Politecnico di Torino

Department of Energy "G.Ferraris"

ScuDo

Scuola di Dottorato - Doctoral School

WHAT YOU ARE, TAKES YOU FAR

Optimizing the operation of an integrated energy system in an office building by means of artificial intelligence-based algorithm

PhD III Level course: Optimization methods for engineering problems

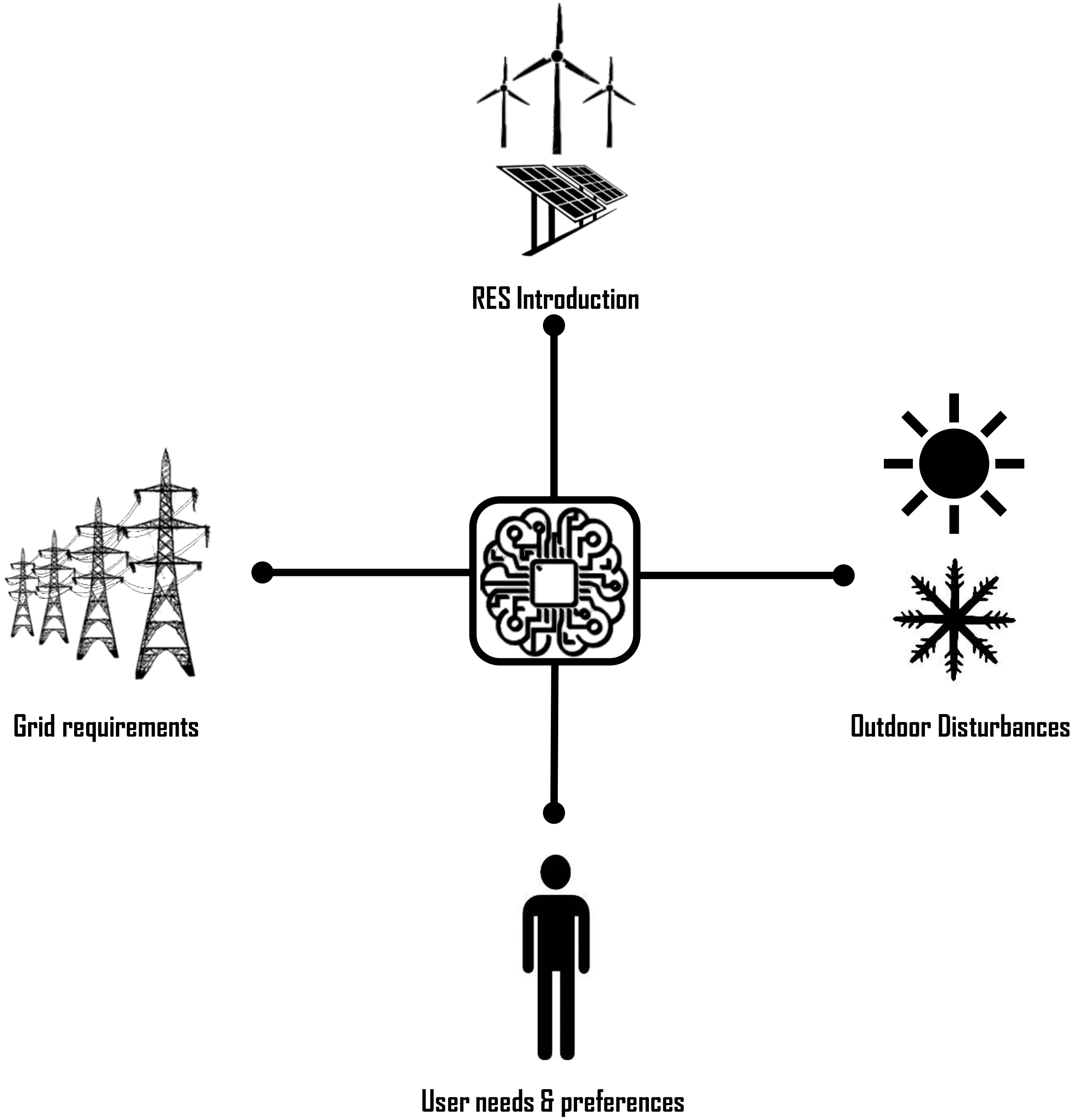
Davide Coraci, PhD student in Energetics
Antonio Gallo, PhD student in Energetics
31 May 2022

Introduction - Problem Statement

1. The increasing penetration of HVAC systems, the introduction of RES and storage has changed the framework of building energy managing → **Energy Flexibility**

2. Classic control strategies (i.e., ON/OFF or PID) are inadequate to adapt to continually changing of preferences of users, grid requirements and disturbances → **Adaptive control**

