limit, the oldest transitions are discarded. Before the training begins, the algorithm initialises the replay buffer (Line 3).

For online exploration, the policy used is given by:

$$a_k = \mu\left(s_k|\theta^\mu\right) + \mathcal{N}_k \tag{6.2}$$

where $\mathcal{N}$ is the *Ornstein-Uhlenbeck* (O-U) process (Line 8) [105]. This value is random noise, added to the policy to ensure the exploration and prevent the algorithm from converging to locally-optimal solutions. In each transition, (s_k, a_k, r_k, s_{k+1}) are stored in mini-batches B and the action for the next state is obtained from the target actor network (Lines 10 and 11).

To train the target critic network, the immediate reward is calculated by (Line 12):

$$y_i = r_i + \gamma Q'\left(s_{i+1}, \mu'\left(s_{i+1}|\theta^{\mu'}\right)|\theta^{Q'}\right). \tag{6.3}$$

The critic network in each transition is updated using (Line 13):

$$L = \frac{1}{N}\sum_i^N \left(y_i - Q(s_i, a_i|\theta^Q)\right)^2 \tag{6.4}$$

which minimises the loss between estimated reward in (6.3) and the expected reward that the critic network gives.

The actor network is trained by a deep deterministic policy gradient, given by (Line 14):

$$\nabla_{\theta^\mu} J \approx \frac{1}{N}\sum_i^N \nabla_a Q(s, a|\theta^Q)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu}\mu(s|\theta^\mu)|_{s_i} \tag{6.5}$$

As we see, the gradient is calculated to obtain the expected reward from transitions in a mini-batch with respect to the weight of the actor network.

Finally, the target critic network and actor network are updated smoothly using the following equations (Line 15):

$$\theta^{\mathcal{Q}'} \leftarrow \tau\theta^{\mathcal{Q}} + (1-\tau)\theta^{\mathcal{Q}'}, \tag{6.6a}$$

$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'}. \tag{6.6b}$$

In the next section, DDPG algorithm will be implemented on our specific problem of optimal control of an HVAC system in PCM-buildings.

6.3 Implementation of DDPG

In this section, we describe our implementation of the DDPG algorithm on our specific optimisation problem in PCM buildings. In the first part, we summarise the assumptions and the hyperparameters that are used in the DDPG algorithm to fit our problem. In the second part, the quality of the optimal solutions that the DDPG gives is assessed against the MADP method by simulating 294 episodes.

6.3.1 DDPG algorithm settings

To implement the DDPG algorithm, we use the Reinforcement Learning Toolbox in MATLAB. The thermal model that we developed in Chapter 3 is used as a learning environment for the DDPG Algorithm. The algorithm is run for 294 episodes. In more detail, each episode represents a day with its own weather conditions. Each episode is 24 hours with one-hour time steps. Here, we only focus on heating, so days that require cooling are excluded (but could be treated the same way). The observation is the indoor temperature of the building, and the action which, has a continuous form, captures the electricity use of the HVAC system.

In contrast to [25], the neural network that we used for each actor and critic has a much more complex structure because of the nonlinear behaviour of the PCM. In particular, each actor and critic network consists of four hidden layers. In the actor network, two layers have 128 neurons, and in the critic network, there is one layer with 128 neurons. In both the actor network and the critic network, the *Adam optimiser* is used for training [106].

The hyperparameters that are used in the DDPG algorithm are listed in Table 6.1.

6.3.2 Performance and benchmarking against MADP

We run the DDPG algorithm for 294 episodes, each representing a day when heating is required. Fig. 6.1 shows the episode rewards and average rewards during the training of the DDPG algorithm over the episodes. We observe that episode rewards fluctuate during the training. Two reasons are identified for this behaviour: (i) each episode is characterised by different weather conditions and (ii) the O-U

Table 6.1: Hyperparameters used in DDPG algorithm.

Hyperparameters used in DDPG	value
Mini-batch size	128
Actor learning rate	1E-04
Critic learning rate	1E-03
Gradient Threshold	1
Sample time	1 hour
Target smooth factor	1E-03
Experience buffer length	1E+05
Noise variance	0.1
Noise variance decay rate	1E-06

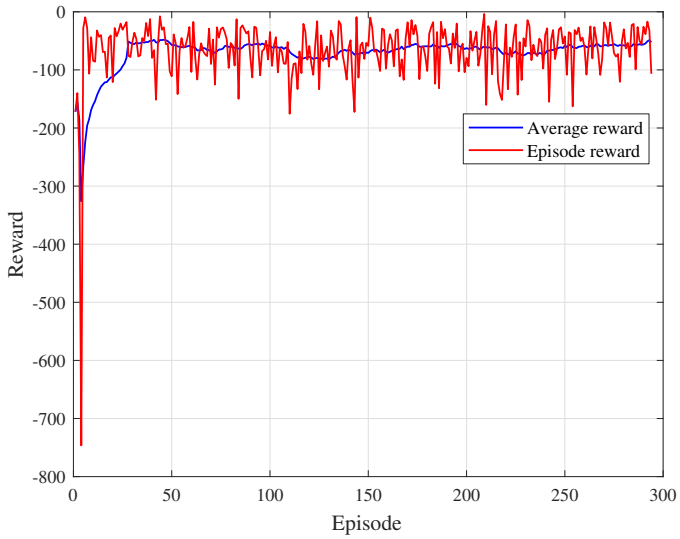


Figure 6.1: Average rewards and episode rewards during the training process of the DDPG algorithm.

Table 6.2: Evaluation of the DDPG method over 42 typical weeks in Sydney ($\lambda = 0.95$)

	Average discomfort (°C / hour)	Cumulative discomfort (°C)	Cumulative number of the HVAC operating hours	Cumulative electricity cost ($)
DDPG	1.1	7804.0	1246.7	696.2
MADP	1.0	6883.5	1448.0	715.6
HVAC system with deadband relay	4.0	28480.6	6373.0	3256.0

noise that is added for policy exploration [1]. Therefore, in each episode, the electricity consumption and discomfort level vary, which leads to fluctuating rewards in each episode. The average reward illustrated in Fig. 6.1 is the average cumulative reward over 24 consecutive episodes that reflects the changing trend of the reward. Observe that the average reward tends to improve steadily during the training process.

To evaluate the performance of the DDPG algorithm, we use MADP as a benchmark against which to measure DDPG. Note that the MADP's performance is in effect an *unachievable* upper bound on the performance of DDPG because MADP's access to the exact thermal dynamics of a building can never be assured; in simulation, we have no way to assess the effects of model errors on the MADP

[1] Exploration means that the algorithm deliberately doesn't follow the optimal policy precisely so that it can explore as-yet-unknown portions of the state space.

method. Moreover, to get a better sense of the benefits of using the DDPG algorithm, we compare it to a simple deadband relay for controlling the HVAC system. To make our comparison fair, identical weather conditions, reward parameters and initial temperatures are used in DDPG, MADP and the deadband control. To assess the quality of the policies computed by these three methods, four criteria are considered, namely:

1. The difference in the average discomfort level over 294 episodes,
2. The difference in cumulative discomfort level,
3. The difference between the cumulative number of the HVAC system operating hours, and
4. The difference in the cumulative electricity cost.

The results for 42 typical weeks (294 typical days) in Sydney, are summarised in Table 6.2.

DDPG vs. MADP

We first compare computation time, then the quality of the policies returned by each algorithm.

We cannot directly compare the runtime of the MADP and the DDPG for simulations of this work as we used different computing platforms to speed up the computation. The MADP method directly incorporates the system dynamics through its internal model, while DDPG has to rely on an external simulation of the system under control to learn a high-quality policy. Therefore, in solving the offline HVAC control problem in PCM-buildings, the DDPG is much faster than the MADP algorithm. Using a computing platform with Intel(R) Core(TM) i7-7500U CPU at 2.70GHz, 64-bit operating system, and 16GB RAM, the runtime of DDPG for 294 episodes is almost one hour and a half.

For the first metric, for the DDPG, the average deviation from the desired temperature $20\,^{\circ}\mathrm{C}$ over 294 episodes is $1.1\,^{\circ}\mathrm{C}/\mathrm{hour}$ compared to $1.0\,^{\circ}\mathrm{C}/\mathrm{hour}$ for the MADP. As we can see, the results are very close to each other. This shows that DDPG can overcome the challenge of continuous action variables, representing duty cycle length, in the HVAC control setting. For the discomfort part, the cumulative sum of comfort penalties for DDPG is $7804\,^{\circ}\mathrm{C}$ and for MADP is $6883.5\,^{\circ}\mathrm{C}$, which are very similar over the duration of the simulation.

For the third metric, the difference between the number of duty cycles of the HVAC system between the DDPG and the MADP is computed by summing the duty cycles of each algorithm, taking the absolute difference. The cumulative duty cycles in the DDPG is 1246.7 and in the MADP is 1448. This implies that the DDPG can control the action continuously and gives an efficient control of the HVAC system compared to the binary control that results from the MADP.

For the fourth metric, we look at the cumulative electricity cost that results from two algorithms. The simulation results show that the cumulative electricity cost in DDPG is \$ 696.2 versus \$ 715.6 for MADP. These values for DDPG show that it produces very good quality control policies that are close to the quality of those computed using MADP.

Fig. 6.2 shows the comparison for one particular episode that simulates the optimal operation of the HVAC system for a typical winter's day in Sydney. Observe that the indoor (controlled) temperature

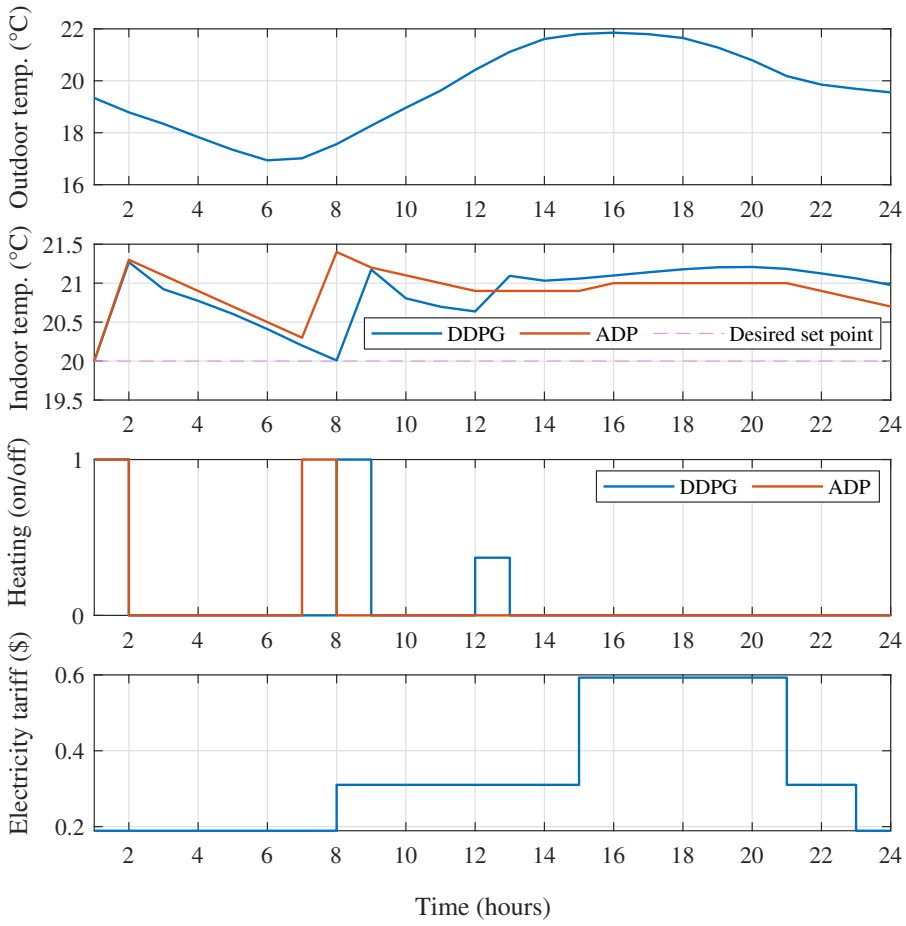


Figure 6.2: Outdoor temperature (top), indoor temperature of the building using the DDPG algorithm vs MADP, the HVAC schedule using the DDPG algorithm vs MADP over a winter's day in Sydney and the electricity time-of-use tariff (ToU) (bottom).

from both the DDPG and the MADP follow each other very closely. Moreover, we can observe a similar number of duty cycles of the HVAC system in Fig. 6.2. However, we can see that in the DDPG algorithm, the HVAC system operates slightly more than for the ADP. This difference in performance comes at the expense of developing a thermal model of each individual building required in the MADP approach. Furthermore, a comparison of the HVAC system operation with the electricity time-of-use tariff shows the desired performance of both the DDPG and MADP algorithms as they push the operation of the heating system into hours with lower electricity price (e.g. before 9am).

Throughout this section, we see that DDPG strikes a good trade-off between the quality of the policies it generates and the simplicity of its implementation, compared to methods that explicitly model system dynamics like MADP.

DDPG vs. Deadband policy

To compare, we consider a conventional HVAC system operating with a deadband control. We simulate an identical home with an identical HVAC system. The only difference is that the HVAC system is controlled using a deadband controller as opposed to optimal scheduling used in the DDPG and MADP. The deadband range is set between $20\,^{\circ}\mathrm{C}$ and $26\,^{\circ}\mathrm{C}$, with a setpoint of $23\,^{\circ}\mathrm{C}$. We record the

values, using same measures as in Section 6.3.2 (refer to Table 6.2). The results show considerable benefits from using optimisation over simple deadband control. Specifically, using DDPG over 42 weeks reduces the cumulative discomfort by $20\,677\,^{\circ}\mathrm{C}$, cumulative operation of the heating system by 5126 hours, and cumulative electricity cost by $2560. Overall, the results confirm the superiority of DDPG over the deadband control.

6.4 Summary

This chapter addresses the major weakness of model-based approaches, such as MADP, which require an explicit thermal model of the building under control. This requirement makes a plug-and-play implementation of the energy management algorithm in an existing smart meter difficult due to the wide variety of building design and construction types. Moreover, due to the need for discretisation, the MADP methods can not directly apply to continuous control problems (such as the HVAC system control). To overcome this difficulty, we developed a model-free actor-critic on-policy RL method based on DDPG. The DDPG algorithm can learn policies in continuous action spaces without access to the full dynamics of the building, and the thermal model of the building is served as a learning environment of the DDPG algorithm. We demonstrate the competitive performance of DDPG by evaluating its performance over 294 days of different weather conditions and comparing this to an MADP method that is known to provide high-quality control policies. The results show that the DDPG model-free RL compares well to MADP, with only a small drop-off in performance. This is balanced against the ability to deploy DDPG in settings where full access to the underlying dynamics of the system are not available, which is typically the case for HVAC in multi-zone buildings. The founding premise of this work is using DDPG as a model-free RL method for energy management in PCM buildings.

This chapter is based on [95], which I was the first author.

Chapter 7

PV Self-Consumption in PCM-buildings using MADP

The cost reductions of rooftop solar photovoltaic (PV) and government policies that support PV uptake have resulted in a dramatic increase in the deployment of rooftop solar PV across Australia over the past ten years. As a result, there are now over two million residential rooftop solar systems with a total capacity exceeding $8\,\mathrm{GW}$ in 2020 [107]. On the other hand, reduced feed-in-tariff (the financial remuneration for exporting excess solar generation to the grid) drive householders' interest in increasing PV self-consumption, with battery storage being the most popular choice. However, battery storage is still expensive; hence there is an interest in other loads that can serve as a *solar sponge*.

One possible candidate is the thermal inertia of the building. However, as discussed in the previous chapters, in Australia, buildings' thermal inertia is low and needs to improve with available technologies such as *phase change material* (PCM). Integrating PCM into a lightweight building's envelope improves its thermal inertia significantly. Like battery storage, PCMs can be charged and discharged. This happens during the phase-change from solid to liquid or vice versa. There are some benefits to considering PCM as an alternative to electrochemical batteries. For instance, the specific PCM considered in this research is plant-based[1] that is more environmentally friendly and more easily disposable compared to batteries. Moreover, it has a long lifetime (almost 80 years) compared to batteries (about 15 years). Another benefit is that PCMs are easy to install, making them suitable for retrofitting buildings.

