15th sdewes Conference COLOGNE 2020

Towards Coordinated Energy Management in Buildings using Deep Reinforcement Learning

Giuseppe Pinto[a], Silvio Brandi[a], Alfonso Capozzoli[a]
José Ramón Vázquez-Canteli[b], Zoltán Nagy[b]

[a] BAEDA Laboratory, DENERG, Politecnico di Torino, Italy

[b] Intelligent Environments Laboratory, CAEE, The University of Texas at Austin, USA

POLITECNICO DI TORINO
Dipartimento Energia

TEXAS
The University of Texas at Austin

BAEDA Lab
Building Automation Energy Data Analytics

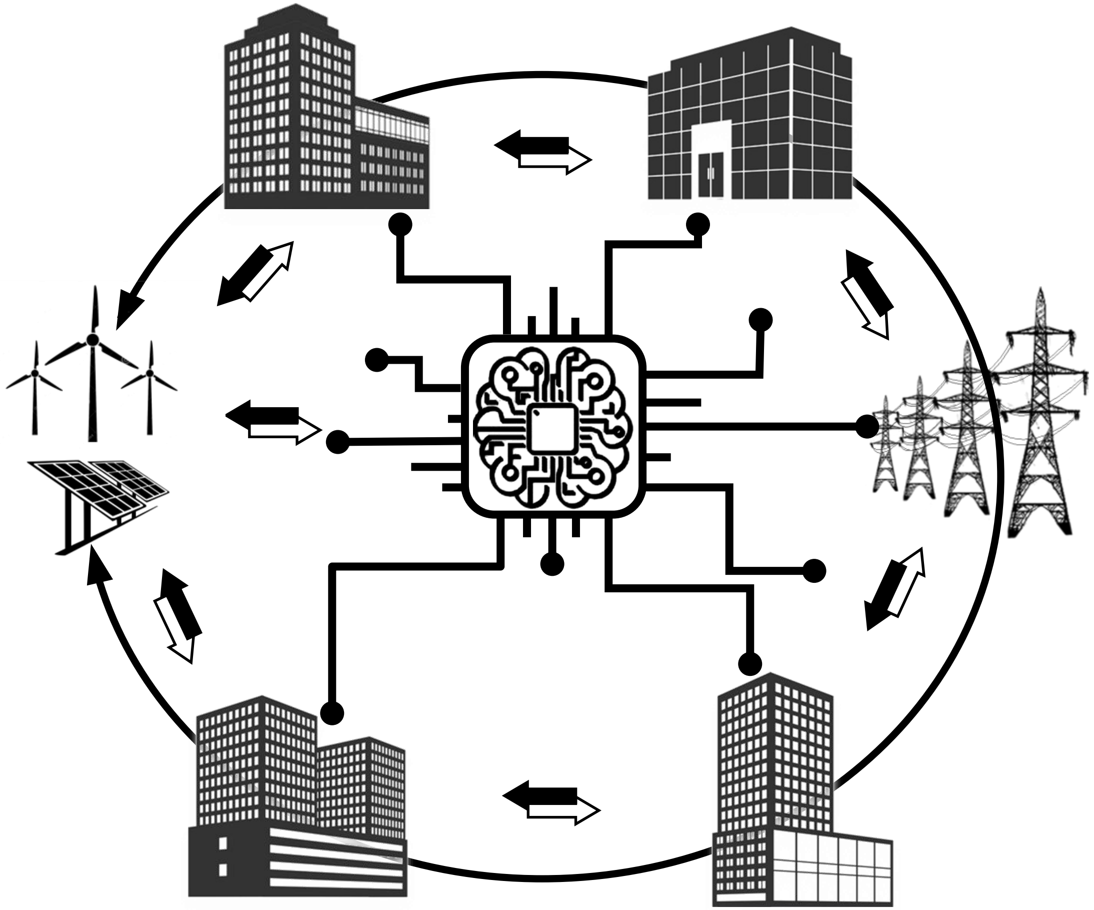
IEL INTELLIGENT ENVIRONMENTS LABORATORY

Problem Statement

1. Increasing penetration of multi-sources HVAC systems, RES and storage in buildings → **Energy flexibility**

2. Control Strategies must be able to handle both users preferences and RES uncertainties → ***Adaptive Control***

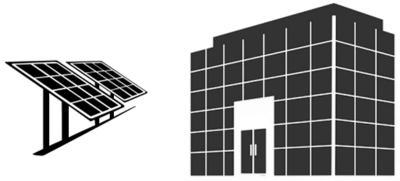
3. Single building management can not always lead to optimal solutions (e.g peak shifting instead of peak shaving) → ***Coordination among buildings***