3. Model-based control strategies (e.g., Model Predictive Control (MPC)) were explored, showing excellent ability in improving comfort conditions and energy efficiency in buildings. However, their application is limited since requires the definition of an accurate model of the environment to be controlled → **Model-free and data-driven control strategies**

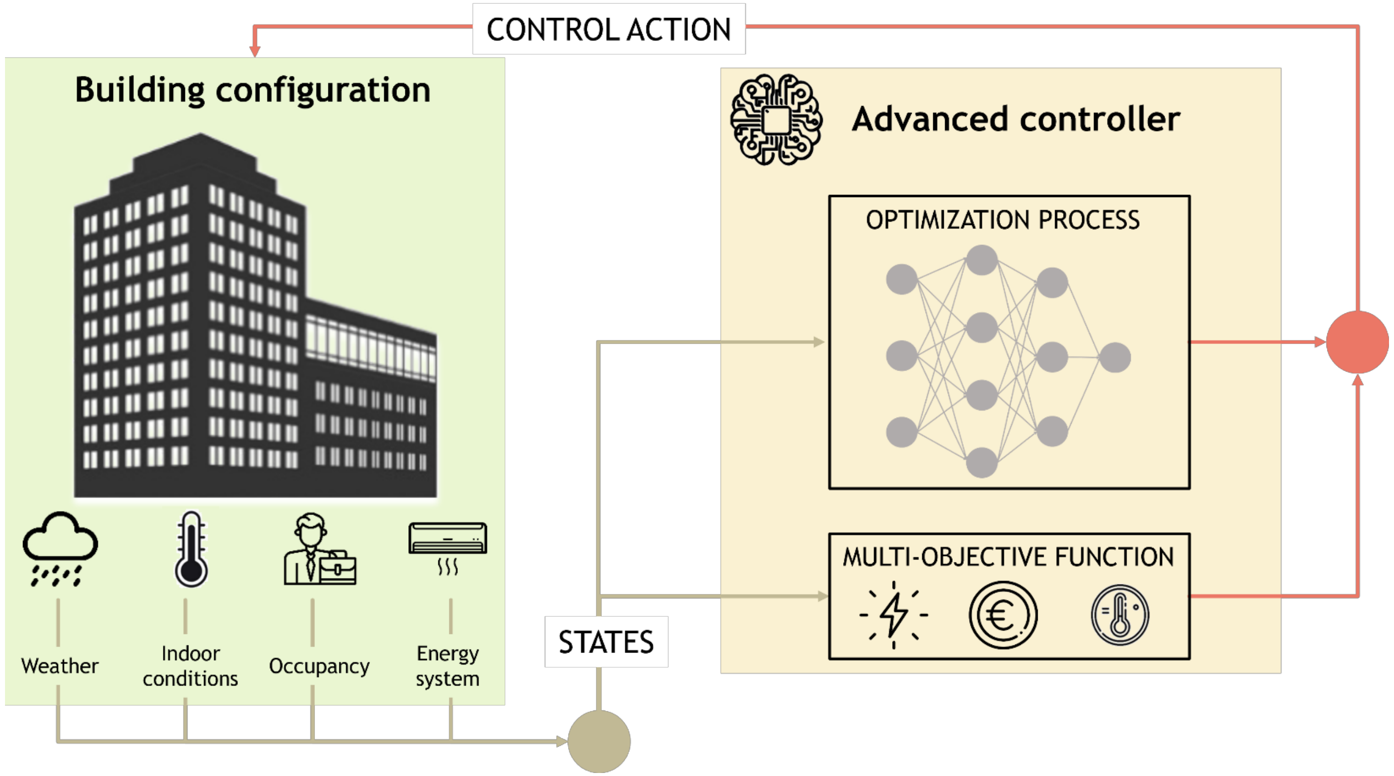


Introduction – Problem Statement

Given their data-driven nature, model-free advanced control strategies based on Artificial Intelligence performs well due to their ability to **adapt** and to **optimise** the operation of complex energy systems.

Furthermore, such control strategies can handle **multi-objective problems**, ensuring an optimal trade-off when conflicting objectives are involved, such as minimize energy consumption while enhancing indoor comfort conditions in buildings.