As discussed in Chapter 2, the literature on optimal control of buildings with the active application of PCMs is scarce. Most of the existing research uses simulation-based optimisation and inverse problem-based simulation methods that are not feasible to apply for large-scale optimisation. The second literature gap is the co-optimisation of PCMs with other distributed energy resources such as rooftop PV, particularly in the context of maximising PV self-consumption.

Against this background, this work illustrates how PCMs can be used as an alternative to battery storage to increase PV self-consumption and reduce electricity costs in the context of home energy

[1] https://phasechange.com/biopcm/

management. In more detail, this chapter investigates the extent to which PCM can be used as an alternative to battery storage systems to increase self-consumption of PV generation. In particular, we explore the electricity cost-savings and increase in PV self-consumption that can be achieved by using PCMs in building insulation and the operation of the HVAC system optimised by a *home energy management system* (HEMS). We consider a HEMS with an HVAC system, rooftop PV, and a PCM layer integrated into the building envelope. The objective of the HEMS optimisation is to minimise the electricity cost while maximising PV self-consumption and maintaining the indoor building temperature in the preferred comfort range. As a case study, we analyse 210 residential buildings (in total) in five Australian capital cities with different climatic conditions. For solving this large nonlinear problem, we use *multi-timescale approximate dynamic programming* (MADP) that we developed in Chapter 5 .

The chapter's main contribution is a techno-economic analysis of the viability of PCMs as an alternative to battery storage to increase PV self-consumption and reduce electricity costs. To our knowledge, this work is the first attempt to:

1. Illustrate the potential economic benefits of PCMs for a large number of residential buildings (210 dwellings in total). The benefits of using PCM are evaluated and quantified by calculating (i) electricity cost-saving (ii) increase in PV self-consumption.

2. Exploit the optimal performance of PCMs by defining the problem as an optimisation problem and solving it using a powerful and computationally efficient method (MADP developed in Chapter 5) that can handle the nonlinearity of the optimisation problem.

3. Comprehensively cover different climatic conditions and the variability of end-user electricity demand and PV generation. To that end, we use historical data (the calendar year of 2019) of weather, PV generation and residential demand.

This chapter progresses as follows: first, we demonstrate the storage capacity of PCM using a crude approximation to provide a sense of the PCM's energy storage capacity. Next, the cost function of the problem is formulated. Afterwards, we describe the case studies used in simulations. Next, we present and discuss the simulation results. And finally, the chapter concludes with a summary.

7.1 PCM as an energy storage medium

The PCM acts as an energy storage medium that absorbs and stores heat, which enables preheating and precooling of the building. To quantify the amount of heat the PCM can store, we need to introduce enthalpy. The enthalpy of a system is the sum of the system's internal energy and the product of its pressure and volume. Since the pressure and volume are constant, the variation in enthalpy equals the variation in the internal energy of the PCM. By calculating the change in enthalpy as the temperature of the PCM changes from T_1 to T_2, we can assess the amount of energy stored or released from the

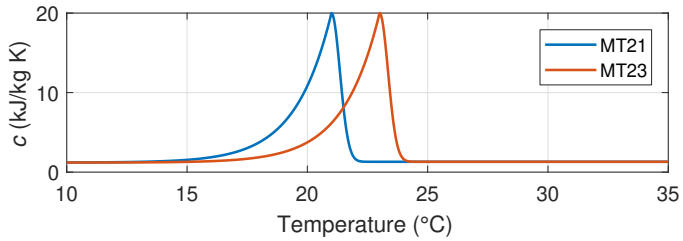


Figure 7.1: Specific heat capacity characteristics of PCMs with melting points of $21\,^{\circ}\mathrm{C}$ (MT21, blue), and $23\,^{\circ}\mathrm{C}$ (MT23, red).

PCM as follows:

$$\Delta H = \int_{T_1}^{T_2} m_{\mathrm{pcm}} c_{\mathrm{pcm}}\,(T)\,dT \quad \text{(J)}, \tag{7.1}$$

where c_{pcm} is the specific heat capacity in $\mathrm{J\,kg^{-1}\,K^{-1}}$ illustrated in Fig. 7.1, and m_{PCM} is the total mass of the PCM in kg used in the building envelope ($2644\,\mathrm{kg}$ in our case).

The enthalpy as a function of temperature along with the specific heat capacity for the melting point of $21\,^{\circ}\mathrm{C}$ is illustrated in Fig. 7.2. For ease of comparison, we have converted the units to more familiar units of kilowatt-hours (kWh). When the PCM temperature changes from $15\,^{\circ}\mathrm{C}$ to $25\,^{\circ}\mathrm{C}$, the PCM stores almost $40\,\mathrm{kWh}$ worth of thermal energy. Considering only the operating range of the PCM between $20\,^{\circ}\mathrm{C}$ to $24\,^{\circ}\mathrm{C}$ (occupant comfort temperature range), the storage capacity of the PCM is about $21\,\mathrm{kWh}$. However, to be able to directly compare that to a storage capacity of an electrochemical battery, we need to consider the COP of the air conditioner to convert the thermal energy into equivalent electrical energy of the HVAC system. Thus, assuming a COP of 4.5, the PCM can store an equivalent of $5\,\mathrm{kWh}$ worth of electrical energy of the HVAC system (assuming no leakage of thermal energy). That is, of course, a crude approximation, which nevertheless gives a sense of the PCM's energy storage capacity.

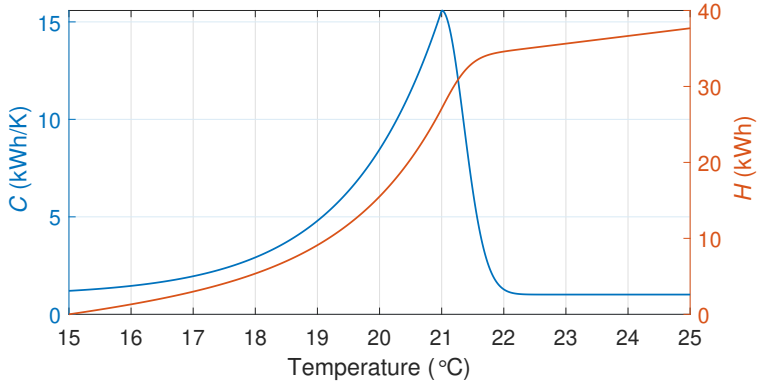


Figure 7.2: Total heat capacity C (left axis) and enthalpy H (right axis) of a PCM with a melting point of $21\,^{\circ}\mathrm{C}$. (Note that these are *total* not *specific* values.)

To further illustrate how a PCM acts as an energy storage medium, we show the thermal performance of a building in Sydney on a typical winter day in Fig. 7.3. The figure compares the indoor temperature and the HVAC system operation of a building with a PCM versus the same building without PCM. The PCM *state of charge* (SOC) on the bottom plot shows the amount of energy stored in the PCM.

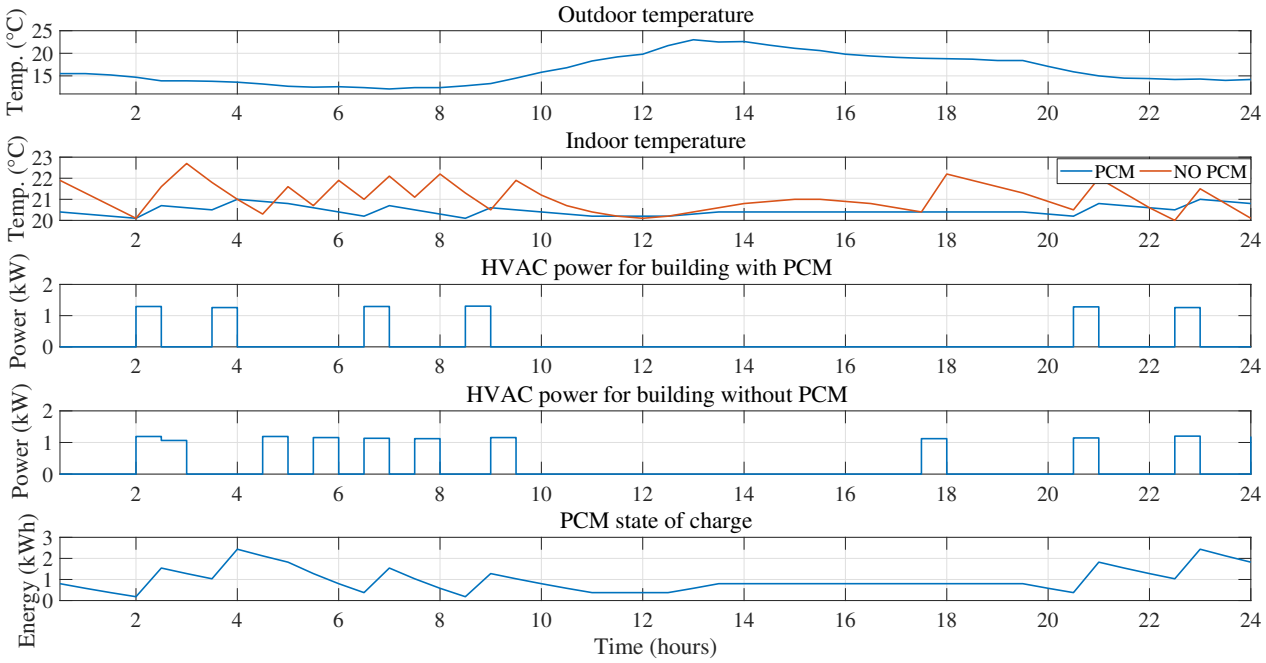


Figure 7.3: Thermal performance of a building for a typical winter day in Sydney. From top to bottom: outdoor temperature, indoor temperature of a building with PCM vs a building without PCM, HVAC power for HEMS with PCM, HVAC power for HEMS without PCM, and PCM state of charge.

The energy is corrected for the COP, so the SOC varies between $0.2\,\mathrm{kW\,h}$ and $2.5\,\mathrm{kW\,h}$ for the indoor temperature between $20.1\,^{\circ}\mathrm{C}$ and $21\,^{\circ}\mathrm{C}$ experienced on the day (assuming SOC is zero at $20\,^{\circ}\mathrm{C}$). Observe how the operation of the HVAC systems increases the SOC of the PCM without a noticeable change in the temperature. On the other hand, in the case without the PCM, the HVAC operation results in a much more pronounced temperature variation, which can negatively affect the occupants' thermal comfort. More importantly, using PCM reduces the HVAC consumption considerably. In particular, it eliminates the operation of the HVAC system during peak hours at 17:30.

7.2 HEMS formulation

Same as Chapter 5, we formulated the problem as HEMS optimisation problem . However, in addition to the HVAC system, and a PCM layer, HEMS has a rooftop PV system. Similar to Chapter 5, we formulated the problem as a MADP (equation (5.1)). However, the cost function is different. The objective is minimising the home's electricity cost while maximising PV self-consumption and maintaining the building's thermal comfort. Therefore, the cost function consists of the cost of *importing* electricity from the grid and the income from *exporting* electricity to the grid:

$$C_k(s_k, x_k) = c_k^{\mathrm{ToU}} p_k^{+} - c^{\mathrm{FiT}} p_k^{-}, \tag{7.2}$$

where p_k^{+} represents total electricity demand at k, and p_k^{-} is PV generation at k. The feed-in tariff c^{FiT} is assumed to be fixed 0.09 $\$/\mathrm{kWh}$, while the electricity tariff $c_{\mathrm{g},k}$ is time-of-use, i.e. it changes with time, as shown in Fig. 5.1. Observe that the electricity cost is always higher than the feed-in tariff,

which means that the optimal strategy is to use as much as possible the power generated by the PV system; that is, minimising cost is equivalent to *maximising PV self-consumption*.

To solve the HEMS optimisation problem, we use the MADP method that we have developed in Chapter 5.

7.3 Case studies

We perform simulations for five Australian capital cities (Sydney, Brisbane, Melbourne, Adelaide and Perth). For each city, we analysed 50 buildings with the same construction (and hence the same thermal performance) but with different demand and solar generation profiles. For Perth, we consider only 10 buildings due to scarcity of data.

7.3.1 Demand data

We assume identical underlying electricity demand for each city that excludes the HVAC demand. The electricity demand of the HVAC system is the output of either a deadband controller or the HEMS, depending on the scenario. To capture the variability that exists in real-world customer demand profiles, we generate fifty random demand profiles, one for each house in each city, using the Bayesian non-parametric method developed in our previous work [108]. The method first constructs a Markov chain model using the electricity demand from the empirical data (Ausgrid's Smart Grid, Smart City data[2] in our case). Then, it synthesises statistically representative demand profiles for an individual house by subsequently sampling the Markov chain model. The generated demand profiles are such that the aggregated demand profile matches well with the aggregated demand profile of the observed data in the Smart Grid Smart City dataset. The demand profile represents the electricity consumption of a medium residential building in Australia, with an average annual electricity demand of approximately $4.7\,\mathrm{MWh}$ (around $12.8\,\mathrm{kWh}$ per day).

7.3.2 Climate data

The temperature data is from the Australian Bureau of Meteorology. We used outdoor dry bulb temperature data from the neighbouring weather stations (Observatory Hill for Sydney, Brisbane City for Brisbane, Olympic Park for Melbourne, West Terrace for Adelaide and Perth Metro for Perth). The data is for the calendar year 2019. The minimum, maximum and average dry-bulb outdoor temperatures are given in Table 7.1.

7.3.3 PV generation data

For PV generation, we use data provided by Solar Analytics.[3] We selected fifty postcodes located in each city and extracted the corresponding solar profiles from the Solar Analytics dataset. Average

[2] www.data.gov.au

[3] www.solaranalytics.com

Table 7.1: Minimum, maximum and average dry-bulb outdoor temperature for the calendar year 2019 from the Australian Bureau of Meteorology.

City	**Minimum** (°C)	**Maximum** (°C)	**Average** (°C)
Sydney	6.2	39.3	18.8
Brisbane	7.6	41.0	21.6
Melbourne	2.4	43.2	15.7
Adelaide	2.3	46.2	17.5
Perth	2.1	41.8	18.6

annual values of PV system generation p^-_{ave}, is given in Table 7.3.

7.3.4 Output variables

Variables of interest in our study are (i) electricity cost-saving; (ii) PV self-consumption, defined as:

$$SC = \frac{\sum_{k=1}^{K} \min(p_k^+, p_k^-)}{\sum_{k=1}^{K} p_k^-} \times 100, \tag{7.3}$$

where p_k^+ represents total electricity demand at time k, and p_k^- is PV generation at time k.

7.3.5 Selection of PCM melting point

To select the optimal PCM melting point, we investigate its impact on electricity cost-saving and PV self-consumption. To do so, we run yearly simulations for all the sites in all the cities, considering melting points of $21\,^{\circ}\mathrm{C}$ (MT21), and $23\,^{\circ}\mathrm{C}$ (MT23). The results are summarised in Table 7.2. In all the cities except Brisbane, the melting point temperature of $21\,^{\circ}\mathrm{C}$ results in more than $40\,\%$ in electricity cost-saving; for Melbourne, this value is as high as $70\,\%$. In Brisbane, MT21 reduces the electricity cost by only $2.4\,\%$. Overall, the difference in cost-savings between MT21 and MT23 is smaller in cities with warmer winters compared to cities with colder winters. On the other hand, different melting points have a small impact on the reduction in PV self-consumption. The same trend is observed for two different PV capacities ($5\,\mathrm{kW}$ and $8\,\mathrm{kW}$). Against this backdrop, we chose MT21 for further analysis.

Table 7.2: Electricity cost-saving and PV self-consumption reduction for PCM melting points of $21\,^{\circ}\mathrm{C}$ (MT21) and $23\,^{\circ}\mathrm{C}$ (MT23), and PV sizes of $5\,\mathrm{kW}$ and $8\,\mathrm{kW}$.