Coordinated Energy Management Objectives

Single Building Level

Exploit energy flexibility to match production and demand

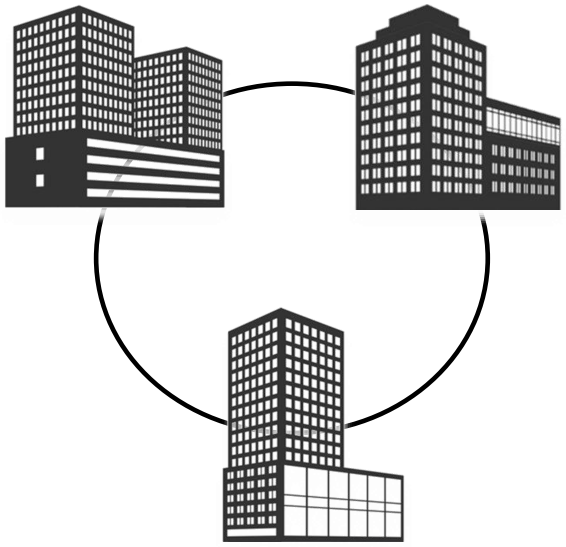


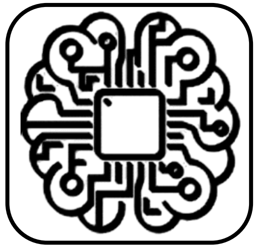
Grid Level

Reduce peak demand and increase grid stability

Cluster Level

Optimise cluster energy consumption



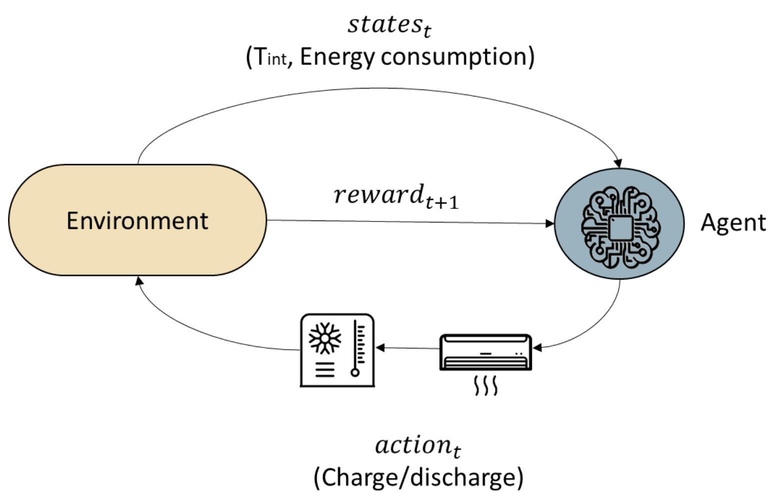


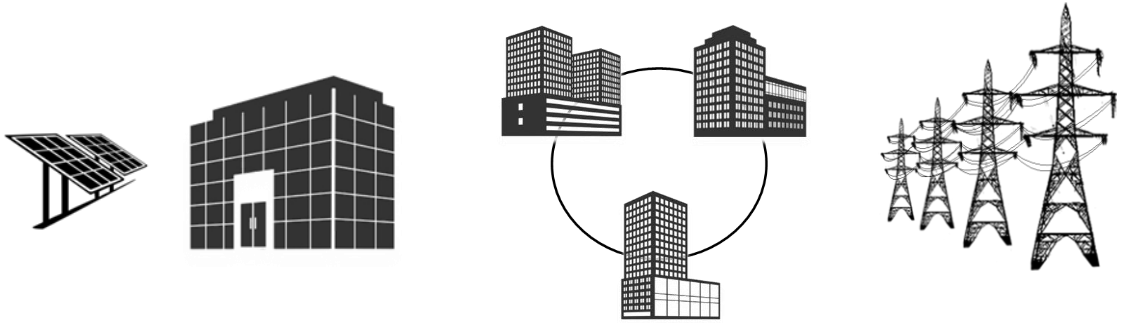
Control Level

Automate the system management according to to the objective function

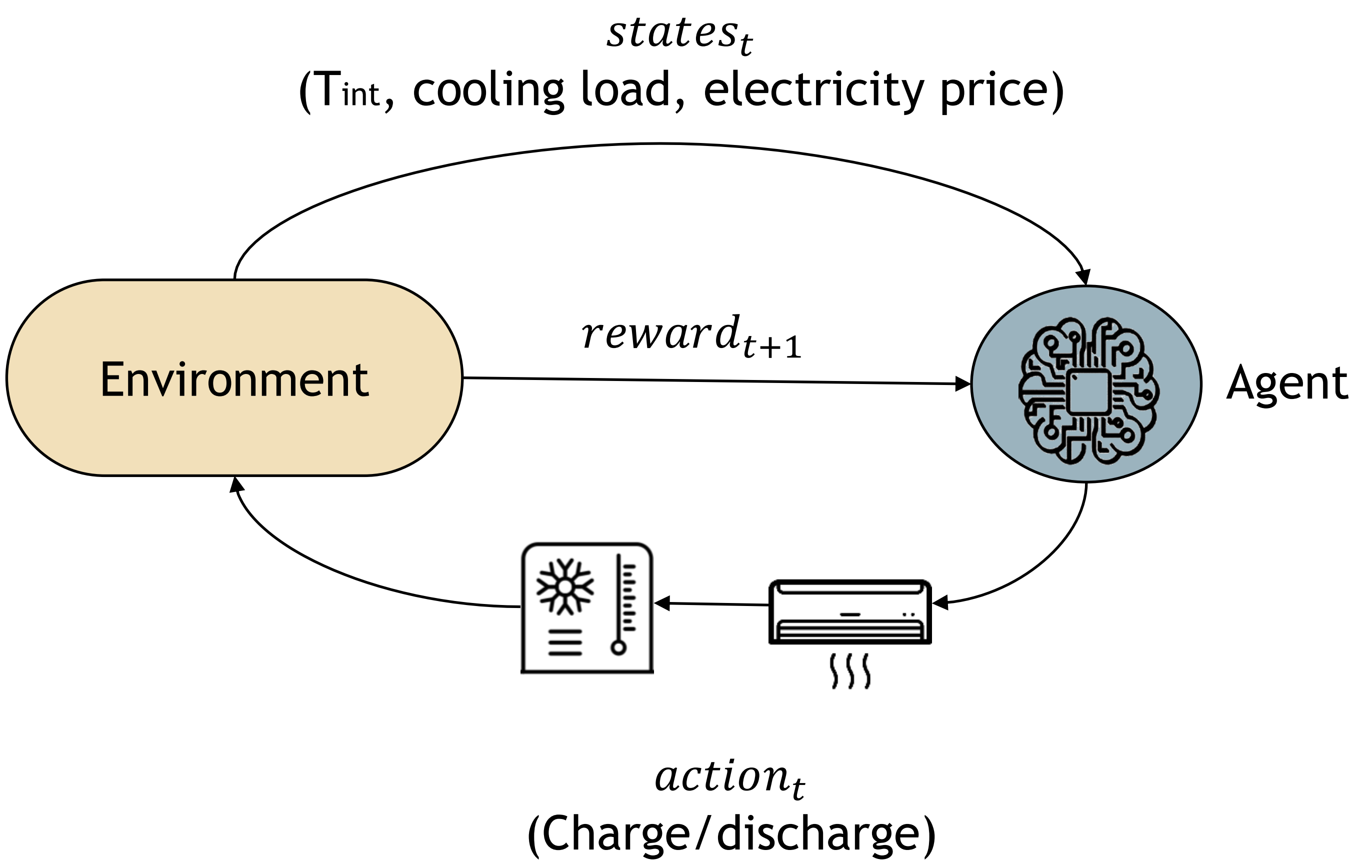
Paper Contribution

- Use of adaptive and model-free control strategy (**Deep Reinforcement Learning**) with multiple buildings
- Use of a novel simulation environment to test Reinforcement Learning control strategies in the built environment (**CityLearn**)
- Assessment of the potential **benefits of a coordinated approach** at 3 levels:
  - Single Building Level
  - Cluster Level
  - Grid Level