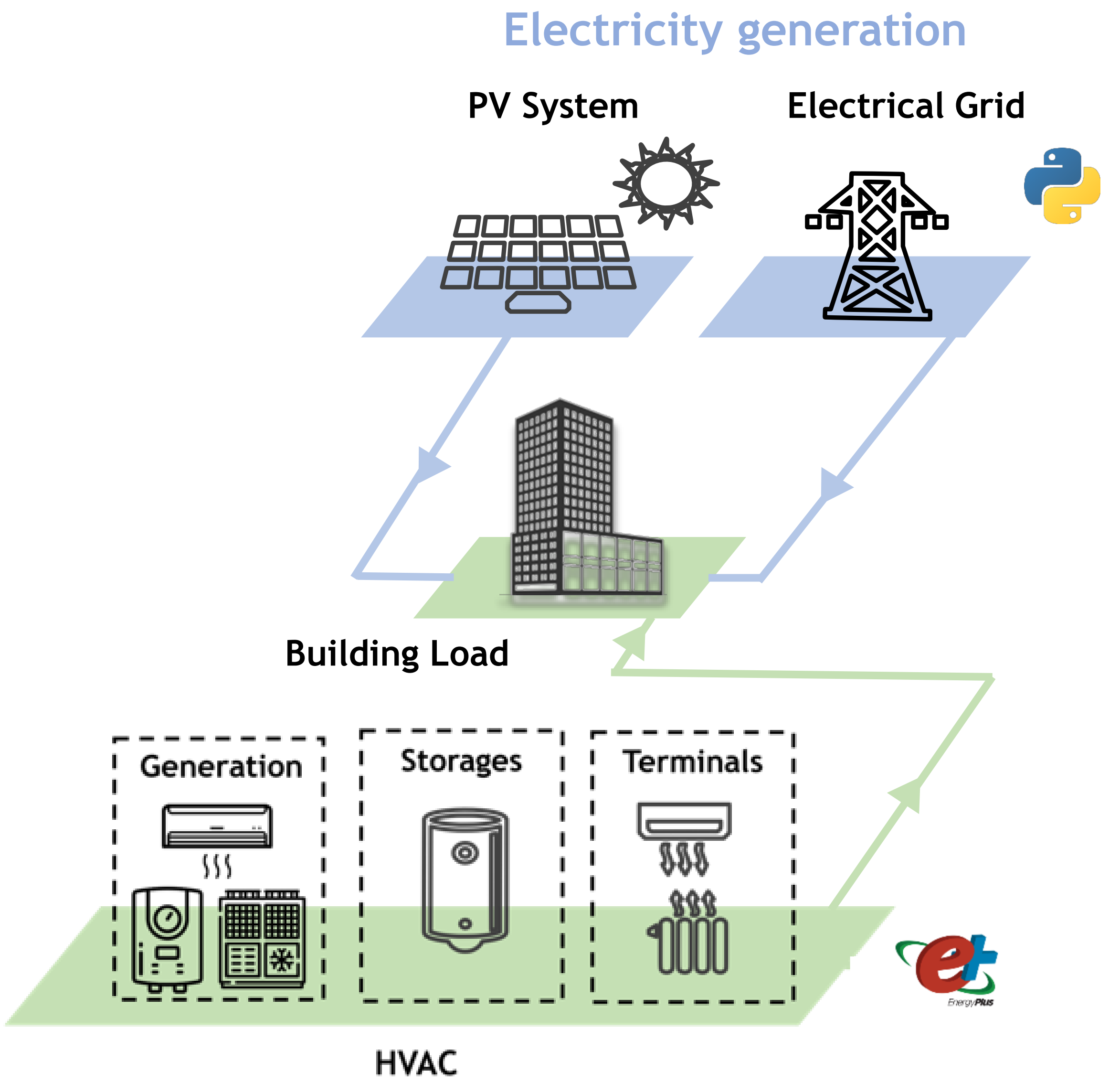
In this work it was designed an adaptive and model-free strategy, based on **Deep Reinforcement Learning.**



Work contribution

- **Application of an adaptive and model-free control strategy (Reinforcement Learning)** to control and optimize the operation mode of the HVAC system (chiller mode, TES charge/discharge) to:
  1. enhance building comfort condition;
  2. reduce costs associated with electricity withdrawn from grid;
  3. reduce daily peak load demand.

- **Perform an automated optimization process** for DRL hyperparameters **using** the Python library **Optuna.**



Methods: Reinforcement Learning

It is a **learning technique** that aims at realizing control agents.