City	Electricity cost-saving				PV self-consumption reduction			
	$5\,\mathrm{kW}$		$8\,\mathrm{kW}$		$5\,\mathrm{kW}$		$8\,\mathrm{kW}$	
	MT21	MT23	MT21	MT23	MT21	MT23	MT21	MT23
	$\bar{x}(\$)$	$\bar{x}(\$)$	$\bar{x}(\$)$	$\bar{x}(\$)$	$\bar{x}(\%)$	$\bar{x}(\%)$	$\bar{x}(\%)$	$\bar{x}(\%)$
Sydney	174.98	113.09	167.64	107.81	1.50	1.49	1.13	1.05
Brisbane	100.58	98.23	91.72	88.73	1.45	2.73	0.99	1.91
Melbourne	306.08	180.40	291.90	171.91	2.30	1.66	1.77	1.18
Adelaide	253.94	168.89	239.40	158.00	2.56	2.64	1.89	1.83
Perth	153.89	109.35	138.09	81.56	2.65	3.73	1.92	1.84

7.3.6 Simulation scenarios

The simulation scenarios consider two different HVAC controls, HEMS and deadband control (DB), both with or without PCM; this gives four scenarios, capturing all four combinations:

1. **DB** (deadband control without PCM)
2. **DB-PCM** (deadband control with PCM)
3. **HEMS** (HEMS without PCM)
4. **HEMS-PCM** (HEMS with PCM)

The setpoint is $21\,^{\circ}\mathrm{C}$ and $23\,^{\circ}\mathrm{C}$ for the heating and cooling mode, respectively, with a deadband of $\pm 1\,^{\circ}\mathrm{C}$. All simulations are run for a whole year with a half-hourly resolution.

7.4 Results and discussion

The results are split into two parts: first, we compare all four scenarios in terms of the HVAC consumption and PV self-consumption, which serves as a baseline for further analysis. Next, we focus specifically on the impact of PCMs on the performance of the HEMS used to control the HVAC system.

7.4.1 Scenario comparison: summary statistics

Table 7.3 summarises the performance of all four scenarios for all five cities, showing PV generation p^{-}_{ave}, underlying electricity demand d_{ave} (the same for all sites), HVAC consumption $d_{\mathrm{HVAC,ave}}$, and PV self-consumption SC_{ave}. The reported values are averages across all sites (fifty for Sydney, Brisbane, Melbourne and Adelaide, and ten for Perth).

As expected, PCMs reduce the HVAC consumption both for HEMS and deadband control. However, there is a significant difference between the two control approaches. For deadband control, PCMs reduce the HVAC consumption by $4.2\,\%$ on average, with a maximum of $6.7\,\%$ in Brisbane and a minimum of $1.9\,\%$ in Melbourne. For HEMS control, on the other hand, PCMs reduce the HVAC consumption by $31.9\,\%$ on average, varying between $28.2\,\%$ in Brisbane and $34\,\%$ in Sydney. The reduction in HVAC consumption is even more pronounced when comparing HVAC and deadband control in buildings with PCM. Adding a HEMS reduces the HVAC consumption by $37.2\,\%$ on average, with a maximum of $39.3\,\%$ in Sydney and a minimum of $32.4\,\%$ in Brisbane.

By contrast, PCMs reduce PV self-consumption in all cases except for Brisbane, where PCMs increase PV self-consumption by $0.5\,\%$ in the case with the deadband control. However, the reduction in PV self-consumption is higher when HEMS is used for HVAC control; on average, PCMs reduce PV self-consumption by $2.1\,\%$, whereas the reduction with the deadband control is only $0.5\,\%$.

Table 7.3: Scenario comparison for a 5kW PV system showing PV generation (p^{-}_{ave}), underlying electricity demand (d_{ave}), HVAC consumption ($d_{\text{HVAC,ave}}$), and PV self-consumption (SC_{ave}). The reported values are averages across all sites (fifty for Sydney, Brisbane, Melbourne and Adelaide, and ten for Perth).

City	p^{-}_{ave} (kWh)	d_{ave} (kWh)	$d_{\text{HVAC,ave}}$ (kWh)				SC_{ave} (kWh)			
			Deadband control		HEMS		Deadband control		HEMS	
			NO PCM	PCM	NO PCM	PCM	NO PCM	PCM	NO PCM	PCM
Sydney	7916.2	4685.3	2031.8	1952.1	1794.5	1184.8	2284.6	2256.9	2218.1	2099.4
Brisbane	9192.5	4685.3	1650.9	1540.3	1449.0	1040.7	2506.8	2557.3	2448.9	2315.6
Melbourne	7332.4	4685.3	3572.3	3503.9	3187.3	2132.5	2361.8	2328.8	2344.9	2176.3
Adelaide	8294.5	4685.3	3455.3	3304.5	3053.2	2104.8	2656.7	2581.3	2621.1	2408.7
Perth	9506.4	4685.3	2032.2	1951.9	1799.8	1200.1	2643.7	2627.6	2623.8	2371.9

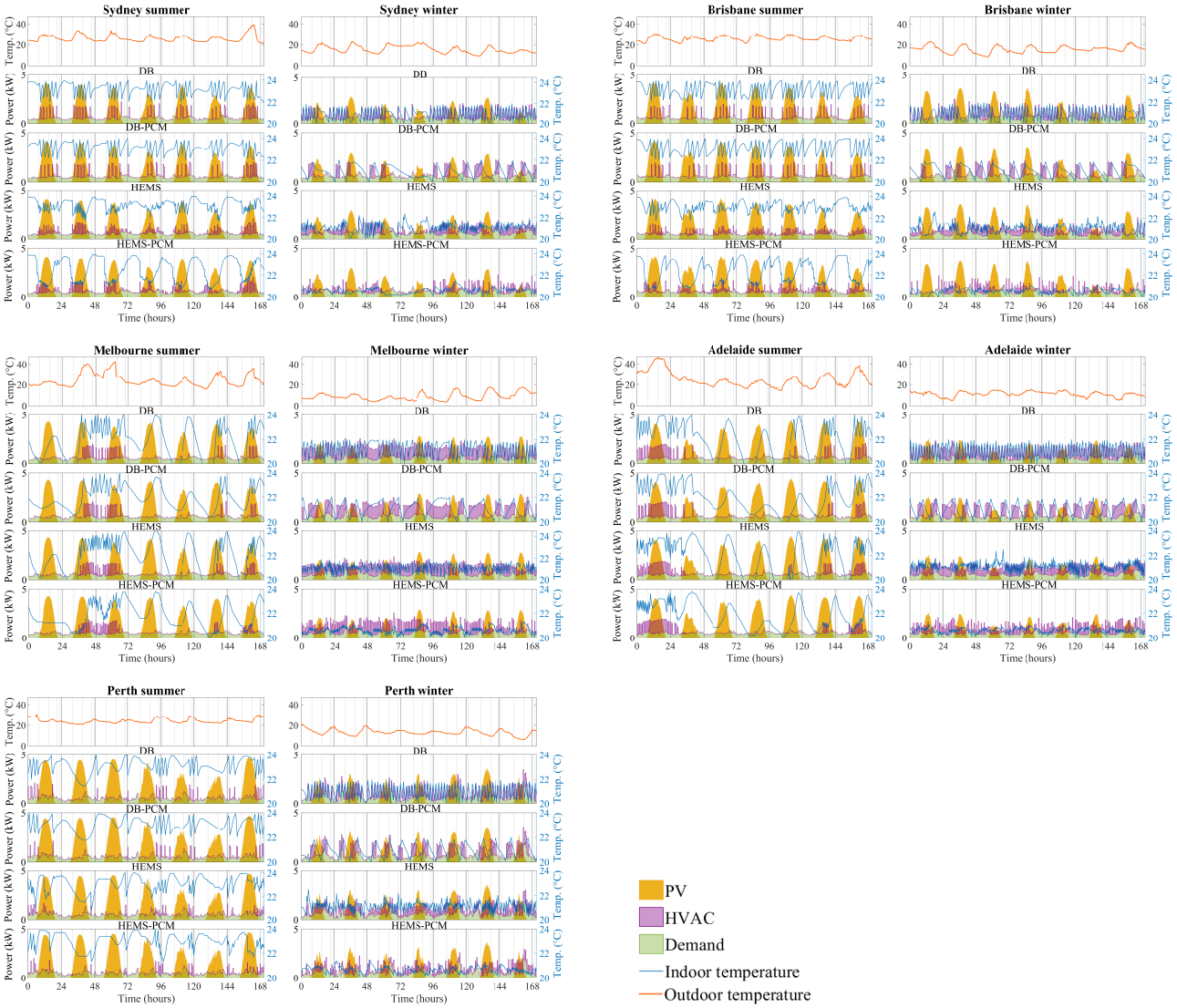


Figure 7.4: Scenario comparison for Sydney, Brisbane, Melbourne, Adelaide, and Perth for a typical summer week (left), and a typical winter week (right). From top to bottom: outdoor temperature, indoor temperature, PV generation for a 5 kW system, electricity demand, and HVAC consumption (average values across all the sites in a city). The minor grid on the x-axis indicates the times of the changes in the time-of-use tariff, as shown in Fig. 5.1 (22:30-7:30 off-peak, 7:30-14:30 and 20:30-22:30 shoulder, and 14:30-20:30 peak).

Table 7.4: Impact of PCMs on HEMS performance: the baseline scenario is a building with a HEMS but no PCM. The reported values are averages across all sites (fifty for Sydney, Brisbane, Melbourne and Adelaide, and ten for Perth). The comparison is made for two performance metrics: electricity cost-saving and PV self-consumption reduction. The comparison is made for two PV system sizes, $5\,\mathrm{kW}$ and $8\,\mathrm{kW}$.

City	Electricity cost-saving								PV self-consumption reduction					
	$5\,\mathrm{kW}$				$8\,\mathrm{kW}$				$5\,\mathrm{kW}$			$8\,\mathrm{kW}$		
	$\bar{x}(\%)$	$\bar{x}(\$)$	$\mathrm{SE}_{\bar{x}}(\$)$	$\sigma_x(\$)$	$\bar{x}(\%)$	$\bar{x}(\$)$	$\mathrm{SE}_{\bar{x}}(\$)$	$\sigma_x(\$)$	$\bar{x}(\%)$	$\mathrm{SE}_{\bar{x}}(\%)$	$\sigma_x(\%)$	$\bar{x}(\%)$	$\mathrm{SE}_{\bar{x}}(\%)$	$\sigma_x(\%)$
Sydney	13.35	174.98	0.90	6.34	17.56	167.64	1.01	7.15	1.50	0.02	0.16	1.13	0.01	0.11
Brisbane	10.60	100.57	0.83	5.90	16.96	94.46	0.71	5.05	1.45	0.03	0.18	1.05	0.02	0.15
Melbourne	16.18	306.10	1.76	12.43	18.51	291.90	2.10	14.89	2.30	0.03	0.18	1.77	0.03	0.18
Adelaide	18.96	253.97	1.33	9.38	26.72	239.41	1.33	9.38	2.55	0.03	0.19	1.89	0.02	0.14
Perth	16.62	153.88	9.66	30.55	27.57	138.08	10.24	32.40	2.65	0.19	0.59	1.92	0.14	0.43

7.4.2 Scenario comparison: typical summer and winter weeks

Fig. 7.4 illustrates the performance of each scenario for a typical summer and winter week in each city. Each plot shows outdoor and indoor temperatures, PV generation, electricity demand and HVAC consumption (average values across all the sites in a city). The minor grid on the x-axis indicates the times of the changes in the time-of-use tariff, as shown in Fig. 5.1 (22:30-7:30 off-peak, 7:30-14:30 and 20:30-22:30 shoulder, and 14:30-20:30 peak) to highlight the load-shifting potential of a HEMS.

A few trends are notable. First, Fig. 7.4 clearly shows the reduced HVAC demand due to the PCM, which is more pronounced in the HEMS cases. Second, the PCM smooths out temperature fluctuations, which is due to the increased thermal inertia. This is most obvious in the DB cases; the PCM's thermal inertia results in a much less frequent on/off toggling of the HVAC system, which results in a smoother temperature profile. Third, controlling the HVAC system by the HEMS shifts the HVAC demand from peak hours to off-peak and shoulder hours to reduce the electricity cost. That is quite pronounced, for example, in winter in Brisbane. Comparing DB and HEMS cases shows that the demand in the peak period (14:30-20:30) in the HEMS case is significantly reduced and shifted to the shoulder period (7:30-14:30). Finally, the high thermal inertia due to the PCM enables preheating and precooling, which results in an even more pronounced shift in the HVAC demand. That is clearly visible, for example, in summer in Sydney (compare HEMS with HEMS-PCM). Observe how the HVAC demand in the HEMS-PCM case is shifted from the peak period (14:30-20:30) to the shoulder period (7:30-14:30).

Fig. 7.4 also illustrates how the choice of the PCM melting temperature affects the optimal HEMS performance. Consider, for example, the summer week in Sydney. Observe how the HVAC demand in the HEMS-PCM case is shifted to earlier in the day compared to the HEMS case. That behaviour appears sub-optimal because running the HVAC in the middle of the day to use the free PV generation would cost less. This seemingly suboptimal behaviour can be explained by considering the thermal properties of the PCM. When the PCM is fully melted, it acts as an additional insulation layer that traps the heat. Because of that, the HVAC has to cool the building down to the PCM melting point ($21\,^{\circ}\mathrm{C}$ in our case) to prevent the indoor temperature from rising above the upper limit of the comfort range ($24\,^{\circ}\mathrm{C}$). That goes to show that a PCM with a higher melting point would perform better in summer.

To investigate this behaviour of PCM in more detail, we plot Fig. 7.5 that demonstrates outdoor and indoor temperatures, PV generation, electricity demand and HVAC consumption (average values across all the sites in a city) for HEMS and HEMS-PCM for a summer and winter week. Moreover, for more observation, we add the plot of the PCM state of charge to Fig. 7.5.

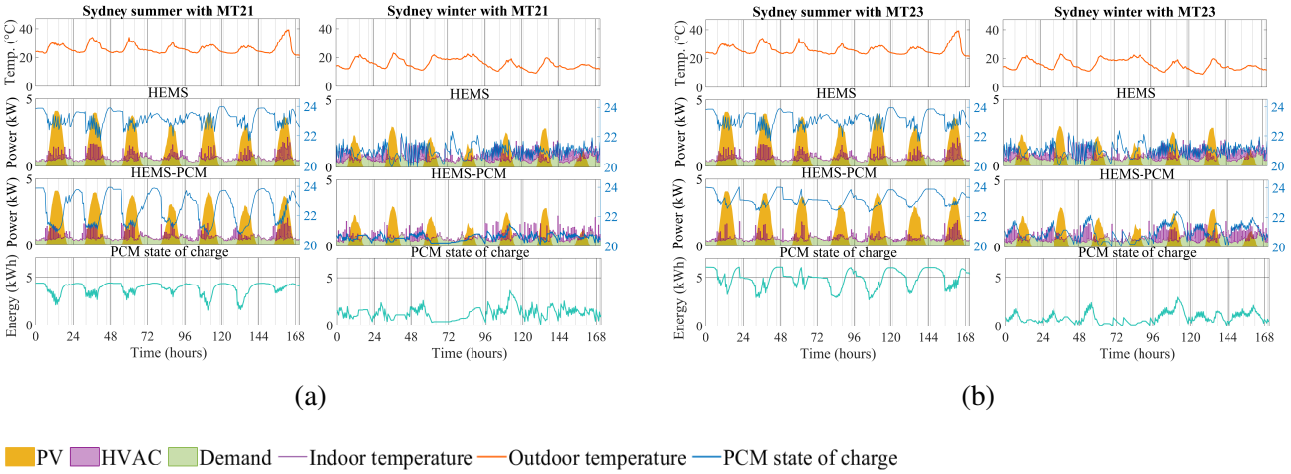


Figure 7.5: Scenario comparison (only HEMS and HEMS-PCM) for Sydney,(a): with MT 21, (b): with MT23, a typical summer week (left), and a typical winter week (right). From top to bottom: outdoor temperature, indoor temperature, PV generation for a $5\,\mathrm{kW}$ system, electricity demand, HVAC consumption, and PCM state of charge (average values across all the sites in Sydney). The minor grid on the x-axis indicates the times of the changes in the time-of-use tariff, as shown in Fig. 5.1 (22:30-7:30 off-peak, 7:30-14:30 and 20:30-22:30 shoulder, and 14:30-20:30 peak).