Methods: Reinforcement Learning



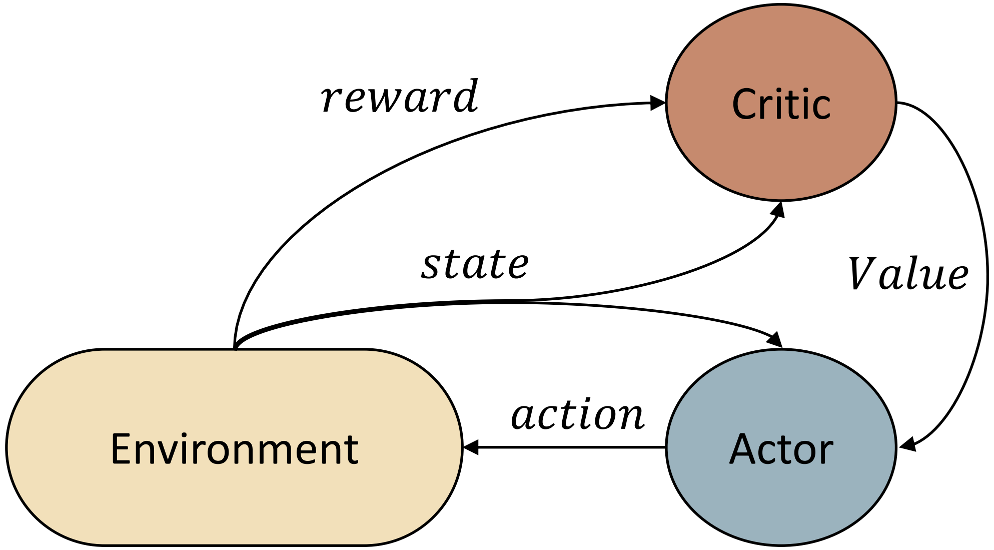
- RL agent learns a policy through a trial-and-error interaction with the environment

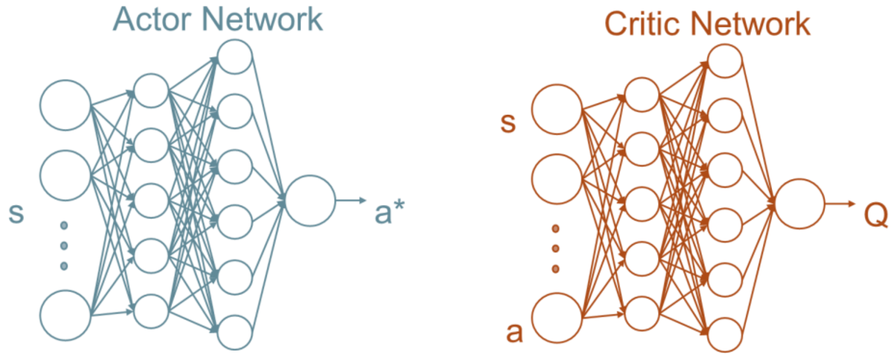
- The problem can be defined by two functions, namely:
  - State-value function $v_\pi(s)$
  - Action-value function $q_\pi(s, a)$

- These functions are used to show the expected return of a control policy π

Goal: find the optimal control policy π to maximise the return (sum of rewards)

From Reinforcement Learning to Deep Reinforcement Learning





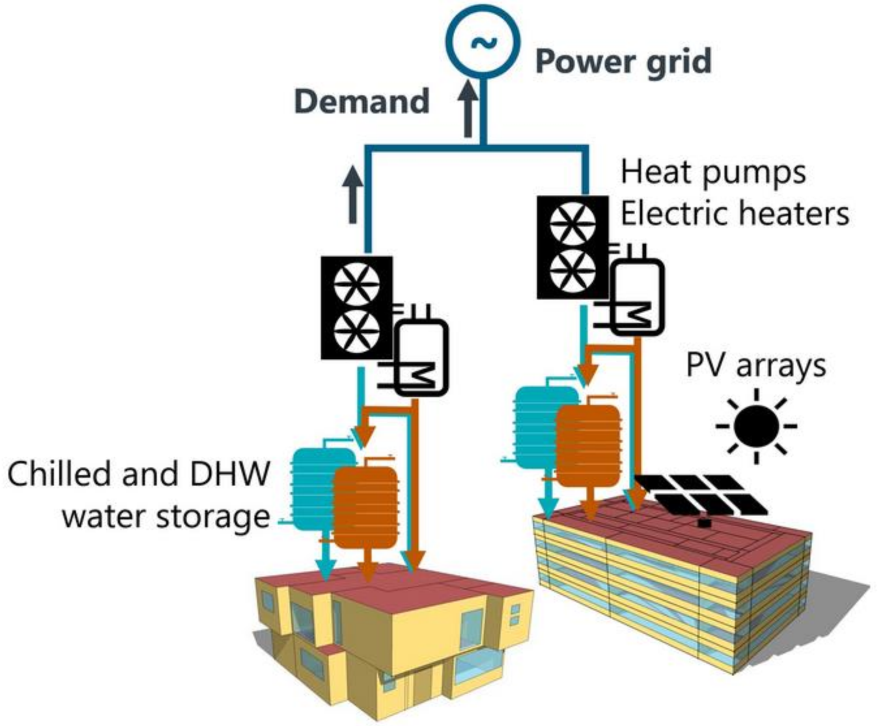
Soft Actor-Critic Algorithm

- Exploits Deep Neural Networks and Actor-Critic Architecture
- Optimises policies to maximise both expected return and expected entropy of the policy

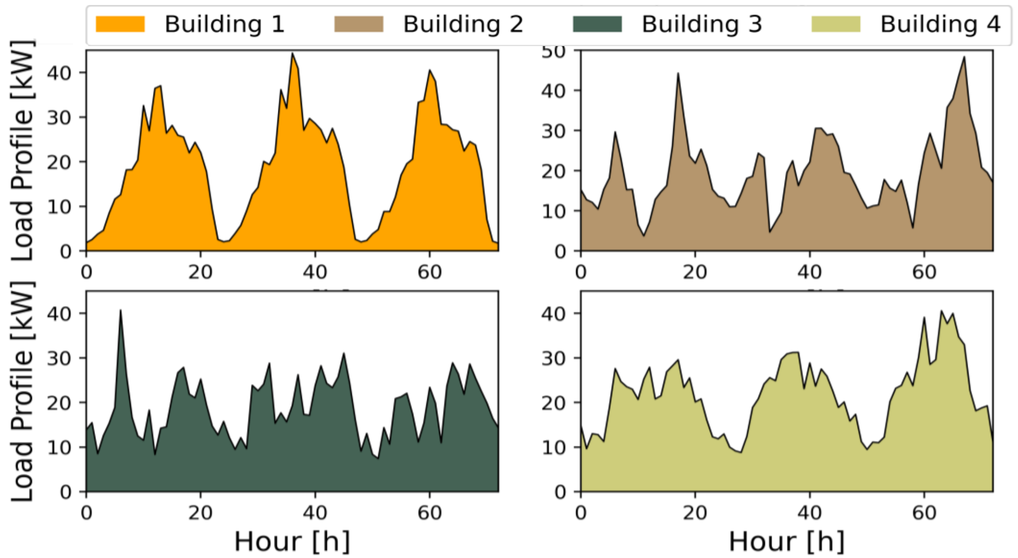
$$\pi^* = argmax_{\pi}\mathbb{E}_{\pi}[\sum_{t=0}^{\infty}\gamma^t(r_t + \alpha\mathcal{H}_t^{\pi})]$$