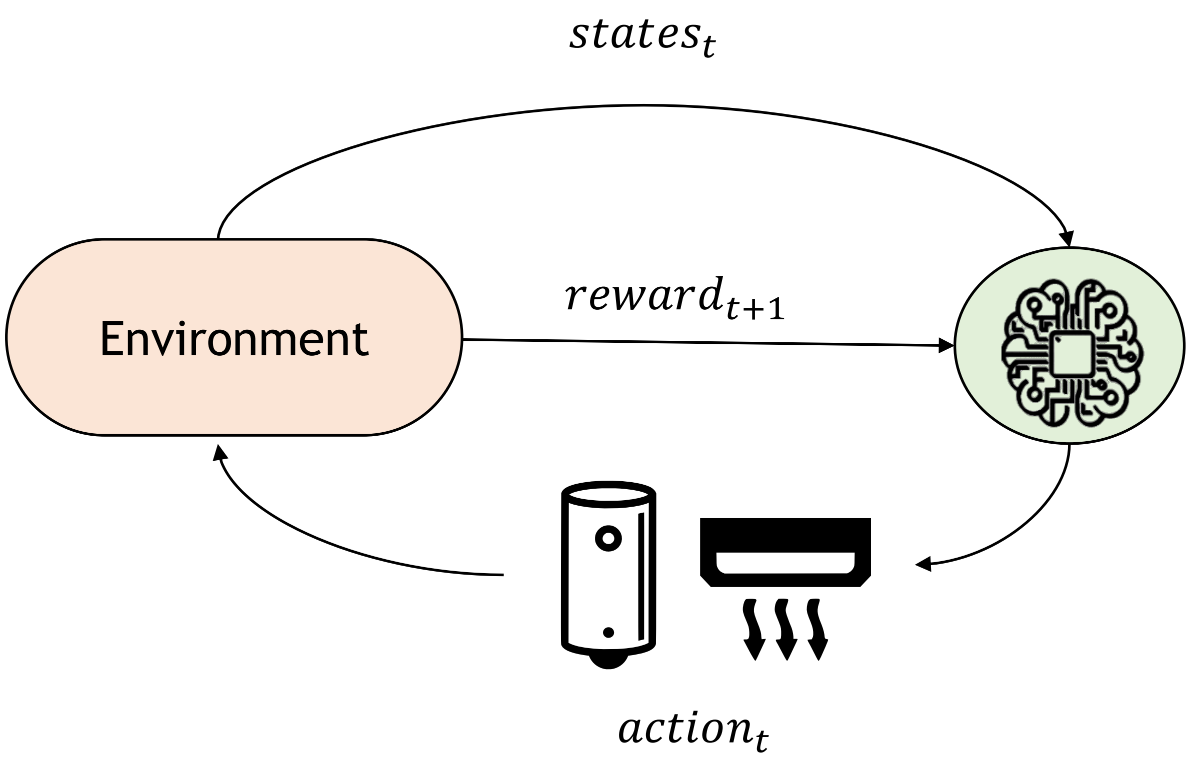
RL agent is trained through trial-and-error interaction with the environment to learn an **optimal control policy π** that maximizes the objective function, called **reward function.**

Two functions are used to define the problem and show the expected return of the control policy:

- State-value function $v_\pi(s)$
- Action-value function $q_\pi(s, a)$

Usually, state-value and action-value functions are approximated with **neural networks → Deep Reinforcement Learning.**

However, the **performances** of DRL algorithms relies heavily on the choice of several features (i.e., hyperparameters), that require an **optimization process** that is carried out through a **trial-and-error process** or in **automated way**.

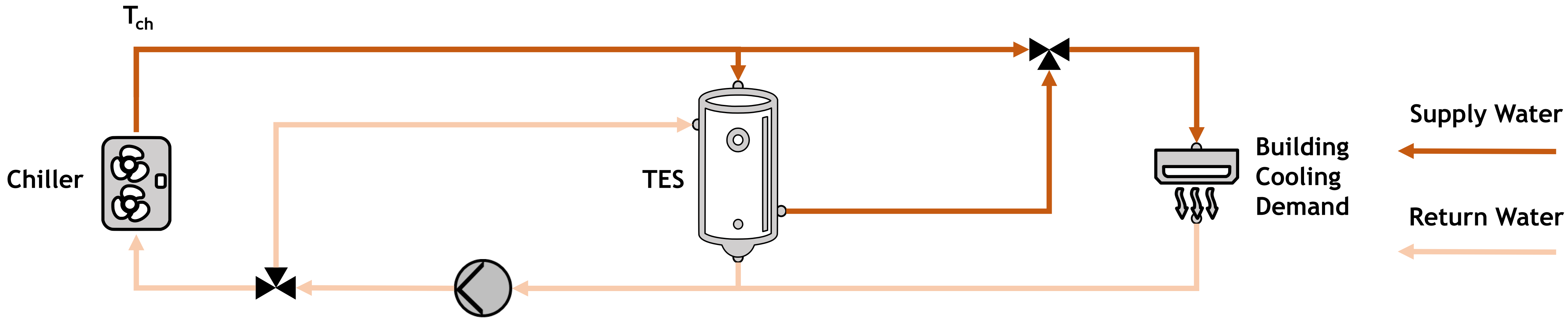


Case study

The building under observation is located in Turin, Italy, and it is divided into three heating zones with 10 workers each plus an unconditioned technical room.