In the summer week, using MT21 (Fig. 7.5a), the HVAC demand in HEMS-PCM is slightly less than the HVAC demand in HEMS (almost $2.1\,\%$), which results in a marginal electricity saving ($2\,\%$). The reason is that if the PCM temperature goes beyond its melting point on a very hot summer day, it is no longer effective to reduce the HVAC consumption. As explained above, it acts like extra insulation that reverses the PCM effectiveness and increases the HVAC system operation. Based on this fact, and for an optimal operation of the HVAC system in HEMS-PCM, the optimal policy is to precool the building in off-peak hours to keep the PCM temperature around melting point ($21\,^{\circ}\mathrm{C}$) during the day to avoid undesired behaviour of the PCM. Whereas in HEMS (without PCM), the indoor temperature can increase up to around $24\,^{\circ}\mathrm{C}$, there is less or no requirement for the HVAC system operation early in a day during the off-peak period. As a result, the HVAC system can operate during the day when there is an excess generation of electricity is available from the rooftop PV system. Looking at the summer week with MT23, we observe that PCM with MT23 acts much better than MT21. The reason is that for the effective performance of PCM, the indoor temperature can increase up to near to the PCM melting point, which is $23\,^{\circ}\mathrm{C}$. Therefore, compared to the MT21 performance on the summer week, there is much less demand to precool the building in off-peak hours. In this case, the HVAC system mainly operates when the PV generation is variable. We verify our observations further by looking at the electricity cost and the HVAC system consumption. In HEMS-PCM with MT23, compared to HEMS, we can save about $25.4\,\%$ in electricity cost and have $31.5\,\%$ less HVAC demand. Moreover, from the PCM state of charge plot in Fig. 7.5a, we observe that the stored energy in the PCM with MT23 is $27\,\%$ higher than stored energy in PCM with MT21.

Against this, for winter week (Fig. 7.5b), PCM with a lower melting point (in this work, MT21) is more effective than PCM with a higher melting point (in this work, MT23). By integrating PCM with MT21, in HEMS-PCM, we can save $20.2\,\%$ in the electricity cost and $37\,\%$ reduction in the HVAC demand compared to HEMS. While using PCM with MT23, the cost saving is about $10\,\%$ less, and we have only a $14.2\,\%$ reduction in the HVAC system consumption compared to HEMS. The reason is in HEMS-PCM with MT21, with the reduction in the outdoor temperature, the optimal policy tries to keep the indoor temperature around $20\,^{\circ}\mathrm{C}$. This is the indoor temperature, which requires minimum operation of the heating system while it is within the comfort range of the householders. This temperature ($20\,^{\circ}\mathrm{C}$) is very close to the melting point of the PCM with MT21. Therefore decrease in the outdoor temperature results in solidifying of PCM around the melting point and releasing a considerable amount of heating during the off-peak period (in winter during the off-peak period, usually is when the outdoor temperature is very low). This entails much less requirement for the heating system operation to maintain the indoor temperature in the comfort range. Whereas, in the HEMS-PCM with MT23, having PCM temperature around $20\,^{\circ}\mathrm{C}$ is far from its melting point ($23\,^{\circ}\mathrm{C}$), and there is not a noticeable release of heating to the building interior. Therefore, the HVAC system operates much more frequently during off-peak hours to keep the indoor temperature in the comfort range. The PCM state of charge for MT21, over the winter week, is $56.7\,\%$ higher than the PCM state of charge with MT23.

However, because we consider the cost for the whole year, the benefit of a PCM with a lower melting point in winter outweighs the suboptimal performance in summer.

7.4.3 Impact of PCMs on HEMS performance

Table 7.4 summarises the analysis of the impact of PCMs on the performance of a HEMS using two performance metrics: electricity cost-saving and the reduction in PV self-consumption for two PV system sizes ($5\,\mathrm{kW}$ and $8\,\mathrm{kW}$). The table shows for each city: the average value across all sites $\bar{x}$, standard error of the mean value $\mathrm{SE}_{\bar{x}}$ and standard deviation σ_x. For electricity cost-saving the values are given in \$ and also in %, while for the PV self-consumption reduction, all the values are given in %.

The key observation is that PCMs reduce the electricity cost by between $10.6\,\%$ in Brisbane and $19\,\%$ in Adelaide. Increasing the PV size from $5\,\mathrm{kW}$ to $8\,\mathrm{kW}$ reduces the electricity cost even further, by between $17\,\%$ in Brisbane and $27.6\,\%$ in Perth. By contrast, PCMs reduce self-consumption by between $1.5\,\%$ in Brisbane and $2.7\,\%$ in Perth. Increasing the PV size to $8\,\mathrm{kW}$ increases PV self-consumption somewhat, but compared to the base case, PV self-consumption is still reduced. The results are summarised in Figs. 7.6 and 7.7 for a $5\,\mathrm{kW}$ and $8\,\mathrm{kW}$ PV system, respectively.

7.5 Summary

The chapter has explored the potential of PCMs in building insulation to serve as a *solar sponge* to increase the self-consumption of rooftop solar PV when a HEMS optimises the operation of the HVAC system. We used our previously developed MADP method to address the non-convexity of the

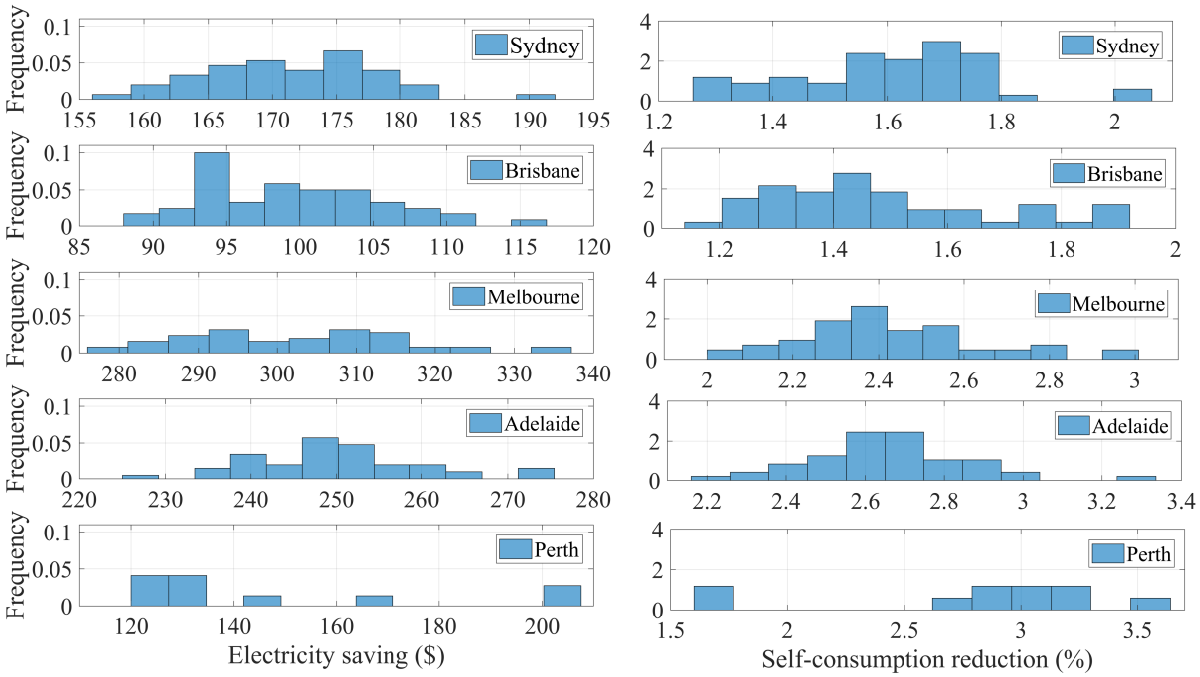


Figure 7.6: Comparison of HEMS with PCM vs HEMS without PCM: histogram of cost-saving (left), and reduction in PV self-consumption (right) for a PV system size of 5 kW.

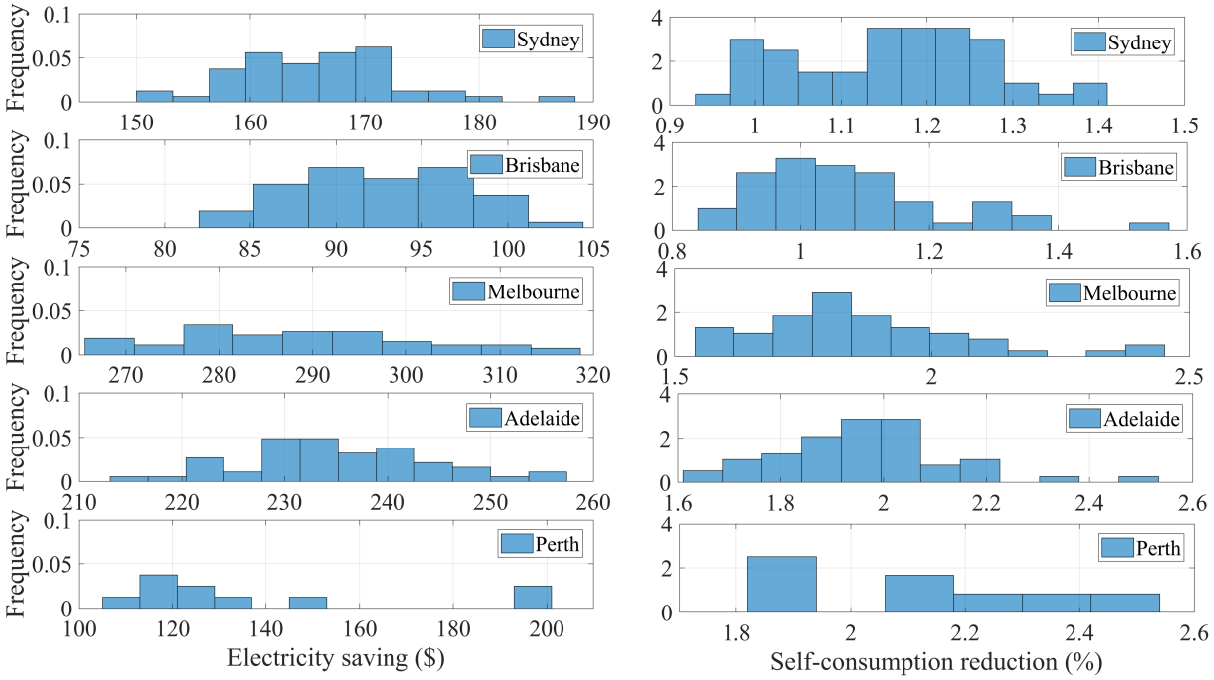


Figure 7.7: Comparison of HEMS with PCM vs HEMS without PCM: histogram of cost-saving (left), and reduction in PV self-consumption (right) for a PV system size of 8 kW.

optimisation problem used for HVAC scheduling (Chapter 5). This method is computationally efficient and can deal with the nonlinear characteristics of the PCM.

Our study is in the context of Australia, including 50 sites in each city of Sydney, Brisbane, Melbourne, Adelaide, and 10 sites in Perth (due to paucity of data). We used the electricity cost-saving and the PV self-consumption as performance metrics for our analysis. For the building' thermal dynamic, we used the thermal model of the building that we developed in Chapter 3. While we use a

simplified building model, we believe that the results broadly indicate the role PCMs can play in home energy management.

We defined four scenarios of DB, DB-PCM, HEMS, HEMS-PCM. The HVAC consumption is either an output of applying HEMS control or the deadband control. The identical underlying electricity demand is considered for each city that excludes the HVAC system consumption. Using real-life weather conditions, PV generation and electricity demand data, so the results are statistically representative.

First, we quantify the amount of thermal energy that can be store in the PCM. With the rough calculate, we demonstrated that considering the occupant comfort range of $20\,^{\circ}\mathrm{C}$ to $24\,^{\circ}\mathrm{C}$, and a COP of 4.5, the PCM can store 5kwh worth of electrical energy of the HVAC system.

Afterwards, we run the simulations with two different melting points of PCM (MT21 and MT23) and two different sizes of the PV system ($5\,\mathrm{kW}$ and $8\,\mathrm{kW}$). Based on the yearly simulation results, we found that PCM with MT21 performed better considering the defined metrics than MT23. However, PCM with a higher melting point (MT23) performs better than PCM with a lower temperature (PCM21) for a very hot summer day. In contrast, in winter, MT21 outperformed MT23.

After selecting the optimal melting point of PCM, we run the simulations for all four defined scenarios. We find that using PCM can reduce annual electricity costs by between $10.6\,\%$ in Brisbane and $19\,\%$ in Adelaide. However, somewhat surprisingly, using PCM reduces PV self-consumption by between $1.5\,\%$ in Brisbane and $2.7\,\%$ in Perth. The hypothesis was that the thermal mass of the PCM would allow preheating and precooling of the building by shifting the HVAC operation to the middle of the day, thus *increasing* PV self-consumption. However, the results show that adding PCMs to the building envelope actually *reduces* it. While the HEMS does shift the operation of the HVAC system to midday, this effect is overshadowed by the overall reduction in HVAC operation. Adding a PCM to a building with an HVAC system controlled by a HEMS, namely, reduces the HVAC consumption by around $30\,\%$ in all five cities. While that results in a significant cost-saving, it reduces PV self-consumption by around $1\,\%$ to $3\,\%$. The sensitivity analysis showed that increasing the size of the PV system from $5\,\mathrm{kW}$ to $8\,\mathrm{kW}$ increased the cost-saving, which was expected but had only a limited impact on the PV self-consumption. The PV self-consumption increased by less than one percentage point in all cities but remained lower compared to the base case (HEMS without PCM). Moreover, using HEMS control on average (over all cities, all sites) reduces the HVAC consumption by $31.9\,\%$, while applying DB control, the HVAC demand reduction is $4.2\,\%$ on average.

This chapter is based on [109], which I was the first author.

Chapter 8

PV Self-consumption in PCM-buildings using RL

This chapter is based on the deep deterministic policy gradient (DDPG) method developed in Chapter 6. However, the home energy management system (HEMS), in contrast to the HEMS in Chapter 6, and similar to the HEMS in 7, has rooftop PV generation. The electricity consumption in the HEMS includes both the underlying electricity load (excluding the HVAC consumption) and the HVAC system consumption.

This work investigates many case studies (8 in total) to analyse the impact of the PV system, size of the PV system, PCM, different electricity tariffs such as time-of-use tariff and flat tariff, lack or presence of feed-in tariff, in combination with different weighting factors on electricity consumption, occupants thermal comfort (indoor temperature deviation from the desired setpoint), and PV system self-consumption in the reward function (λ in equation (8.1)). All case studies are conducted on a five-day set of summer in Sydney.

In summary, the contributions of this chapter are twofold:

1. Demonstrating the efficacy of the developed DDPG method with sensitivity analysis through various case studies.

2. Simulating a real-world application of the DDPG algorithm in HEMS with PCM, the HVAC system and a rooftop PV system, using historical data of weather, PV generation and residential demand.

This chapter progresses with a brief of the DDPG algorithm settings used in this chapter, followed by a description of the different case studies. Next, the simulation results are showcased and discussed. Finally, the chapter concludes with a summary.

8.1 DDPG algorithm settings

In this work, we use the DDPG method developed in Chapter 6. Similarly, the thermal model of the building described in Chapter 3 is used as a training environment for the DDPG agent. However, we make a slight change in the reward function (equation (6.1)) as represented in (8.1). More importantly, the electricity demand includes both underlying electricity demand and the HVAC consumption. Moreover, as a complementary to the case study in Chapter 6, we consider a summer month (cooling mode of the HVAC system) as of weather conditions focus for the case studies. Therefore, the setpoint of the HVAC system is set to $T_{\mathrm{s}} = 23\,^{\circ}\mathrm{C}$.

Similar to Section 6.3, each training episode consists of a day time horizon with an hour time resolution. The state observation is the indoor temperature of the building. However, similar to the DDPG agent in [110], a time observation state is added as an extra state observation to the DDPG algorithm. The time observation state makes it easy to track the time of the day during the simulations.