Case Study: CityLearn environment

The CityLearn Challenge
MULTI-AGENT REINFORCEMENT LEARNING FOR INTELLIGENT ENERGY MANAGEMENT



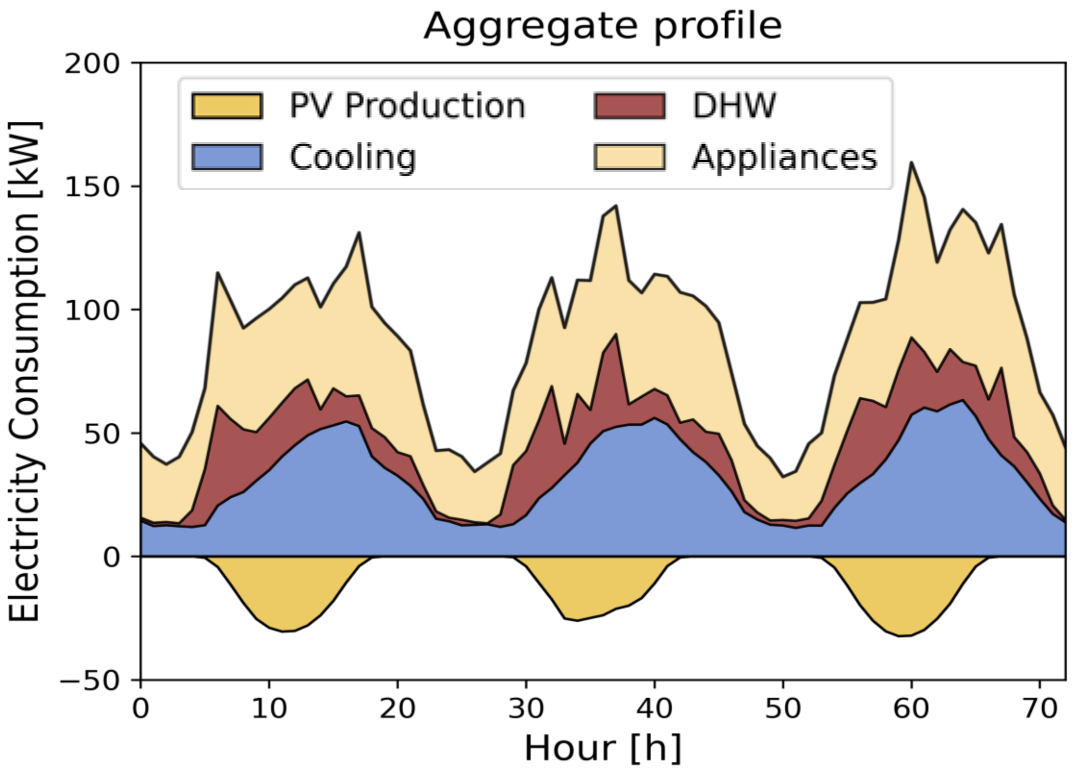
	Type	Surface $[m^2]$	Cold Storage Capacity [kWh]	Hot Storage Capacity [kWh]	PV Capacity [kW]
Building 1	Office	5000	235	50	0
Building 2	Restaurant	230	150	75	25
Building 3	Retail	2300	200	70	20
Building 4	Retail	2100	185	105	0



CityLearn enhancement:

- Variable electricity price
- Detailed modeling of heat pump

Case Study: Control Strategies



	Cold Storage Capacity [kWh]	Hot Storage Capacity [kWh]	PV Capacity [kW]
Cluster	770	300	45