The thermal side of energy system and the building were modeled in **EnergyPlus**, and it was developed a **BCVTB** to allow the interaction with the RBC/DRL controller developed in **Python,** where the **controllers and PV system** (1.2 kW size) **were developed.** The **simulation period** was limited between June and August since it was considered a cooling system.

The price of electrical energy drawn from the grid to operate the energy system is based on a Time-Of-Use (TOU) tariff structure commonly implemented in Italy **(low price** 0,071 €/kWh, **medium price** 0,143 €/kWh, **high price** 0,214 €/kWh).



Modified from: Gallo et al.. 2022

Methodology

The methodological framework of this work is composed by three modules.

RL optimization

Design RL controller
- Define state space
- Define action space
- Define reward function

Automated hyperparameter optimization
- Define objectives for Optuna

Train RL controller
- Update parameters ∂_i

Rule-Based control

Design RBC
- Implementation of If-then rules
- Tuning of threshold for heating system scheduling

Virtual environment

Performance analysis

Build the Pareto front
- Find non-dominated solutions
- Identification of optimal solution

Performance assessment
- Testing of RL controller
- Benchmarking of the RL controller with RBC

Hyperparameters optimization for advanced controller

DRL controllers are characterized by many **hyperparameters** which require appropriate tuning, since they **affect performances of these algorithms**. Therefore, in this project, an **optimization** was carried out on some of the most important DRL hyperparameters by using **Optuna**.

Optuna is an **open-source Python library that automates the search of optimal hyperparameter configuration** in machine learning-based models. This library ensures efficient optimization of hyperparameters by adopting state-of-the-art algorithms for hyperparameter sampling.

Optuna employs records of recommended parameter values and evaluated target values to restrict and optimize the search space of the hyperparameters.

Within the sampling algorithms used in Optuna, in this work, it was chosen the **Tree-structured Parzen Estimator (TPE)**. TPE is based on an iterative process that employs the historical record of evaluated hyperparameters and metrics obtained to create a probabilistic model, which is used to recommend the new hyperparameter set.

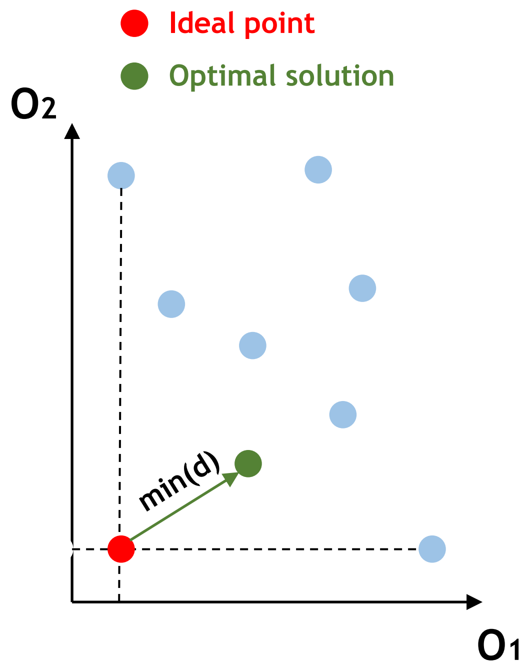
Hyperparameters optimization for advanced controller

Furthermore, Optuna requires as input the **target function to be minimized or maximized as well as the hyperparameter optimization range with the associated incremental step.**

Since we are dealing with a **multi-objective control problem**, it will exist a set of optimal solutions, called Pareto-optimal solutions, ensuring the trade-off between the chosen control objectives (i.e., temperature, efficiency and peak shaving) → **it was necessary to establish a criterion to choose the best solution among the optimal ones.**

Therefore, it was adopted the **criterion of the minimum distance from the so-called ideal point**, i.e., the **point whose coordinates correspond to the minimum of both objective function terms.** Then, it was computed the distance between points corresponding to each Optuna solution and the ideal point, whose coordinates are like [E_{cons}, T_{viol}, $P_{daily,peak}$].

- E_{cons} represents the total energy consumption, calculated in MWh.
- T_{viol} stands as the total temperature violations, calculated in °C. The temperature violations were calculated as the absolute difference between the indoor temperature and the lower or upper limit of the temperature acceptability range [25, 27], when the indoor temperature was lower or higher than these limits.
- $P_{daily,peak}$ represents the daily peak power, evaluated as the maximum of the hourly energy withdrawn from the grid.



Hyperparameters optimization for advanced controller

To better explain, this table reports the values of hyperparameters kept fixed as well as the range for those optimized with Optuna.