The reward function is defined as follows:

$$R_k(s_k, a_k) = \lambda c_{\mathrm{g},k} p_k^{+} - (1-\lambda)\, T_{\text{punish}} \tag{8.1}$$

where

$$T_{\text{punish}} = \begin{cases} -0.25 \times |T_{\mathrm{s}} - T_{\mathrm{in},k}| & \text{for} \quad |T_{\mathrm{s}} - T_{\mathrm{in},k}| < 1 \\ -|T_{\mathrm{s}} - T_{\mathrm{in},k}| & \text{otherwise} \end{cases} \tag{8.2}$$

and

$$p_k^{+} = P_{\text{HVAC},k} + P_{\text{load},k} - p_k^{-} \tag{8.3}$$

The DDPG training using a time-of-use tariff [111]:

$$c_{\mathrm{g},k} = \begin{cases} \$0.09 & \text{for} \quad P_{\text{HVAC},k} + P_{\text{load},k} \leq p_k^{-} \\ \$0.21 & \text{for} \quad 1 \leq k \leq 7, 23 \leq k \leq 24 \\ \$0.39 & \text{for} \quad 8 \leq k \leq 9, 18 \leq k \leq 20 \\ \$0.37 & \text{otherwise} \end{cases} \tag{8.4}$$

The DDPG training using a flat tariff:

$$c_{\mathrm{g},k} = \begin{cases} \$0.09 & \text{for} \quad P_{\text{HVAC},k} + P_{\text{load},k} \leq p_k^{-} \\ \$0.30 & \text{otherwise} \end{cases} \tag{8.5}$$

When $P_{\text{HVAC},k} + P_{\text{load},k} \leq p_k^{-}$, then $c_{\mathrm{g},k}$ is equal to the feed-in tariff, the amount paid that the householders receive by feeding back the electricity to the grid.

The hyper-parameters for the DDPG agent are shown in Table 8.1.

Table 8.1: Hyperparameters used in DDPG algorithm.

Hyperparameters used in DDPG	value
Mini-batch size	128
Actor learning rate	1E-04
Critic learning rate	1E-03
Gradient Threshold	1
Sample time	1 hour
Target smooth factor	1E-03
Experience buffer length	1E+06
Noise variance	0.3
Noise variance decay rate	1E-06

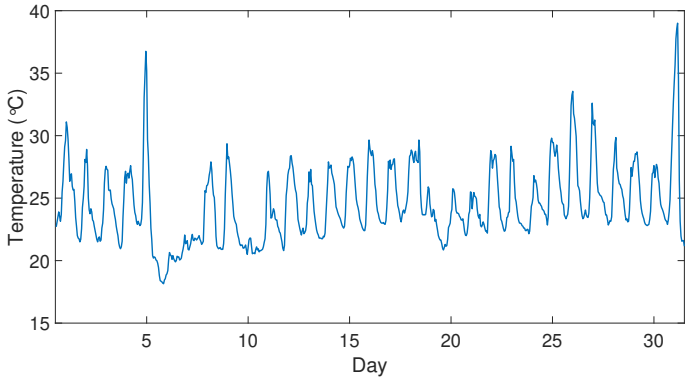


Figure 8.1: Set of temperature training data.

8.2 Implementation

As mentioned before, the reward function (8.1) comprises two parts of electricity cost and discomfort (deviation from the desired setpoint). The share of each part is regulated by λ. A high value for λ reflects the preference of the end-users to minimising the electricity cost rather than comfortability. Finding a balance point between maintaining temperature in the comfort range and reducing electricity cost requires trial and error by running a few training with a different value for λ.

8.2.1 Temperature data

For this work, we use the same data we employed for the city of Sydney in Chapter 7. However, we use only summer month data (January), as illustrated in Fig. 8.1. Since we use hourly time steps in this work, we sample the hourly data from the available half-hourly weather data presented in Chapter 7. Weather data is replicated at least 10,000 times to train the agent for 10,000 episodes.

8.2.2 Solar PV and load data

Similar to Chapter 7, solar PV generation and electricity demand data was added as input data to the optimisation problem. However, unlike Chapter 7, both solar and demand data used in this chapter are from the the Solar Analytics dataset. This data was collected for the month of January from a house in Sydney, Australia. Same as for the weather data, the half-hourly load and PV generation data is

sampled to obtain data with hourly time resolution. Moreover, similar to the weather data used for training, the demand and PV generation data are replicated to use for 10,000 training episodes.

8.2.3 Case studies

The DDPG agents have been trained in different environments with different parameters (8 case studies in total). To find a balance between energy costs and thermal comfort, agents in each environment were trained with a range of λ.

1. **Thermal model**

   The agents are trained with the thermal model described in Chapter 7 using the weather data-set, PV generation data-set, and load data-set as inputs to the DDPG algorithm. The simulations are run with three different values of 0.3, 0.4, and 0.5 for λ. We recorded the optimal scheduling of the HVAC system, the electricity cost and the discomfort using these three values for λ. Based on the results, we find that λ with 0.3 value strikes a good trade-off between the electricity cost and the discomfort. Therefore, the remaining agents are trained with λ values of either 0.28, 0.3, 0.305, 0.33, 0.35 or 0.4.

2. **Thermal Model with PCM Removed** The agents are trained using a thermal model of the building that the PCM layer was replaced with a layer with a heat capacity of $1300\,\mathrm{J\,K^{-1}\,kg^{-1}}$. Agents with λ of 0.3 is trained in this non-PCM environment.

3. **Thermal Model without Solar PV** The agents are trained in an environment without the PV generation system. The agents are trained with values of 0.3 for λ.

4. **Thermal model with increased solar PV** Agents are trained in an environment with a larger PV system size. To calculate the new size of the PV system, we need to calculate the total energy consumption, including the underlying electricity load (excluding the HVAC consumption) and the HVAC system demand. Using the PV load data set (Section 8.2.2), the average daily electricity consumption is about $5.2\,\mathrm{kW\,h}$. For calculating the average HVAC consumption, we need to calculate the total energy storage capacity of the building. To do so, we first convert the total heat capacity of the thermal model ($C_{\text{PCM}}, C_{\text{e}}, C_{\text{a}}$) from $\mathrm{J\,K^{-1}\,kg^{-1}}$ to $\mathrm{kW\,h\,K^{-1}}$. Next, we integrated the total heat capacity over a temperature range that PCM undergoes a phase-change ($22\,^{\circ}\mathrm{C}$ to $28.5\,^{\circ}\mathrm{C}$). The calculations revealed that the total energy storage of the building is approximately $52\,\mathrm{kW\,h}$. Considering the HVAC with a COP of 3, the thermal load is roughly $\frac{52}{3} = 17.3\,\mathrm{kW\,h}$. As a result, the total daily energy consumption is $22.5\,\mathrm{kW\,h}$. According to Section 8.2.2, the daily generation of the solar PV system is approximately $9\,\mathrm{kW\,h}$. To meet the daily energy consumption, taking into account the generation losses, the PV system size needs to be increased by 2.5 times. Therefore, the size of the PV system is considered as $3.75\,\mathrm{kW}$, and the agent, with a λ value of 0.4 is trained.

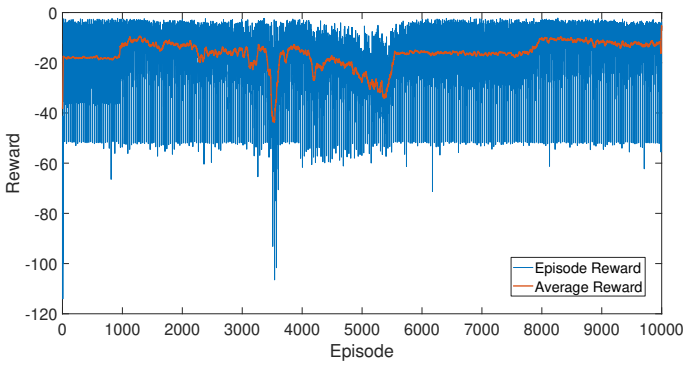


Figure 8.2: Training data from $A_{\text{PCM},0.33}$, showing the episode reward and the average reward (averaging window of 31 episodes).

5. **Thermal model with no feed-in tariff** An agent is trained without having a feed-in tariff for excess generation of the PV system. As a result, this environment will encourage the agent for PV self-consumption, and this may be associated with a potential oversupply of PV generation during the daytime. This agent was trained with a λ value of 0.28.

6. **Thermal model with a flat electricity tariff** In contrast to all other agents, the agent is trained with a flat electricity tariff. In this case, we will be able to see the impact of time-invariant pricing on the optimal scheduling of the HVAC system. The agent is trained with a λ value of 0.33.

7. **Thermal model with PCM with MT24.5** An agent is trained using the thermal model of the PCM building as a learning environment. However, against the baseline scenario, the PCM melting point is $24.5\,^{\circ}\mathrm{C}$. This agent is trained with a λ of 0.305.

8. **Thermal model without a time observation** To train the agent, a baseline thermal environment is used. However, the time observation is removed, and the state observation consists of only the indoor temperature of the building (T_{in}). This agent is trained with a λ of 0.35.

8.2.4 Training

Each agent is trained for 10,000 episodes. Each episode consists of 24 hours time horizon with hourly time steps. As an example for the DDPG training process, the training of $A_{\text{PCM},0.4}$ is shown in Fig. 8.2. As we can see, against the DDG training in Section 6.3.1 of Chapter 6, the training process does not seem to improve steadily like in Fig. 6.1, even after training for 10000 episodes; however, the range of oscillations are acceptable. The reason is, in this work, the agent's learning environment contains additional inputs, such as electricity demand and PV generation, in each episode. We used an Intel(R) Core(TM) i5-8250U CPU for all simulations. On average, a simulation of a case study takes 38,596 seconds (almost 11 hours).

8.2.5 Metrics

A trained agent generates a deterministic policy function. Time and internal temperature are inputs into this function, which generates an action. The DDPG agents produce a continuous action based on a continuous state, and a gradient function determines this continuous action. For state values outside the training range, the action may output a value less than 0 or greater than 1. The action represents the control of an HVAC system, where 0 represents being off, and 1 represents running at full capacity. Consequently, if an action value is greater than 1, it is capped at 1, and if it is less than 0, it is floored at 0. A five day testing period from the training data set was used to evaluate and compare these agents.

The testing data for the evaluation consists of a five-day set of summer dry-bulb outdoor temperature, combined with solar PV generation and underlying electricity load data as shown in Fig. 8.3. Based on the testing data, power usage is calculated from a combination of HVAC power usage from the agent's action, the electricity load and the solar PV generation. The temperature transition between time steps is calculated using the building's thermal model developed in Chapter 3. For numerical analysis of each agent's performance, four metrics are defined as (i) average daily energy consumption, (ii) average daily energy cost, (iii) average hourly deviation from the setpoint and (iv) average daily solar PV self-consumption that are calculated as follows:

$$\text{Net Energy Consumption} = \frac{\sum_{k=1}^{24n}(P_{\text{HVAC},k} + P_{\text{load},k} - p_k^-)}{n} \tag{8.6}$$

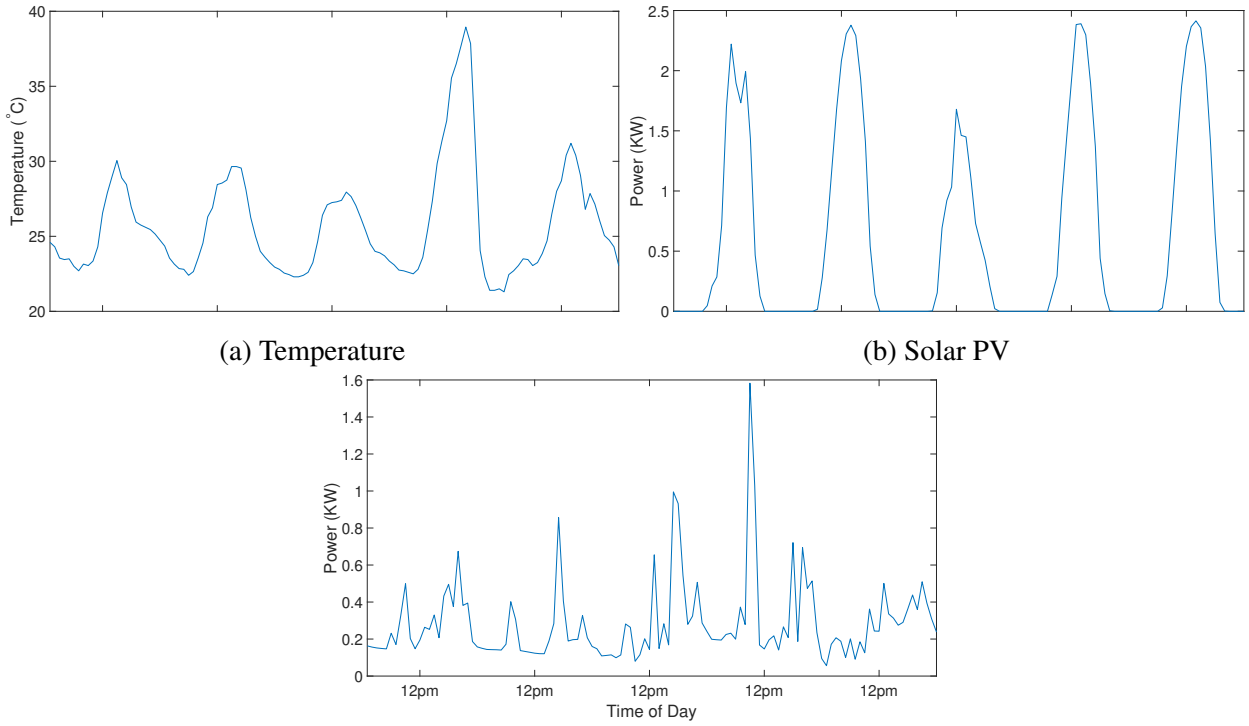
$$\text{Energy Cost} = \frac{\sum_{k=1}^{24n}\left[(P_{\text{HVAC},k} + P_{\text{load},k} - p_k^-) \times c_{\text{g},k}\right]}{n} \tag{8.7}$$

$$\text{Setpoint Deviation} = \frac{\sum_{k=1}^{24n}|T_{\text{in},k} - T_{\text{s}}|}{24n} \tag{8.8}$$

As in the above equations, n represents the number of days in the sample (in this case $n = 5$) and k represents the hour period. During hour k, $P_{\text{HVAC},k}$ represents the energy consumption of the HVAC system, $P_{\text{load},k}$ represents the energy consumption of the building load, and p_k^- represents the energy produced by the solar PV system. The internal temperature of the building during hour k is represented by $T_{\text{in},k}$. The cost of electricity during hour k is represented by $c_{\text{g},k}$.

Even though these metrics provide some quantitative information about the agents' performance, direct comparisons between them can be difficult. It is apparent that there are intrinsic links between these metrics: a reduction in setpoint deviation could potentially indicate increased consumption and costs, and a rise in PV self-consumption implies higher power consumption.

The ideal approach would be to train the agents to maintain a consistent deviation from the setpoint. However, this would require a lot of trial and error, making training time-consuming. For comparison, one agent from each variant is chosen such that the deviation from the setpoint is relatively consistent across the selected agents. Moreover, some environments are quite different; For instance, in an environment with a larger PV system, the power usage will be considerably lower.



(a) Temperature

(b) Solar PV

(c) Load

Figure 8.3: Testing data.

8.3 Results and discussion

In this section, we analyse the trained agents' responses to different melting points of PCM, solar PV system size, and financial incentives. Moreover, we discuss the use of DDPG as an optimisation technique.

The results of testing the selected agents can be found in Table 8.2. Except for $A_{\text{NoPV},0.3}$ and $A_{\text{IncrPV},0.4}$, agents showed relatively high PV self-consumption. This is due to the increase in the size of solar PV systems since the Solar Analytics dataset was collected, which was primarily for consumption rather than export.

The metrics provide information about the agents' performance and their environment. For instance, $A_{\text{IncrPV},0.4}$ has a relatively high negative net energy consumption because of the larger size of the PV system in its environment. Instead of observing metrics of performance, the focus of this work is on the scheduling strategy (i.e. shifting thermal load) that would minimise costs and maximise PV generation.

8.3.1 Impact of varying PCM

Integration of PCM into the building envelope promotes precooling, as expected. This can be seen by comparing Fig. 8.4a and 8.4b. The PCM results in $2.08\,\mathrm{kW\,h}$ lower energy consumption. The total energy cost and average set-point deviation are very similar in the testing sample.

Using PCM reduces energy consumption by extending the time at which a comfortable temperature can be achieved for the same amount of heat absorption. However, as stated before, the optimal performance highly depends on PCM's melting point. When using MT24.5 PCM, the deviation from

Table 8.2: Comparison of metrics between the selected agents from each case study.