$$R = -\beta * \sum_{i=1}^{n} e_i^2 - \rho * \sum_{i=1}^{n} (c_{el} * e_i)$$

e_i^2 → Peak term

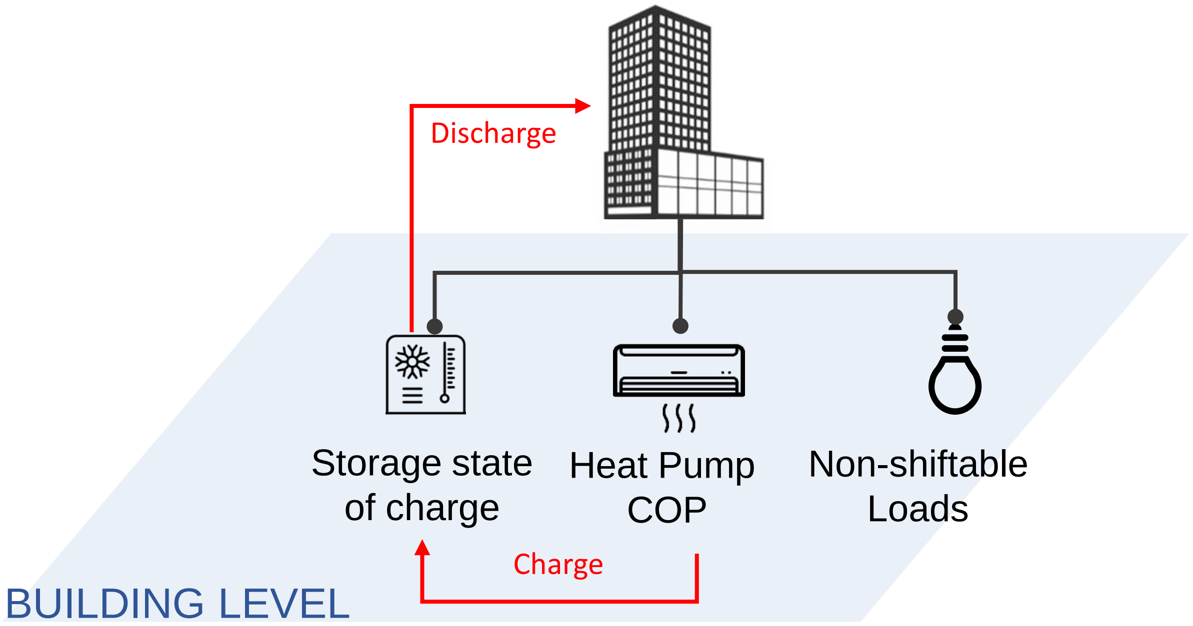
$c_{el} * e_i$ → Cost term

β & ρ → Relative weights

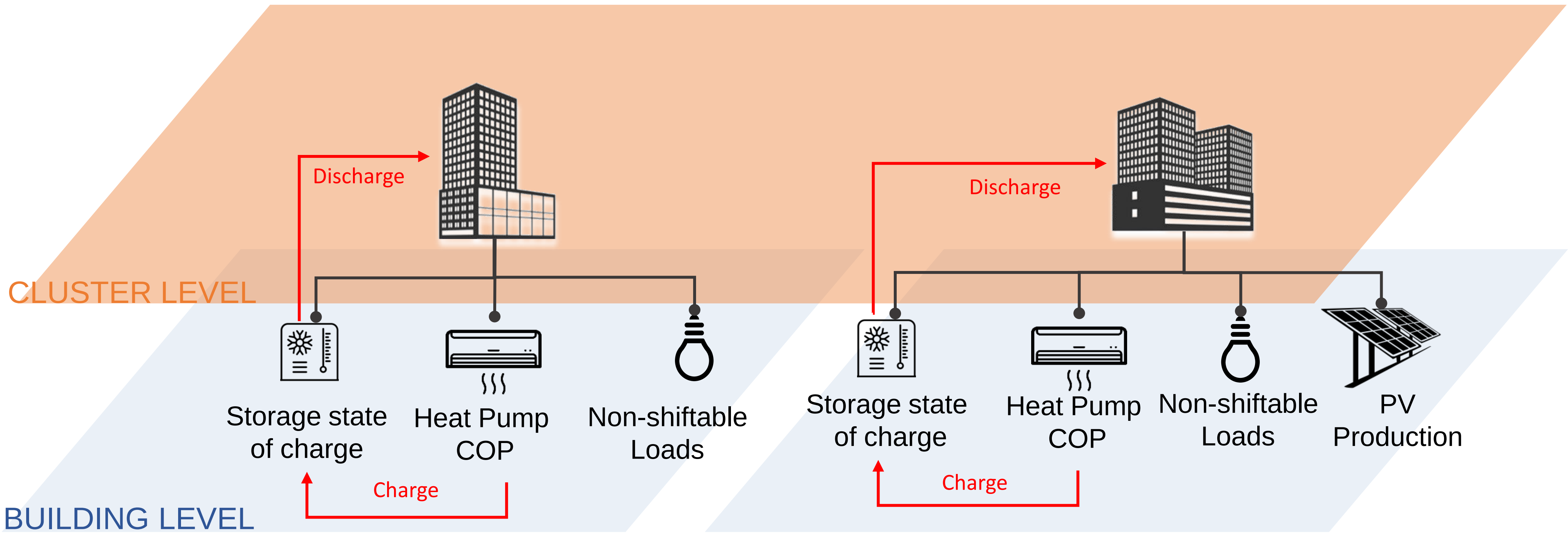
Baseline → Manually optimised RBC: take full advantage of variable electricity price

DRL Controller → Maximise the reward function

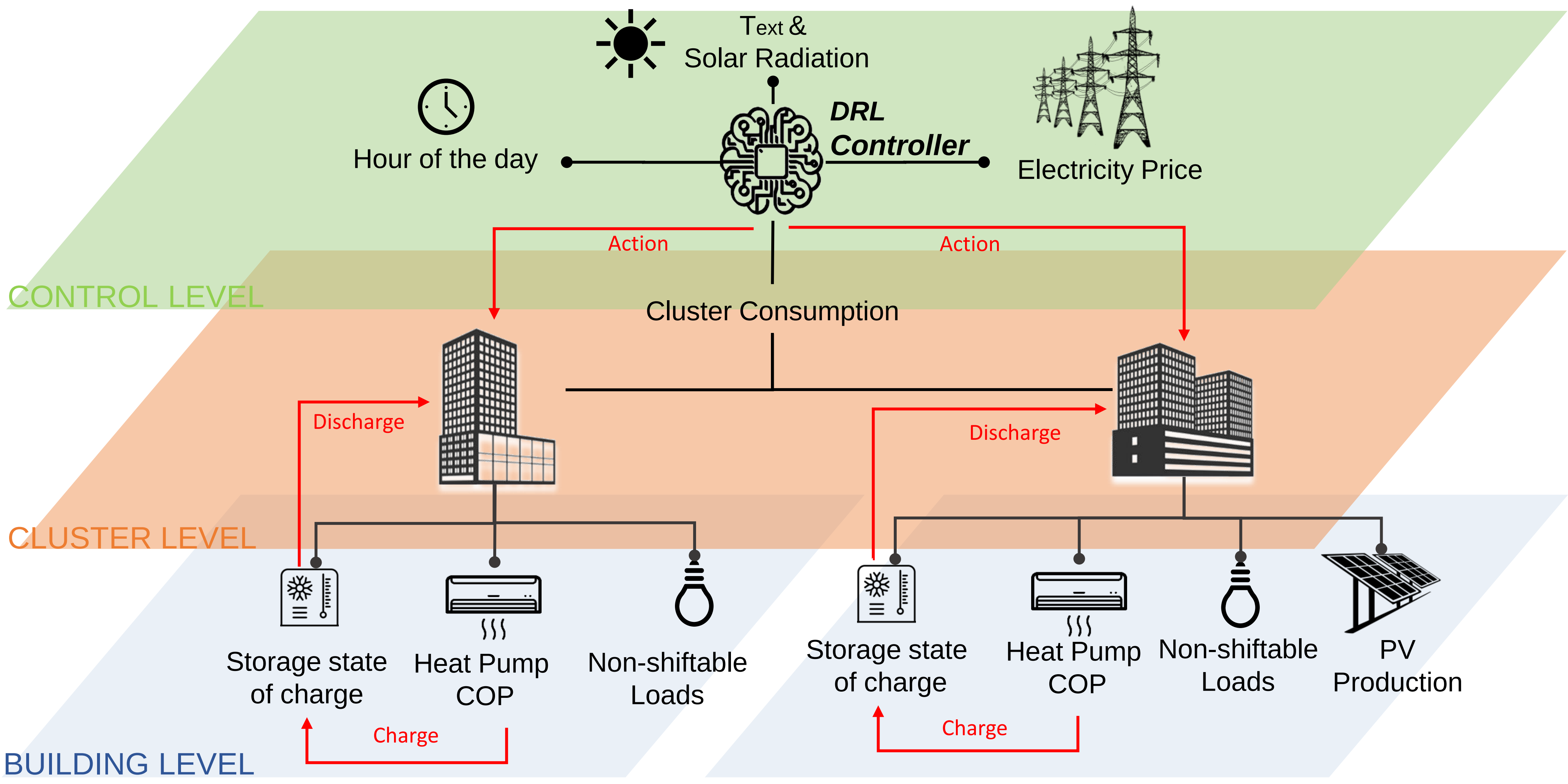
Framework


Discharge
Storage state
of charge
Heat Pump
COP
Non-shiftable
Loads
Charge
BUILDING LEVEL