DRL Hyperparameters	Value
DNN architecture	2 layers
Episode length	90 days
Boltzmann temperature coefficient (α)	0.1
Discount factor	[0.9, 0.95, 0.99]
Learning rate	[0.001, 0.002, ..., 0.005]
β	[0.01, 0.015, 0.02, ..., 0.05]
δ	[0.5, 1, 1.5, ..., 5]
θ	[5, 6, 7, ..., 15]
Batch size	128
Neurons per hidden layer	256
Number of episodes	30

RB controller

The performances of DRL controller were compared with a baseline controller, made of **two agents** that decide the operation mode **(RBC-1)** and if providing cooling energy to the building **(RBC-2)**

The **RBC-1** operates the cooling system in different operation mode (**chiller mode** or **charging/discharging TES**) depending on the electricity **price (low/high) and TES temperature.**

- TES was charged if **the electricity price is low** and until the SOC is max (=1)
- TES was **discharged** whenever it is **supplied cooling energy to the building** and **if the TES SOC is not zero**
- The thermal system operates in **chiller mode** whenever it is **supplied cooling energy to the building** and **if the TES SOC is zero**

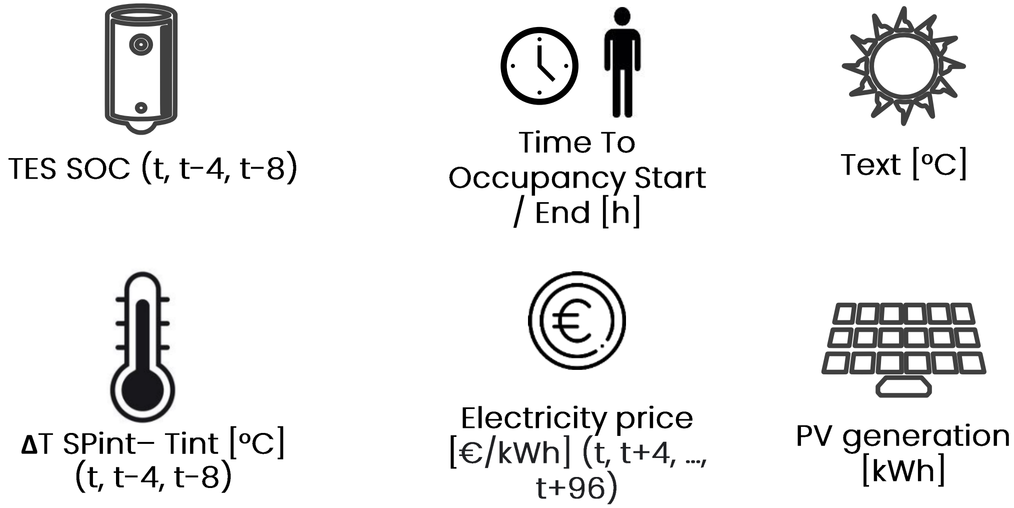
The RBC-2 control logic consists of **two parts**, a **pre and post first switch ON phase**, where the agent starts to supply cooling energy to the building according to specific **indoor temperature conditions and to the time of the day**. The energy supply is interrupted **when occupants leave the building.**

Design of DRL Controller

For the source building, the goal of the designed controller is to **optimize control of the indoor temperature** during the occupancy period while **reducing the electricity bills (chiller + pump energy) and shaving the energy peaks.** The setpoint was set equal to 26 °C and the temperature acceptability ranges between 25 °C and 27 °C.

The **action-space** is discrete, and it includes 5 couple of action indicated as [system operation mode, cooling fraction].

The **state-space** is composed in this way



Design of DRL Controller

The reward that the agent receives after having taken actions at each control time step depends on three values, that are weighted introducing coefficients β, δ, θ :

- energy-related term ($\boldsymbol{r_E}$)
- peak-related term ($\boldsymbol{r_P}$)
- Temperature-related term ($\boldsymbol{r_T}$)

The **energy-related term** refers to the electricity withdrawn from the grid for the chiller and pump, considering the electricity prices from TOU schedule

$$\boldsymbol{r_E = c_E * (E_{CHILLER} + E_{PUMP})}$$

The **peak-related term** refers to the mean daily peak power withdrawn from the grid for the chiller and pump

$$\boldsymbol{r_P = P_{PEAK,DAILY}}$$

The **temperature-related term** has different expressions depending on the indoor temperature values, computed or not as the difference between upper/threshold temperature and the temperature setpoint during occupancy periods