Variant	Name	Deviation from Setpoint (°C)	Net Energy Consumption (KWh)	Energy Cost ($)	PV Self-Consumption (%)
Standard	$A_{\text{PCM,0.33}}$	0.71	-5.17	7.40	38.39
No PCM	$A_{\text{NoPCM,0.3}}$	0.51	-3.09	7.10	50.43
No PV	$A_{\text{NoPV,0.3}}$	1.09	64.28	18.83	N/A
Increased PV	$A_{\text{IncrPV,0.4}}$	0.68	-109.05	-2.47	24.31
No Feed-In	$A_{\text{NoFiT,0.28}}$	0.69	-10.37	7.84	52.06
Flat Tariff	$A_{\text{Flat,0.33}}$	0.71	2.91	8.93	47.96
MT24.5 PCM	$A_{\text{PCM2,0.305}}$	0.55	-4.92	7.86	37.75
No Time Obs.	$A_{\text{NoTime,0.35}}$	0.27	-2.88	7.74	48.36

the setpoint is 0.16°C lower than in the standard PCM environment, while it uses $0.25\,\mathrm{kW\,h}$ more energy per day. Fig. 8.4a and 8.4g show this clearly. Possibly a more comprehensive approach to measuring thermal comfort would be to weigh higher temperatures as less comfortable than lower ones.

Taking into account the total load, energy consumption is calculated as $P_{\text{HVAC}} + P_{\text{load}} - p_k^-$. Consequently, the agent is discouraged from operating at times with high loads, as that would reduce the reward from that particular time step. The agent is trying to maximise rewards over the entire episode (1 day), but at peak load times, reducing energy consumption becomes a higher priority than maintaining indoor temperature close to the HVAC system setpoint.

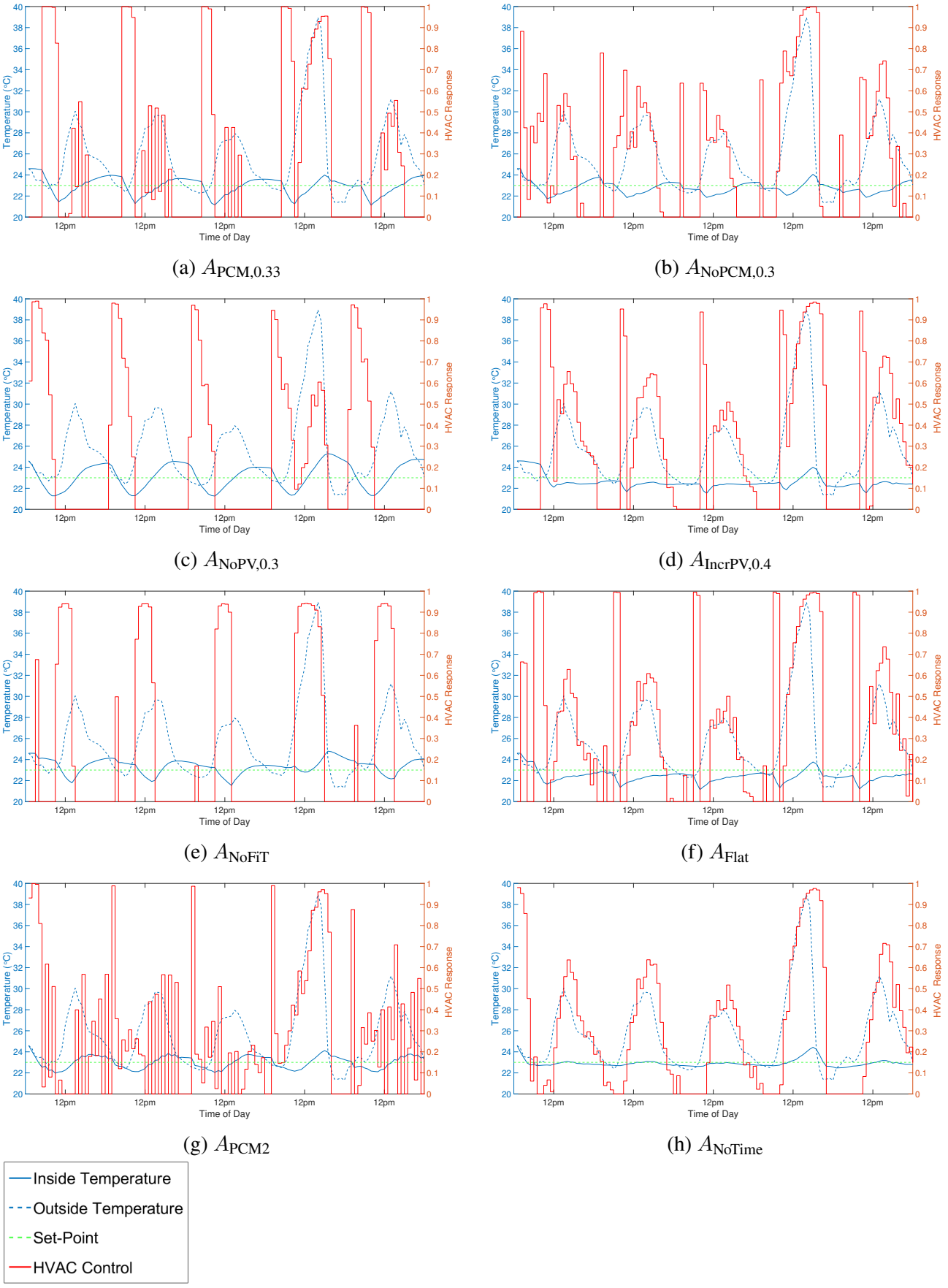
PV self-consumption appears to be reduced when thermal capacity increases. Table 8.2 shows that MT24.5 PCM environment had the lowest self-consumption of 37.7%, compared to standard PCM environment of 38.4% and non-PCM environment of 50.4%. The amount of energy removed during precooling increases as the thermal capacity increases. The agent's responses indicate that less cooling is needed during the middle of the day when solar PV generation is at its peak.

PCM can reduce their energy costs as long as it is able to maintain a comfortable temperature until the evening when the temperature drops outside and provides natural cooling. However, this characteristic is not useful in utilising PV generation, as self-consumption declined as more thermal load was shifted.

8.3.2 Impact of varying solar PV

In the absence of a PV system, the net energy consumption and cost increased. In addition, a larger PV system reduced energy consumption and cost. This can be observed in Table 8.2, comparing the results of $A_{\text{PCM,0.33}}$, $A_{\text{NoPV,0.3}}$ and $A_{\text{IncrPV,0.4}}$. The agent without the PV system mainly relies on precooling to maintain the internal temperature in the desired range. While in the agent with the PV generation, midday cooling, when the solar PV generation is available, can be used as a supplementary strategy. Therefore, a large difference was observed between scenarios with and without the PV system. It is worth noting, however, changing the size of the PV system does not have an impact on HVAC scheduling (refer to Fig. 8.4c and 8.4d).

The increase in generation capacity, $A_{\text{IncrPV,0.4}}$ in Fig.8.4d compared to $A_{\text{PCM,0.33}}$ in Fig. 8.4a, the



(a) $A_{\text{PCM},0.33}$ (b) $A_{\text{NoPCM},0.3}$

(c) $A_{\text{NoPV},0.3}$ (d) $A_{\text{IncrPV},0.4}$

(e) A_{NoFiT} (f) A_{Flat}

(g) A_{PCM2} (h) A_{NoTime}

Figure 8.4: Graphical testing results of the selected agents.

HVAC system works frequently and cycles from cooler to warmer throughout the day to maintain the internal temperature in a building very near the setpoint. Hence, this agent is not effective in utilising increased PV generation.

Despite this, there are benefits to feeding-in excess power at the feed-in tariff rate, and these benefits increase as the system size increases. An alternative strategy is not to use precooling but to use generated electricity for bulk cooling instead at midday. In part, the length of the training episode may deter agents from engaging in this alternate strategy. By cooling in bulk at midday, a daily reward may not maximise its benefits since these benefits may only be realised the following day, while outside of the training episode.

8.3.3 Impact of varying financial incentives

A time-of-use tariff plays a major role in precooling strategy. With the time-of-use tariff employed in this chapter, economically, it is more beneficial to consume less energy during the day. The reason is the electricity has a higher price during the day. In addition, the excess electricity generated by the PV system that is not required for maintaining the indoor temperature in the desired comfort range can be sold back to the grid under the feed-in tariff. The first reason seems to be the primary reason as the agent in A_{Flat}'s strategy demonstrates acting like a thermostat. This can be confirmed further in Table 8.2 as the corresponding PV self-consumption is nearly 48%.

With the removal of the feed-in tariff, excess generation of the PV system could not be sold to the grid, so keeping the temperature close to the setpoint during the day was the sole incentive. As a result, primary precooling was replaced by a more balanced precooling schedule, refer to Fig. 8.4e. Furthermore, this shift led to a significant increase in PV self-consumption, from 38.4% by $A_{\text{PCM,0.33}}$ to 52.1% by $A_{\text{NoFiT,0.28}}$.

Removing the feed-in tariff, the agent was forced to maximise PV self-consumption to maximise the reward. Maximising PV self-consumption resulted in an increase in peak energy consumption, which is concerning.

8.3.4 DDPG as an optimisation technique

In the model-free RL, there is no need for a model of the system dynamics. However, the model is used as a learning environment for the RL agent. In light of this, in this work, the thermal model of the building developed in Chapter 3 is used as a training environment for the DDPG agent. The agent gets only two pieces of information: state observations (in this case, the indoor temperature and time of day) and the reward. A desired performance of the DDPG can be achieved by defining a more comprehensive reward function. However, in this chapter, we aim to illustrate the potential of DDPG as an optimisation method, despite very basic input information

However, complexity is almost always at the expense of computation time. Unlike a simple reward function, a complex reward function requires more training since each component of the reward function must be explored in relation to the agents' actions. The same holds for the state observation;

additional variables in the state (such as the outdoor temperature) bring more accurate predictions about the expected reward but require more time for taring the agent.

Several trials and errors were conducted in order to determine the training duration. It was found that training sessions with a smaller number of episodes (3000, 5000) still produced acceptable solution quality, but it was not robust for all agents. Since an agent has a limited quantity of action space and a considerable portion of its reward function is based on the outside temperature - an aspect beyond its control, gross optimal scheduling strategies could be determined relatively quickly. By further training, we were able to smooth the responses and fine-tune the optimisation.

The training horizon for the agents was 24 hours, meaning that the agents optimised decisions to suit that time frame. Although this was the case, the agents managed to chain together consecutive days of responses in a logical manner. Optimisation could be conducted over a longer time frame, with longer training time being the only downside. The increased training time is due to the backward propagation of information relevant to the critic (Q-value function) since the critic is estimating future rewards of all subsequent time steps until the end of the episode. In other words, moving backwards in time toward earlier state observations will lead to a greater degree of uncertainty in the Q-value function. As the training episode lengthens, more trials are required for good critic predictions at the beginning of the training.

Compared to other research on PCM and DDPG optimisation, this study implements DDPG using a time observation. Despite being trained with timing-sequential data, the agent does not automatically take time-sequenced actions without time as part of the state observation. The gradient parameters are updated using randomly sampled (s_k, a_k, r_k, s_{k+1}) tuples, and these random samples ignore the order in which the tuples are added to memory. As a result, the agent only takes action based on observation of the current state. If this observation only includes the internal temperature, a temperature in the morning that is the same as a temperature in the evening will result in the same response, even though these are very different situations and should require independent actions. A clear example of this is the unique response of A_{NoTime}, Fig. 8.4h.

8.4 Summary

In this chapter, the sensitivity of the developed DDPG method in Chapter 6 is examined thoroughly in a range of scenarios. The scenarios are designed to explore how the PV system, its size, the PCM, feed-in tariffs, various tariff designs, and the preferences of the householders will impact the DDPG algorithm outputs. In the sensitivity analysis, electricity cost, PV self-consumption, and total temperature discomfort are used as metrics.

It has been observed that PCM shifts cooling demand away from hours when PV generation is highest. Thus, PCM integration with HEMS lowers PV-self consumption, as observed in Chapter 7.

During the hot summer days, solar PV generation significantly reduced electricity costs. In particular, when the cost associated with the thermal discomfort far exceeded the cost of running the HVAC system. Agents with different sizes of PV systems gave the same results, showing that the

HVAC system's scheduling is not affected by the PV system's size. Selling any excess PV generation back to the grid was more beneficial in this instance. It is evident that focusing only on scheduling may not be sufficient to analyse the impact of an increase in PV system size.

Using a flat tariff unlike a time-of-use tariff, the HVAC system operates in accordance with outside temperatures. As a result, PV self-consumption increases, but at the expense of higher peak electricity demand. The PV self-consumption can be improved by eliminating feed-in tariff and additionally get benefits of using the thermal mass of the building to shift the cooling demand.

This chapter is based on [112], which I co-supervised.

Chapter 9

Conclusion

This chapter presents an overview of the contributions made in this dissertation in light of the research questions outlined in Chapter 1. To begin with, Section 9.1 presents a summary of the results discussed in the preceding technical chapters. Next, Section 9.2 outlines general ideas and directions of potential future work.

9.1 Summary of results

As discussed in Chapter 1, in recent years, demand response (DR) programs have proved helpful in managing peak demand and meeting sustainability goals, enabling efficient use of the smart grid. Heating, ventilation, and air conditioning (HVAC) loads in buildings constitute a large proportion of the total energy consumption of households. Accordingly, a flexible and efficient operation of these devices can aid power utilities in meeting load management objectives while reducing consumers' electricity bills. The emergence of promising new technologies, such as phase change materials (PCMs), lightweight buildings envelope can serve as thermal energy storage, which improves energy efficiency and allows occupants to offer grid services like peak demand reduction.

To begin with, Chapter 2 reviewed the potential of using buildings envelopes as thermal energy storage for demand response. This is followed by a review of PCM application in buildings starting from PCM thermo-physical properties, different types of PCM, and the type of PCM used in this thesis. The review is extended to literature that investigated the optimal performance of PCM and co-optimisation of PCM with distributed energy resources (DERs). Alas, despite the plethora of literature on energy-saving and seeking the optimal performance of PCM that can achieve in PCM-integrated buildings, still PCM has not found their way into real-world home energy management systems (HEMS) applications. The literature generally lacks to address the search for the optimal performance of PCM as an optimisation problem and, more importantly, highlight the challenge that the nonlinear feature of PCM puts in the corresponding optimisation problem in HEMS. Therefore, Chapter 2 continues by introducing HEMS as a way to formulate the PCM HEMS optimisation problem. This is followed by reviewing methods that are extensively used for solving the HEMS problem. The methods including LP, MILP, MINLP,

and heuristic approaches such as GA, to name a few, are incapable of solving the PCM HEMS problem as the distinctive nonlinear feature of the PCM makes the corresponding PCM HEMS problem a nonlinear nonconvex optimisation problem.

Before exploring the main challenges imposed by solving the nonlinear, nonconvex optimisation problem of the PCM HEMS, the complexity of PCM performance should be addressed. Therefore, Chapter 4 conducted a wide range of simulations to demonstrate that PCM optimal performance considerably depends on different factors such as geographical location, PCM melting point, duration of precooling and preheating, setting points of the HVAC system, thickness and location of PCM in the buildings envelope. Reduction in the HVAC demand, increase in the demand shift and improving the thermal comfort were used to define PCM's optimal performance. The benefit of using PCM in terms of the HVAC demand reduction and the shift was witnessed in all cities. However, using PCM became more encouraging for the cities of Hobart and Melbourne as the results showed the annual HVAC demand was reduced even when compared with heavy constructions like brick-wall buildings. Moreover, It was observed that there is a correlation between the optimal melting point of PCM and the average outdoor air dry bulb temperature (°C) of the city where PCM is used. The cities with lower average temperatures have a lower optimal melting point for PCM.

Along these lines, this dissertation addresses three main challenges that hamper the realistic implementation of HEMS in PCM-integrated buildings.

The first challenge lies in addressing a nonlinear nonconvex nature of the HEMS optimisation problem in PCM-integrated building enforced by the nonlinear feature of PCM specific heat capacity. Overcoming this challenge, therefore, results in solving an optimisation method in HEMS that consists of a layer of PCM as a thermal storage system, rooftop solar photovoltaic (PV) system as a generation, and the HVAC system as a controllable device. The objective is to minimise the electricity cost while maximising the PV self-consumption and maintaining occupants thermal comfort. The objective function is formulated in a way to account for customers preferences in terms of trade-off electricity cost and thermal comfort.