Framework


CLUSTER LEVEL
BUILDING LEVEL
Discharge
Storage state of charge
Heat Pump COP
Non-shiftable Loads
Charge
Discharge
Storage state of charge
Heat Pump COP
Non-shiftable Loads
PV Production
Charge

Framework

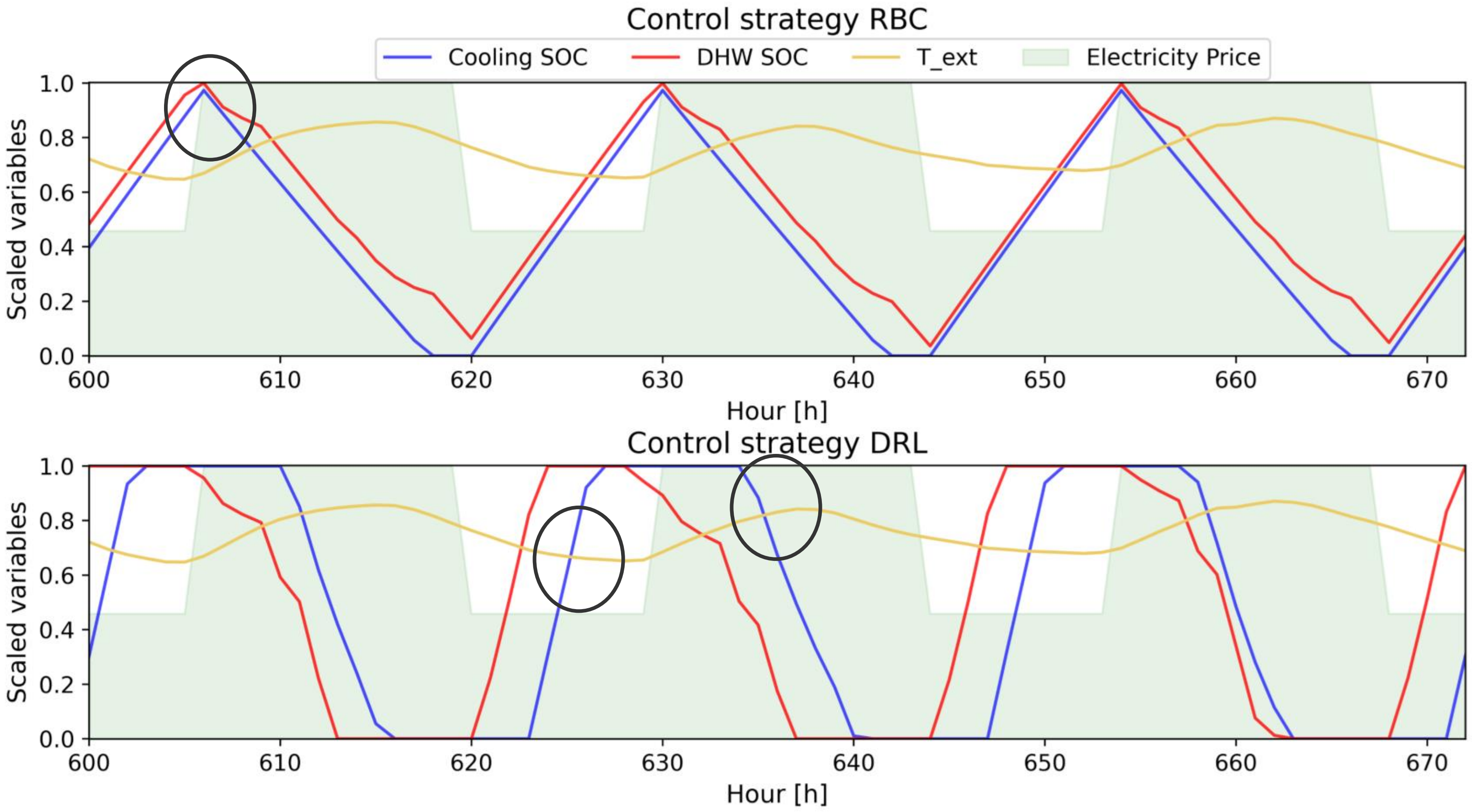


Text &
Solar Radiation
DRL
Controller
Hour of the day
Electricity Price
Action
Action
CONTROL LEVEL
Cluster Consumption
Discharge
Discharge
CLUSTER LEVEL
Storage state
of charge
Heat Pump
COP
Non-shiftable
Loads
Storage state
of charge
Heat Pump
COP
Non-shiftable
Loads
PV
Production
Charge
Charge
BUILDING LEVEL