$$\boldsymbol{r_T = f(\Delta T)}$$

The general expression of the reward is

$$\boldsymbol{r = -(\delta * r_E + \beta * r_T + \theta * r_P)}$$

Results – Hyperparameters optimization

The **2D-Pareto** fronts are reported below together with the optimal solution

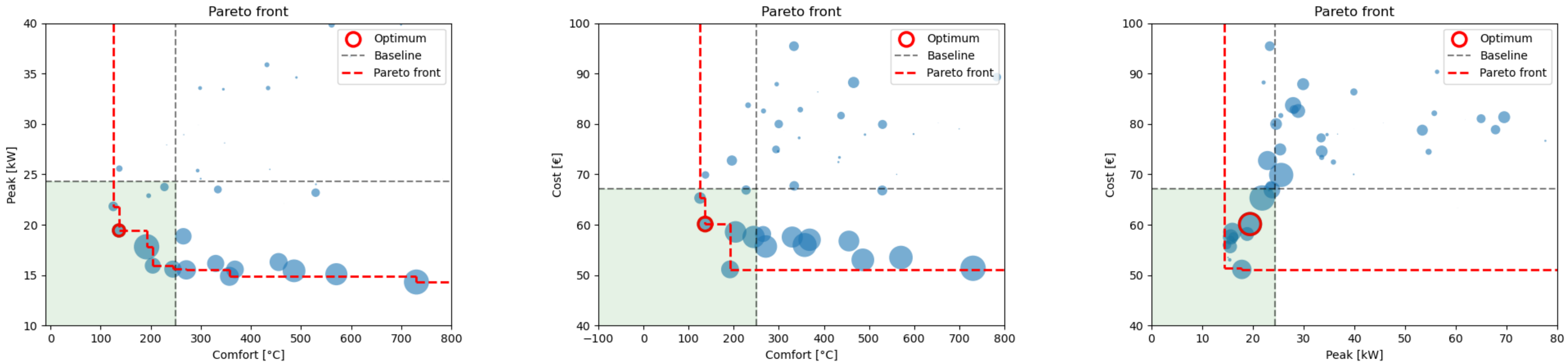
The optimal solution is a **dominated solution** in the **Cost/Peak** plane

Cost and peak are correlated since the electricity demand is the highest during occupied period, which occur during the highest electricity price

On the other hand, peak and comfort and cost and comfort are contrasting objectives

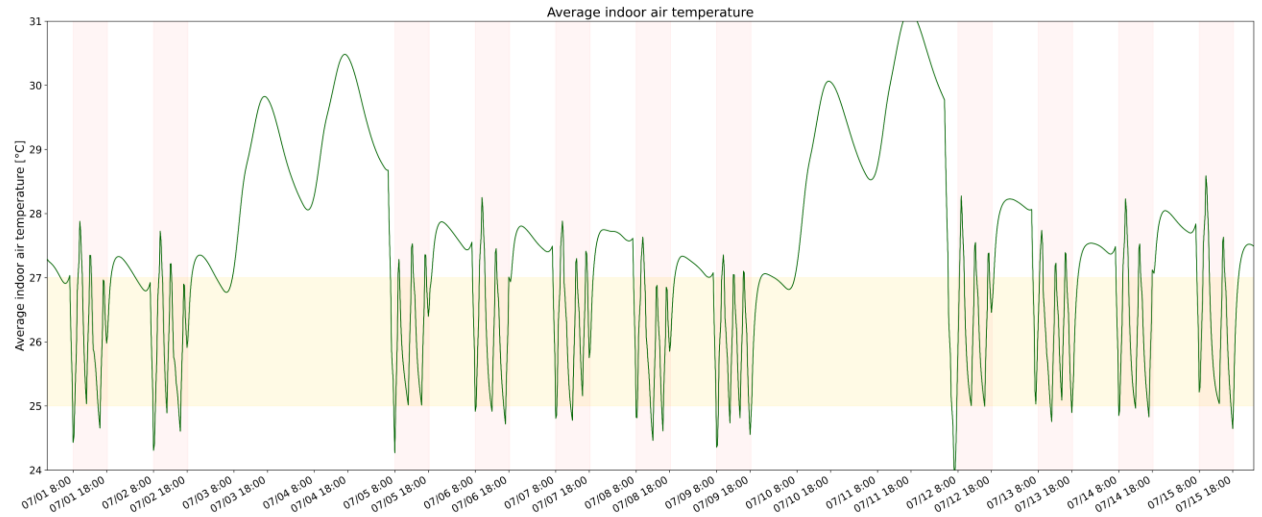
The size of the bubbles is related to the third dimension

Optimized Hyperparameters	Value
Discount factor	0.99
Learning rate	0.001
β	0.05
δ	1.5
θ	5