As a first step towards addressing this challenge, Chapter 3 proposes a simplified thermal model of a typical building in Australia. The model is presented with a second-order ordinary differential equations (ODEs) and is enough to approximate a typical building to an acceptable accuracy in the outcome when its performance compares to EnergyPlus software. This simplified model provides a good balance between complexity and accuracy. As a result, its inclusion in the formulation of the HEMS optimisation problem cannot enforce an extra unnecessary computational burden.

Having the simplified model of the building, in Chapter 5, we formulated the PCM HEMS optimisation problem as an Markov decision process (MDP) and proposed computationally-efficient approximate dynamic programming. The method is based on dynamic programming (DP). However, DP suffers from the curse of dimensionality, which means the runtime of the DP algorithm increase exponentially with an increase in time horizon or variables of the problem. Therefore, to speed up the DP algorithm, a few techniques from AI, such as state-space approximation, multi-time scale MDPs, and artificial neural network (ANN) function approximation, were applied in three separate steps. As a

first step, for approximating the state-space, we rounded the transition functions outputs in each time step to the nearest multiple of 0.1. As a second step, and in an attempt to speed up the DP algorithm performance further, the original large MDP is divided into multiple smaller MDPs across the time horizon of the problem. In more detail, the original MDP is converted into small MDPs that are connected successively and solved iteratively using a backward fashion. It is mathematically proven that the solution from multi-time scale MDPs is precisely the same as that of solving the original large scale MDP using the VI algorithm (refer to the Appendix). And finally, in the third step, identifying the main bottleneck of the algorithm, second-order ODEs transition functions are replaced with ANN that is trained offline. Putting all these three steps together results in the computationally efficient multi-timescale approximate dynamic programming (MADP) algorithm that demonstrated superior computational performance compared to a direct application of DP while maintaining acceptable solution quality.

The second challenge lies in generalising the optimisation method of PCM HEMS to make it applicable to buildings with different design and construction types without affecting the algorithm's performance. The proposed MADP method in Chapter 5 requires an explicit thermal model of the building under control. Considering the wide variety of building design and construction types, developing and verifying the thermal model for each individual building is infeasible. Moreover, due to discretisation in the MADP method, the MADP cannot continuously control devices such as the HVAC system. Therefore there is a possibility that the solution is a suboptimal solution rather than an optimal solution. Overcoming this challenge, in Chapter 6, we developed a model-free actor-critic on-policy reinforcement learning method based on deep deterministic policy gradient (DDPG). The DDPG method, as the name implies, requires no explicit model of building dynamics. Nonetheless, despite the remarkable advantage of model-free RL, we still rely on model-based methods such as MADP for full access to the underlying system dynamics to evaluate their solution quality. The developed DDPG algorithm is numerically demonstrated to converge to an acceptable solution quality by benchmarking against the proposed computationally-efficient MADP.

The third challenge is to solve a large scale optimisation of PCM HEMS, to investigate the feasibility of PCM as an alternative to a battery storage system in PCM HEMS. In particular in Chapter 7, using the MADP method (Chapter 5), we conducted a large number of optimisation (210 in total) on selected residential buildings in five cities of Australia to assess the PCM's impact on electricity cost-saving and PV-self consumption. Surprisingly, the results showed that using PCM reduces PV self-consumption as PCM reduces the overall consumption of the HVAC system. Integrating PCM to the HEMS reduces the HVAC consumption by around $30\,\%$ in all five cities. This work paves the way for householders to consider using PCM but now with the ability to envision their application in integration with PV generation in terms of electricity cost-savings, maximising PV self-consumption and improving the occupant comfort.

And finally, as a complementary to Chapter 6, Chapter 8 conducted extensive case studies (8 in total). The case studies investigate the sensitivity of the DDPG performance to the choice of factors such as end-users preference (trade-off cost and thermal comfort), different PV system size, absence or

integration of PCM, and different electricity tariffs, including time-of-use tariff and flat tariff. The case studies are simulated for five-day set of summer in Sydney. Same as Chapter 7, results showed that using PCM is not advantageous in utilising midday PV generation. As a result, we witnessed a reduction in PV self-consumption and a shift in the HVAC demand to the earlier hours of the day.

Table 9.1, matches the research questions that are outlined in Chapter 1 of this thesis with the relevant chapter that attempted to answer them.

Table 9.1: Map research questions outlined in Chapter 1 to the relevant chapter that answered the research question.

Research Questions outlined in Chapter 1	**Chapter answered the question**
RQ1- What is the impact of different factors, including the geographical location, operational conditions of the HVAC system (duration of precooling and preheating, and setting points of the HVAC system), physical properties of PCM (melting point, thickness, location in the envelope), on PCM performance in particular electricity cost-saving and shift in the space heating and cooling demand, in the context of Australia?	Chapter 4
RQ2- What is the powerful and computationally efficient method to optimise the operation of controllable devices such as the HVAC system in HEMS consisting of PCM as a storage system and rooftop PV system as a distributed generation? In particular, the method needs to handle the nonlinearity of the corresponding optimisation method, deliver an acceptable solution quality and be executable on a smart meter with limited computational power and memory.	Chapter 5
RQ3- What is the simplest model that can be used to be represented by low-order of ODEs (less computational cost) and at the same time serve as an acceptable representative model of a typical lightweight building in Australia? How can we validate the performance of the developed model?	Chapter 3
RQ4- How can we generalise the optimisation method in HEMS with PCM to be used as an embedded algorithm installed in the existing smart meters in buildings regardless of the buildings' design and construction types?	Chapter 6
RQ5- How far can we trust in model-free reinforcement methods compared to model-based methods such as DP?	Chapter 6
RQ6- To what extent can we use PCMs as an alternative to battery storage systems to reduce electricity costs and maximise PV self-consumption?	Chapter 7
RQ7- How can different preferences of householders (trade-off between electricity cost and householders' comfortability), different electricity tariffs design, and the different scale of the PV system affect the HEMS performance in managing the cooling demand on hot summer days in Sydney?	Chapter 8

9.2 Future work

The work presented in this thesis puts a step forward towards a real-world integration of PCM in conjunction with DERs to HEMS. However, there are still many gaps to be filled and challenges to be solved in this matter. Thus, the work presented in this dissertation can be extended into the following directions:

- The thermal model of a typical building developed in this thesis can be improved further by considering a more detailed multi-zone design of the building that includes external heat gains and householders' activities associated with heat gains. In this case, the proposed MADP method (presented in Chapter 5) can be applied as a data-driven tool, using the output data of the building simulation software such as EnergyPlus that can simulate a more complicated building structure.

- In this thesis, in the formulation of the MDP, the exogenous information, such as weather conditions, electricity demand and PV solar generation, are treated as deterministic. To improve the quality of the solution in real-world settings, forecasting methods should be applied as part of solving the HEMS optimisation problem to take into account the stochasticity of the MDP's inputs. However, the developed model-free DDPG RL model of the PCM-building that is presented in this thesis can be used without any substantial change to incorporate the stochasticity of the learning environment.

- The main reason that might hinder using techniques such as MADP and RL for energy management in actual buildings is the lack of experimental results to prove the capabilities of these methods in practice. In more detail, future works require simulating real-time energy management in HEMS at least through a laboratory-level testbed.

- In this thesis, we considered a large number of PCM-integrated buildings with rooftop PV to investigate the PCM impact on electricity cost-saving and PV self-consumption. However, as a further step toward reality, the buildings should be considered with different occupant heat gain activities and, importantly, taking into account different preferences of end-users for trade-off electricity-saving and thermal comfort. The multi-agent model-free RL method can be used as a methodology to solve multiple PCM HEMS to maximise the potential for DR.

- The long-term benefit of PCM application on residential customers' energy cost saving and improving comfort needs to be investigated by considering the cost of PCM installation, simulating a scenario close to practical application. Furthermore, the impact of PCM can be studied in terms of the benefits that it can provide on power distribution networks.

Taken together, the advances presented in this dissertation brings the implementation of PCM integration to buildings several steps closer to reality.

Bibliography

[1] American society of heating, refrigerating and air-conditioning engineers, Handbook, ASHRAE Fundamentals, 2009.

[2] Gregory, Katherine and Moghtaderi, Behdad and Sugo, Heber and Page, Adrian, Effect of thermal mass on the thermal performance of various Australian residential constructions systems, Energy and Buildings 40 (4) (2008) 459–465.

[3] Evola, Gianpiero and Papa, Novella and Sicurella, Fabio and Wurtz, Etienne, Simulation of the behaviour of phase change materials for the improvement of thermal comfort in lightweight buildings, in: Proceedings of Building Simulation,12th Conference of International Building Performance Simulation Association, 2011, pp. 1299–1306.

[4] Karthik Muruganantham, Application of Phase Change Material in Buildings: Field Data vs. EnergyPlus Simulation, Master's thesis, Arizona State University (2010).

[5] International Energy Agency Photovoltaic Power Systems Programme, Snapshot of Global PV Markets 2021, Tech. rep. (2021).

[6] IEA, Renewables 2020, Analysis and forecast to 2025, Tech. rep. (November 2020).

[7] Australian Government, Department of Industry, Science, Energy and Resources, Solar PV and batteries, https://www.energy.gov.au/households/solar-pv-and-batteries, accessed: 2021-09-21 (2021).

[8] Balijepalli, VSK Murthy and Pradhan, Vedanta and Khaparde, Shrikrishna A and Shereef, RM, Review of demand response under smart grid paradigm, in: ISGT2011-India, IEEE, 2011, pp. 236–243.

[9] IEA, Tracking Buildings 2020, https://www.iea.org/reports/tracking-buildings-2020, accessed: 2021-09-21 (2020).

[10] European Union, Heating and Cooling, https://ec.europa.eu/energy/topics/energy-efficiency/heating-and-cooling_en, accessed: 2021-09-21 (2021).

[11] Department of Energy, United States of America, An Assessment of Energy Technologies and Research Opportunities, Tech. rep. (September 2015).

[12] Australian Government, HVAC HESS factsheets, Tech. rep. (September 2013).

[13] Australian Government, Australian Renewable Energy Agency, Comparison of Dispatchable Renewable Electricity Options, Tech. rep. (2018).

[14] Zhao Y, Ruether T, Bhatt AI, Staines J , Australian landscape for lithium-ion battery recycling and reuse in 2020 - Current status, gap analysis and industry perspectives, Tech. rep. (2021).

[15] Keerthisinghe, Chanaka and Verbič, Gregor and Chapman, Archie C, A fast technique for smart home management: ADP with temporal difference learning, IEEE Transactions on Smart Grid 9 (4) (2018) 3291–3303.

[16] H. Tischer, G. Verbič, Towards a smart home energy management system-a dynamic programming approach, in: Innovative Smart Grid Technologies Asia (ISGT), IEEE PES, 2011.

[17] de Souza, Antonio Carlos Zambroni and Castilla, Miguel, Microgrids design and implementation, Springer, 2019.

[18] D. Azuatalam, S. Mhanna, A. Chapman, G. Verbič, Optimal HVAC scheduling using phase-change material as a demand response resource, in: Innovative Smart Grid Technologies-Asia (ISGT-Asia), IEEE, 2017.

[19] Bellman, Richard, Dynamic programming, Science 153 (3731) (1966) 34–37.

[20] Bellman, Richard E and Dreyfus, Stuart E, Applied dynamic programming, Princeton university press, 2015.

[21] Powell, Warren B, Approximate Dynamic Programming: Solving the curses of dimensionality, Vol. 703, John Wiley & Sons, 2007.

[22] Sutton, Richard S, TD models: Modeling the world at a mixture of time scales, in: Machine Learning Proceedings, 1995, pp. 531–539.

[23] Lillicrap, Timothy P and Hunt, Jonathan J and Pritzel, Alexander and Heess, Nicolas and Erez, Tom and Tassa, Yuval and Silver, David and Wierstra, Daan, Continuous Control with Deep Reinforcement Learning, in: Int. Conf. Learning Representations (ICLR '16), 2016.

[24] Sutton, Richard S and Barto, Andrew G, Reinforcement learning: An introduction, MIT press, 2018.

[25] Gao, Guanyu and Li, Jie and Wen, Yonggang, DeepComfort: Energy-efficient thermal comfort control in buildings via reinforcement learning, IEEE Internet of Things Journal 7 (9) (2020) 8472–8484.

[26] Albayyaa, Haider and Hagare, Dharmappa and Saha, Swapan, Energy conservation in residential buildings by incorporating Passive Solar and Energy Efficiency Design Strategies and higher thermal mass, Energy and Buildings 182 (2019) 205–213.

[27] Verbeke, Stijn and Audenaert, Amaryllis, Thermal inertia in buildings: A review of impacts across climate and building use, Renewable and sustainable energy reviews 82 (2018) 2300–2318.

[28] Karlsson, Jonathan and Wadsö, Lars and Öberg, Mats, A conceptual model that simulates the influence of thermal inertia in building structures, Energy and Buildings 60 (2013) 146–151.

[29] Reilly, Aidan and Kinnane, Oliver, The impact of thermal mass on building energy consumption, Applied Energy 198 (2017) 108–121.

[30] Xu, Peng, Peak demand reduction from pre-cooling with zone temperature reset in an office building (2006).

[31] Kuczyński, T and Staszczuk, A, Experimental study of the influence of thermal mass on thermal comfort and cooling energy demand in residential buildings, Energy 195 (2020) 116984.

[32] Aste, Niccolò and Leonforte, Fabrizio and Manfren, Massimiliano and Mazzon, Manlio, Thermal inertia and energy efficiency–Parametric simulation assessment on a calibrated case study, Applied Energy 145 (2015) 111–123.

[33] Al-Sanea, Sami A and Zedan, MF and Al-Hussain, SN, Effect of masonry material and surface absorptivity on critical thermal mass in insulated building walls, Applied energy 102 (2013) 1063–1070.

[34] C. Castellón, A. Castell, M. Medrano, I. Martorell, L. Cabeza, Experimental study of pcm inclusion in different building envelopes, Journal of Solar Energy Engineering 131 (4) (2009) 041006.

[35] Faraj, Khaireldin and Khaled, Mahmoud and Faraj, Jalal and Hachem, Farouk and Castelain, Cathy, A review on phase change materials for thermal energy storage in buildings: Heating and hybrid applications, Journal of Energy Storage 33 (2021) 101913.

[36] Konuklu, Yeliz and Ostry, Milan and Paksoy, Halime O and Charvat, Pavel, Review on using microencapsulated phase change materials PCM in building applications, Energy and Buildings 106 (2015) 134–155.

[37] Zhou, Dan and Zhao, Chang-Ying and Tian, Yuan, Review on thermal energy storage with phase change materials (PCMs) in building applications, Applied energy 92 (2012) 593–605.

[38] Al-Yasiri, Qudama and Szabó, Márta, Incorporation of phase change materials into building envelope for thermal comfort and energy saving: A comprehensive analysis, Journal of Building Engineering (2020) 102122.

[39] Behzadi, Sam and Farid, Mohammed Mehdi, Energy storage for efficient energy utilisation in buildings, International High Performance Buildings Conference at Purdue, July 12-15, 2010 (2010).

[40] Qureshi, Waqar A and Nair, Nirmal-Kumar C and Farid, Mohammad M, Impact of energy storage in buildings on electricity demand side management, Energy conversion and management 52 (5) (2011) 2110–2120.

[41] Halford, Christopher K and Boehm, Robert F, Modeling of phase change material peak load shifting, Energy and Buildings 39 (3) (2007) 298–305.

[42] Khudhair, AM and Farid, MM, Use of phase change materials for thermal comfort and electrical energy peak load shifting: experimental investigations, in: Proceedings of ISES World Congress 2007 (Vol. I–Vol. V), Springer, 2008, pp. 283–288.

[43] Ramakrishnan, Sayanthan and Wang, Xiaoming and Sanjayan, Jay and Wilson, John, Thermal performance of buildings integrated with phase change materials to reduce heat stress risks during extreme heatwave events, Applied energy 194 (2017) 410–421.

[44] Ascione, Fabrizio, Energy refurbishment of existing buildings through the use of phase change materials: Energy savings and indoor comfort in the cooling season, Applied Energy 113 (2014) 990–1007.