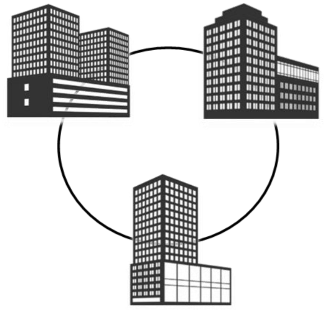
Single Building Level

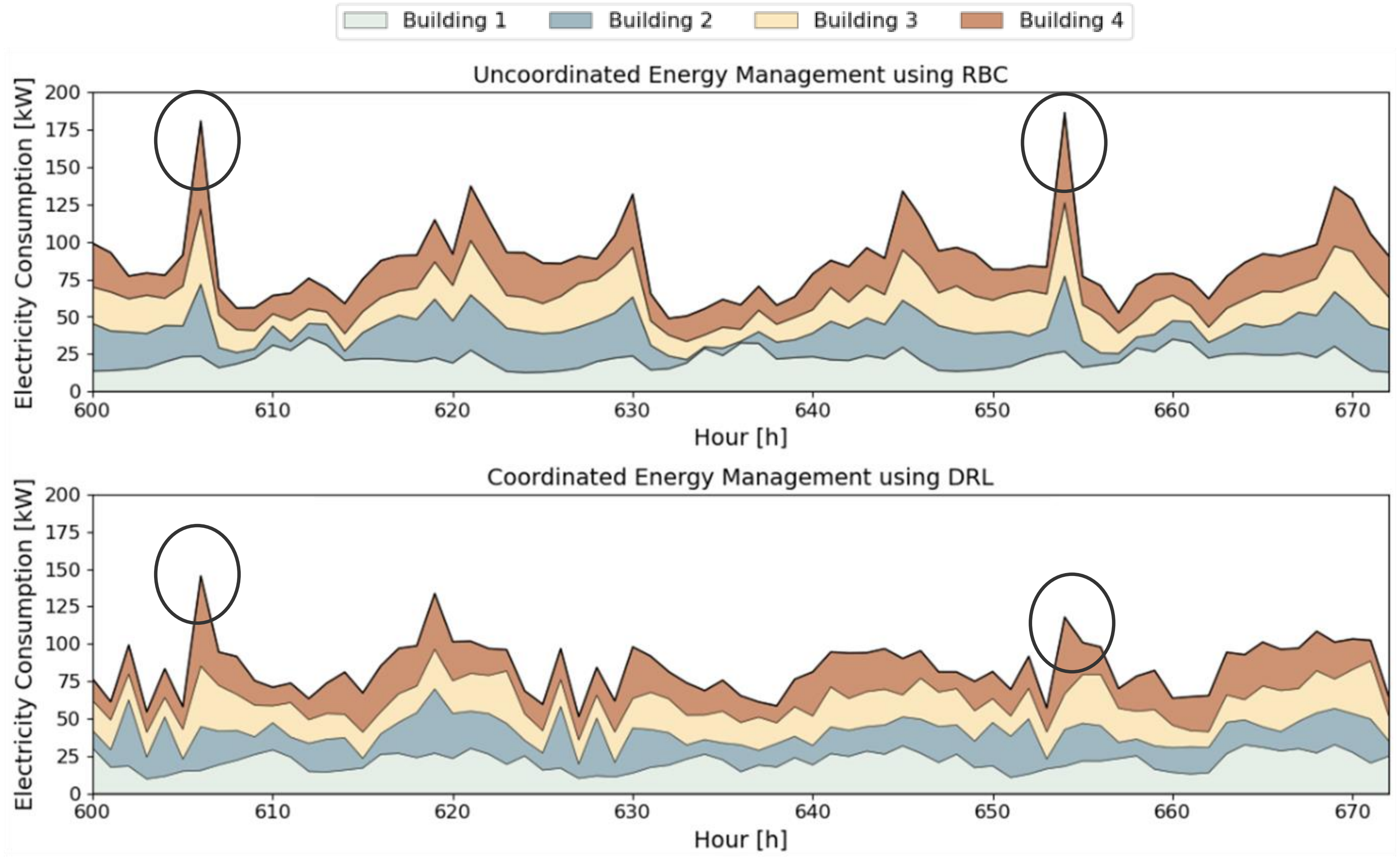
DRL controller automatically learned:

- Electricity price dependency → Cost reduction
- Heat pump temperature dependency → Energy reduction
- Charge/discharge shifting among the two storages → Peak reduction at single building level



Cluster Level



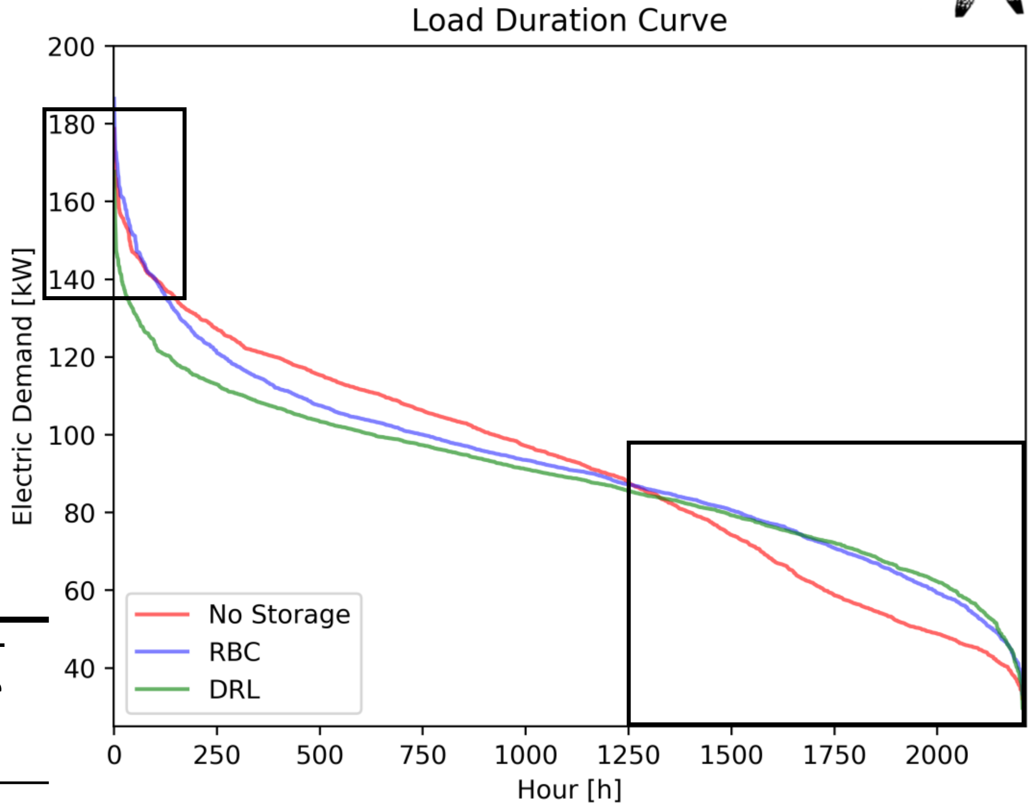


- Shifting charge/discharge process among buildings → Peak reduction at cluster level
- More homogeneous electricity consumption over the day → Renewable integration

Grid Level

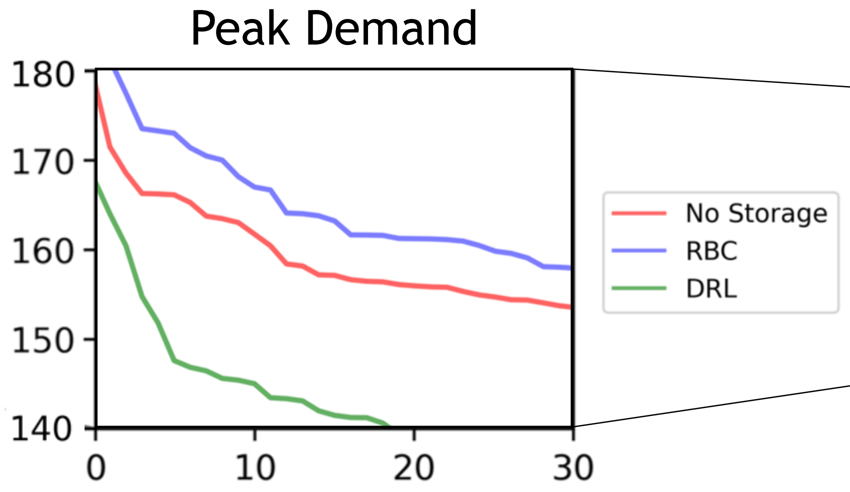
- Storages increase base load in both configurations (RBC and DRL)

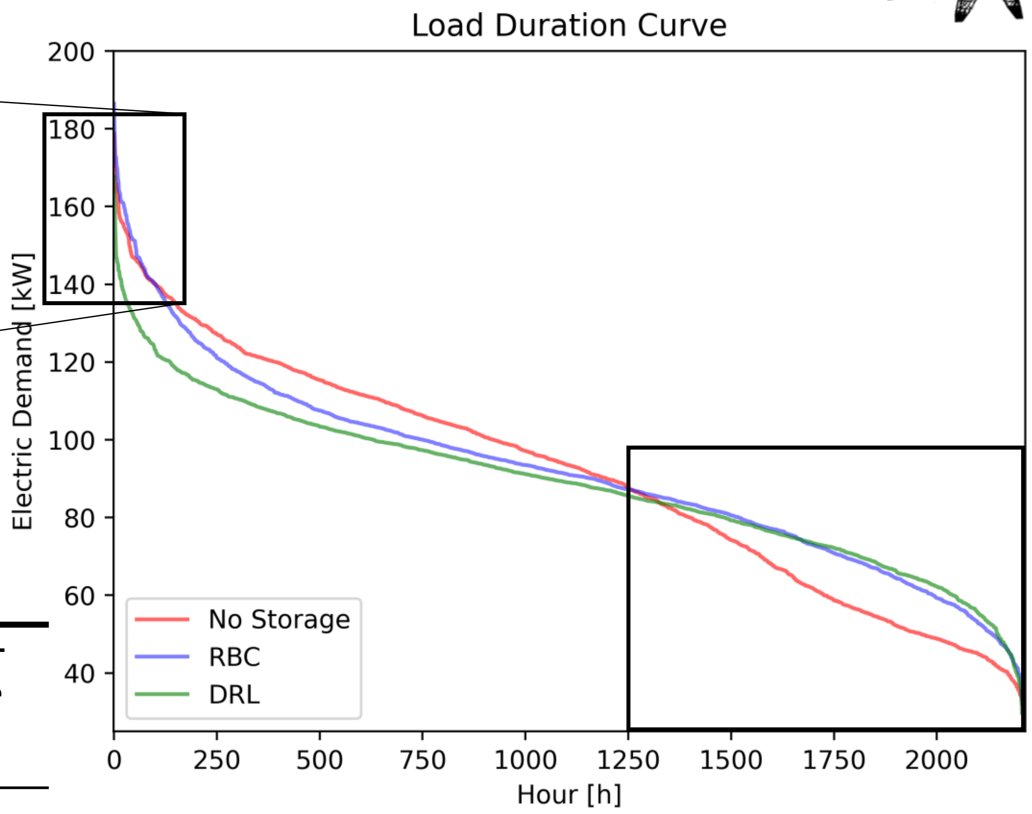
- Control strategies influence peak demand → Double-edged sword



	Energy Consumption	Electricity Cost	Maximum Peak	Average daily peak	Peak-to-average ratio
Manually Optimised RBC	1	1	1	1	1
DRL	0.96	0.99	0.92	0.84	0.95

Results at Grid Level





	Energy Consumption	Electricity Cost	Maximum Peak	Average daily peak	Peak-to-average ratio
Manually Optimised RBC	1	1	1	1	1
DRL	0.96	0.99	0.92	0.84	0.95

Conclusion

“The whole is greater than the sum of its parts”

-Aristotle-

The presented paper exploits model-free RL controller to coordinate the energy management of a cluster of buildings.

“What you measure affects what you do”

-Joseph Stiglitz-

The design of the reward function plays a key role for the DRL controller behaviour. It is necessary to find an optimal trade-off between the advantages of single users and cluster.

“The measure of intelligence is the ability to change”

-Albert Einstein-

The strength of DRL approach is the opportunity provided by its adaptive nature. A large environment involves rapid changes, that requires rapid (and smart) decisions.

Future works

- Analyse the effect of the reward function definition on scalability of the process
- Compare several DRL architectures such as Single-Agent, Multi-agent and Hierarchical Reinforcement Learning
- Analyse the energy exchange among multiple buildings in configuration such as Energy Community or P2P energy trading and the resulting changes in control strategy
- Introduce comfort metrics in CityLearn environment using black-box models

Alfonso Capozzoli
Associate professor at PoliTo

Silvio Brandi
Energy Engineer - PhD Student

Giuseppe Pinto
Energy Engineer - PhD Student

Zoltan Nagy
Assistant Professor at UT Austin

José R. Vázquez-Canteli
Civil Engineer - PhD Student

POLITECNICO DI TORINO
Dipartimento Energia

TEXAS
The University of Texas at Austin

BAEDA Lab
Building Automation Energy Data Analytics

IEL
INTELLIGENT ENVIRONMENTS LABORATORY

15th sdewes Conference COLOGNE 2020

Towards Coordinated Energy Management in Buildings using Deep Reinforcement Learning

Giuseppe Pinto[a], Silvio Brandi[a], Alfonso Capozzoli[a]
José Ramón Vázquez-Canteli[b], Zoltán Nagy[b]

E-mail: alfonso.capozzoli@polito.it
Or: giuseppe-pinto@polito.it

THANKS FOR THE ATTENTION

POLITECNICO DI TORINO
Dipartimento Energia

TEXAS
The University of Texas at Austin

BAEDA Lab
Building Automation Energy Data Analytics

IEL
INTELLIGENT ENVIRONMENTS LABORATORY