[45] Rahimpour, Zahra and Faccani, Alice and Azuatalam, Donald and Chapman, Archie and Verbič, Gregor, Using thermal inertia of buildings with phase change material for demand response, Energy Procedia 121 (2017) 102–109.

[46] Medved, Sašo and Arkar, Ciril, Correlation between the local climate and the free-cooling potential of latent heat storage, Energy and Buildings 40 (4) (2008) 429–437.

[47] Sharifi, Naser P and Shaikh, Ahsan Aadil Nizam and Sakulich, Aaron R, Application of phase change materials in gypsum boards to meet building energy conservation goals, Energy and buildings 138 (2017) 455–467.

[48] Adilkhanova, Indira and Memon, Shazim Ali and Kim, Jong and Sheriyev, Almas, A novel approach to investigate the thermal comfort of the lightweight relocatable building integrated with PCM in different climates of Kazakhstan during summertime, Energy 217 (2021) 119390.

[49] Al-Absi, Zeyad Amin and Hafizal, Mohd Isa Mohd and Ismail, Mazran and Mardiana, Ahmad and Ghazali, Azhar, Peak indoor air temperature reduction for buildings in hot-humid climate using phase change materials, Case Studies in Thermal Engineering 22 (2020) 100762.

[50] Liu, Jiang and Liu, Yan and Yang, Liu and Liu, Tang and Zhang, Chen and Dong, Hong, Climatic and seasonal suitability of phase change materials coupled with night ventilation for office buildings in Western China, Renewable Energy 147 (2020) 356–373.

[51] Bai, Lu and Xie, Jingchao and Farid, Mohammed M and Wang, Wei and Liu, Jiaping, Analytical model to study the heat storage of phase change material envelopes in lightweight passive buildings, Building and Environment 169 (2020) 106531.

[52] M. Hagenau, M. Jradi, Dynamic modeling and performance evaluation of building envelope enhanced with phase change material under danish conditions, Journal of Energy Storage 30 (2020) 101536.

[53] Chan, ALS, Energy and environmental performance of building façades integrated with phase change material in subtropical Hong Kong, Energy and Buildings 43 (10) (2011) 2947–2955.

[54] J. Jia, B. Liu, L. Ma, H. Wang, D. Li, Y. Wang, Energy saving performance optimisation and regional adaptability of prefabricated buildings with pcm in different climates, Case Studies in Thermal Engineering (2021) 101164.

[55] G. Tian, H. Lv, J. Huang, P. Liu, W. Feng, Experimental study on the heat transfer characteristics of different walls with phase change materials in summer, Journal of Building Engineering 44 (2021) 103354.

[56] Bimaganbetova, Madina and Memon, Shazim Ali and Sheriyev, Almas, Performance evaluation of phase change materials suitable for cities representing the whole tropical savanna climate region, Renewable Energy 148 (2020) 402–416.

[57] Barzin, Reza and Chen, John JJ and Young, Brent R and Farid, Mohammed M, Application of PCM underfloor heating in combination with PCM wallboards for space heating using price based control system, Applied Energy 148 (2015) 39–48.

[58] Baniassadi, Amir and Sailor, David J and Bryan, Harvey J, Effectiveness of phase change materials for improving the resiliency of residential buildings to extreme thermal conditions, Solar Energy 188 (2019) 190–199.

[59] M. Saffari and A. De Gracia and C. Fernández and L.F. Cabeza, Simulation-based optimisation of PCM melting temperature to improve the energy performance in buildings, Applied Energy 202 (2017) 420–434.

[60] Cheng, Rui and Pomianowski, Michal and Wang, Xin and Heiselberg, Per and Zhang, Yinping, A new method to determine thermophysical properties of PCM-concrete brick, Applied Energy 112 (2013) 988–998.

[61] Byrd, Richard H. and Nocedal, Jorge and Waltz, Richard A., Knitro: An Integrated Package for Nonlinear Optimisation (2006) 35–59.

[62] Boggs, Paul T. and Tolle, Jon W., Sequential quadratic programming for large-scale nonlinear optimisation, Journal of Computational and Applied Mathematics 124 (1-2) (2000) 123–137.

[63] Sun, Xiaoqin and Lin, Yian and Zhu, Ziyang and Li, Jie, Optimised design of a distributed photovoltaic system in a building with phase change materials, Applied Energy 306 (2022) 118010.

[64] Han, Jinsoo and Choi, Chang-Sic and Park, Wan-Ki and Lee, Ilwoo, Green home energy management system through comparison of energy usage between the same kinds of home appliances, in: 2011 IEEE 15th International Symposium on Consumer Electronics (ISCE), IEEE, 2011, pp. 1–4.

[65] Son, Young-Sung and Pulkkinen, Topi and Moon, Kyeong-Deok and Kim, Chaekyu, Home energy management system based on power line communication, IEEE Transactions on Consumer Electronics 56 (3) (2010) 1380–1386.

[66] Asare-Bediako, Ballard and Kling, WL and Ribeiro, PF, Home energy management systems: Evolution, trends and frameworks, in: 2012 47th International Universities Power Engineering Conference (UPEC), IEEE, 2012, pp. 1–5.

[67] Warren B. Powell, Approximate Dynamic Programming - Solving the Curses of Dimensionality, John Wiley & Sons, Inc., 2007.

[68] Keerthisinghe, Chanaka and Verbič, Gregor and Chapman, Archie C, A fast technique for smart home management: ADP with temporal difference learning, IEEE Transactions on smart grid 9 (4) (2016) 3291–3303.

[69] Rahimpour, Zahra and Verbič, Gregor and Chapman, Archie C, Computationally-Efficient Energy Management in Buildings with Phase Change Materials using Approximate Dynamic Programming, in: 2021 IEEE Madrid PowerTech, IEEE, 2021, pp. 1–6.

[70] Khan, Said G and Herrmann, Guido and Lewis, Frank L and Pipe, Tony and Melhuish, Chris, Reinforcement learning and optimal adaptive control: An overview and implementation examples, Annual reviews in control 36 (1) (2012) 42–59.

[71] Mnih, Volodymyr and Kavukcuoglu, Koray and Silver, David and Rusu, Andrei A and Veness, Joel and Bellemare, Marc G and Graves, Alex and Riedmiller, Martin and Fidjeland, Andreas K and Ostrovski, Georg and others, Human-level control through deep reinforcement learning, nature 518 (7540) (2015) 529–533.

[72] S. Kim, H. Lim, Reinforcement learning based energy management algorithm for smart energy buildings, Energies 11 (8) (2018) 2010.

[73] E. Mocanu, D. C. Mocanu, P. H. Nguyen, A. Liotta, M. E. Webber, M. Gibescu, J. G. Slootweg, On-line building energy optimisation using deep reinforcement learning, IEEE transactions on smart grid 10 (4) (2018) 3698–3708.

[74] Vázquez-Canteli, José R and Nagy, Zoltán, Reinforcement learning for demand response: A review of algorithms and modeling techniques, Applied energy 235 (2019) 1072–1089.

[75] Sutton, Richard S and Barto, Andrew G and others, Introduction to reinforcement learning, Vol. 135, MIT press Cambridge, 1998.

[76] Barrett, Enda and Linder, Stephen, Autonomous hvac control, a reinforcement learning approach, in: Joint European conference on machine learning and knowledge discovery in databases, Springer, 2015, pp. 3–19.

[77] ANSI-ASHRAE, ANSI/ASHRAE Standard 140-2017, https://webstore.ansi.org/preview-pages/ASHRAE/preview_ANSI+ASHRAE+Standard+140-2017.pdf (2017).

[78] Chong, Adrian and Gu, Yaonan and Jia, Hongyuan, Calibrating building energy simulation models: A review of the basics to guide future work, Energy and Buildings (2021) 111533.

[79] Al-Saadi, Saleh Nasser and Zhai, Zhiqiang John, Modeling phase change materials embedded in building enclosure: A review, Renewable and Sustainable Energy Reviews 21 (2013) 659–673.

[80] Underwood, Chris and Yik, Francis, Modelling methods for energy in buildings, John Wiley & Sons, 2008.

[81] U.S. Department of Energy, Engineering Reference — EnergyPlus Version 8.5 Documentation, https://bigladdersoftware.com/epx/docs/8-5/engineering-reference/ (2016).

[82] Laret, L, Use of general models with a small number of parameters, Part 1: Theoretical analysis, in: Proceedings of 7th International Congress of Heating and Air Conditioning CLIMA 2000. Budapest, 2000, pp. 263–276.

[83] Gouda, MM and Danaher, Sean and Underwood, CP, Low-order model for the simulation of a building and its heating system, Building Services Engineering Research and Technology 21 (3) (2000) 199–208.

[84] Crabb, JA and Murdoch, N and Penman, JM, A simplified thermal response model, Building Services Engineering Research and Technology 8 (1) (1987) 13–19.

[85] Tindale, A, Third-order lumped-parameter simulation method, Building Services Engineering Research and Technology 14 (3) (1993) 87–97.

[86] Antonopoulos, KA and Koronaki, EP, On the dynamic thermal behaviour of indoor spaces, Applied Thermal Engineering 21 (9) (2001) 929–940.

[87] Nielsen, Toke Rammer, Simple tool to evaluate energy demand and indoor environment in the early stages of building design, Solar Energy 78 (1) (2005) 73–83.

[88] Kämpf, Jérôme Henri and Robinson, Darren, A simplified thermal model to support analysis of urban resource flows, Energy and buildings 39 (4) (2007) 445–453.

[89] Kämpf, Jérôme Henri, On the modelling and optimisation of urban energy fluxes, Tech. rep., EPFL (2009).

[90] Gouda, MM and Danaher, Sean and Underwood, CP, Building thermal model reduction using nonlinear constrained optimisation, Building and Environment 37 (12) (2002) 1255–1265.

[91] Patankar, Suhas V, Numerical heat transfer and fluid flow, CRC press, 2018.

[92] Vakilaltojjar, Saied Mohammad and Saman, W, Analysis and modelling of a phase change storage system for air conditioning applications, Applied Thermal Engineering 21 (3) (2001) 249–263.

[93] Zhou, Guobing and Yang, Yongping and Wang, Xin and Zhou, Shaoxiang, Numerical analysis of effect of shape-stabilized phase change material plates in a building combined with night ventilation, Applied Energy 86 (1) (2009) 52–59.

[94] Zhu, Na and Hu, Pingfang and Xu, Linghong, A simplified dynamic model of double layers shape-stabilised phase change materials wallboards, Energy and Buildings 67 (2013) 508–516.

[95] Rahimpour, Zahra and Verbič, Gregor and Chapman, Archie C., Actor-critic learning for optimal building energy management with phase change materials, Electric Power Systems Research 188 (2020) 106543.

[96] Masy, Gabrielle, Definition and validation of a simplified multi-zone dynamic building model connected to heating system and HVAC unit, Phd, Université de Liege (2008).

[97] Sasic Kalagasidis, Angela, A multi-level modelling and evaluation of thermal performance of phase-change materials in buildings, Journal of Building Performance Simulation 7 (4) (2014) 289–308.

[98] Kośny, Jan, PCM-enhanced building components: an application of phase change materials in building envelopes and internal structures, Springer, 2015.

[99] Rahimpour, Zahra and Verbič, Gregor and Chapman, Archie C., Using Thermal Inertia of Buildings with Phase Change Material for Demand Response, in: Asia-Pacific Solar Research Conference 2018, Australian PV Institute, 2018.

[100] Tischer, Henning and Verbič, Gregor, Towards a smart home energy management system - A dynamic programming approach, in: 2011 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia), IEEE, 2011.

[101] Pipattanasomporn, M. and Kuzlu, M. and Rahman, S., An Algorithm for Intelligent Home Energy Management and Demand Response Analysis, IEEE Transactions on Smart Grid 3 (4) (2012) 2166–2173. doi:10.1109/TSG.2012.2201182.

[102] Alam, Morshed and Jamil, Hasnat and Sanjayan, Jay and Wilson, John, Energy saving potential of phase change materials in major Australian cities, Energy and Buildings 78 (2014) 192–201.

[103] A. I. Orhean, F. Pop, I. Raicu, New scheduling approach using reinforcement learning for heterogeneous distributed systems, Journal of Parallel and Distributed Computing 117 (2018) 292–302.

[104] Rahimpour, Zahra and Verbič, Gregor and Chapman, Archie C., Energy Management of Buildings with Phase Change Materials Based on Dynamic Programming, in: 13th PowerTech 2019, IEEE, 2019.

[105] Uhlenbeck, George E and Ornstein, Leonard S, On the theory of the Brownian motion, Physical review 36 (5) (1930) 823.

[106] D. P. Kingma, J. Ba, Adam: A method for stochastic optimisation, arXiv preprint arXiv:1412.6980 (2014).

[107] Graham, Paul and Havas, Lisa, CSIRO Projections for small scale embedded energy technologies - Report to AEMO, Tech. Rep. June (2020).

[108] Power, Thomas and Verbič, Gregor and Chapman, Archie C, A Nonparametric Bayesian Methodology for Synthesising Residential Solar Generation and Demand Data, IEEE Transactions on Smart Grid 11 (3) (2019) 2511–2519.

[109] Rahimpour, Zahra and Verbic, Gregor and Chapman, Archie C, Can phase change materials in building insulation improve self-consumption of residential rooftop solar? An Australian case study, arXiv preprint arXiv:2111.05524 (2021).

[110] Yu, Liang and Xie, Weiwei and Xie, Di and Zou, Yulong and Zhang, Dengyin and Sun, Zhixin and Zhang, Linghua and Zhang, Yue and Jiang, Tao, Deep Reinforcement Learning for Smart Home Energy Management, IEEE Internet of Things Journal 7 (4) (2020) 2751–2762.

[111] Origin Energy NSW , Residential Energy Price Fact Sheet for Essential Energy Distribution Zone, `https://www.originenergy.com.au` (2020).

[112] Nicholas Morreau, Optimal Scheduling of HVAC Systems with Phase Change Materials using Deep Reinforcement Learning, Bachelor's thesis, The University of Sydney (2020).

Appendix A

Proof of the multi-timescale approach

Here we briefly explain the proof of the multi-timescale approach presented Section 5.2-B. The proof is based on a generalised Bellman equation [22]

$$V = C + \mathbb{P}^T V. \tag{A.1}$$

The model is valid if it satisfies (A.1) for any $\mathbb{P}$ and C, with $\lim_{i\to\infty} \mathbb{P}^i = 0$, where i is the number of the MDPs in the model. For any valid model, we can update the value function through lookahead or backup operation as follows:

$$V_{k+1} = C + \mathbb{P}^T V_k. \tag{A.2}$$

As long as the model is valid, it converges to the same value function regardless of the number of steps:

$$V_\infty = \sum_{i=0}^{\infty} \mathbb{P}^{T^i} C = V \tag{A.3}$$

To prove that the solution of a multi-timescale MDP is the same as the solution of an one-step MDP, we need to prove that an i-step model formulation satisfies the generalised Bellman equation (A.1).

Theorem 6.1: A multi timescale or an n-step model that has a general form (A.4) satisfies the generalised Bellman equation (A.1).

$$C^{(n)} = \sum_{i=0}^{n-1} \left(\mathbb{P}^T\right)^i C, \tag{A.4a}$$

$$C^{(n)^T} s_k = \mathbb{E}(c_k^{(n)} | s_k), \tag{A.4b}$$

where $c_k^{(n)} = \sum_{i=1}^{n} c_{k+i}$ is the n-step truncated return starting from state s_k.

Proof: We combine $\mathbb{P}$ and C and the initial value s_0 into a matrix M:

$$M = \left(\begin{array}{c|c} s_0 & C^T \\ \hline 0 & \mathbb{P} \end{array}\right)$$

If the vector V is also augmented by adding an initial component whose value is always 1, then the generalised Bellman equation (A.1), can be written as

$$V = M^T V. \tag{A.5}$$

Same as before, we consider model M to be valid if and only if it satisfies (A.5). For any valid model M_i, the composed model $\prod_{i=1}^{n} M_i$ is also valid because

$$\prod_{i=1}^{n} (M_i)^T V = \prod_{i=1}^{n} M_i^T V = V \tag{A.6}$$

Note that M has been constructed such that it is valid only if the corresponding $\mathbb{P}$ and C are valid. Therefore, (A.6) proves the validity of the n-step model (A.4). ■