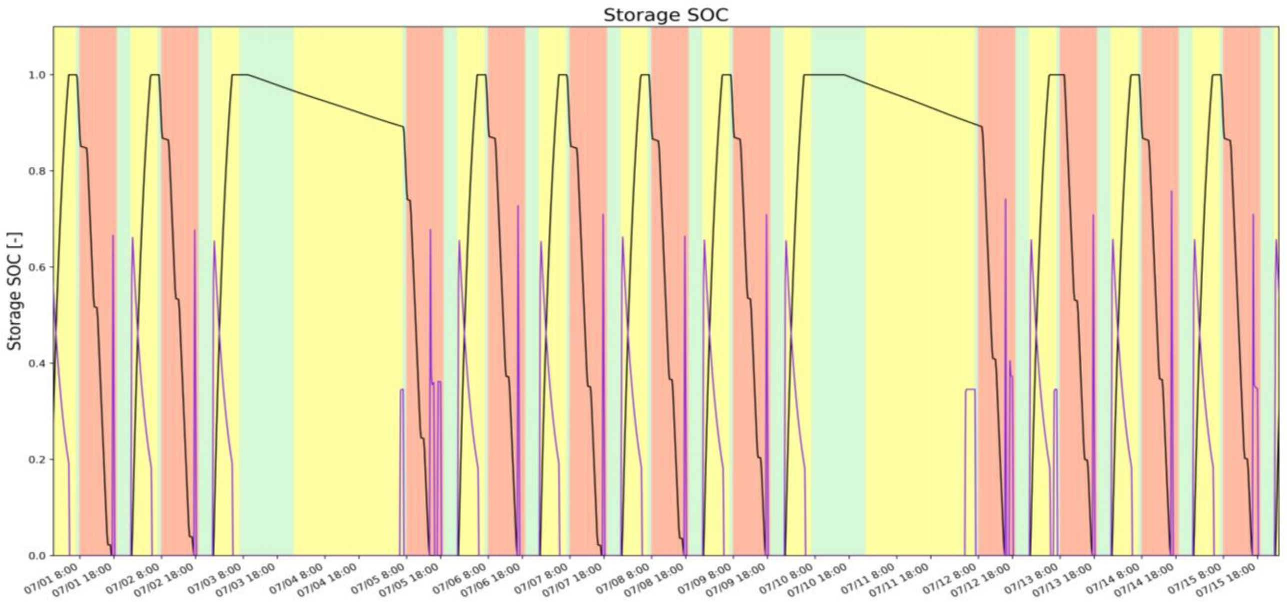


Politecnico di Torino
1859
Department of Energy "G.Ferraris"

Results – RBC Controller



The indoor temperature profile, the storage SOC and the energy consumption of the chiller [in kWh] are reported in these figures from RBC controller



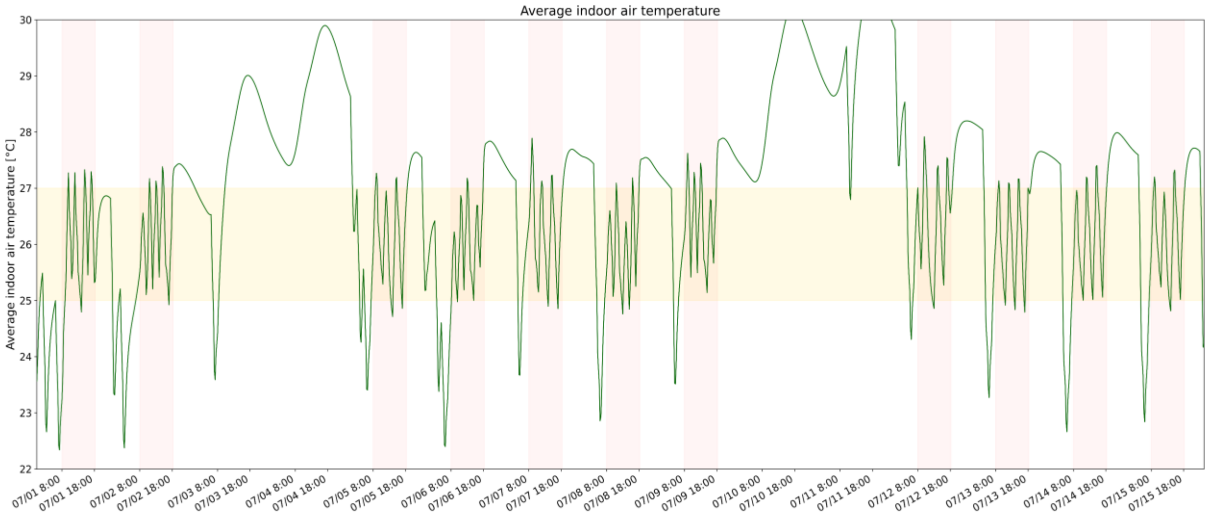
Legend

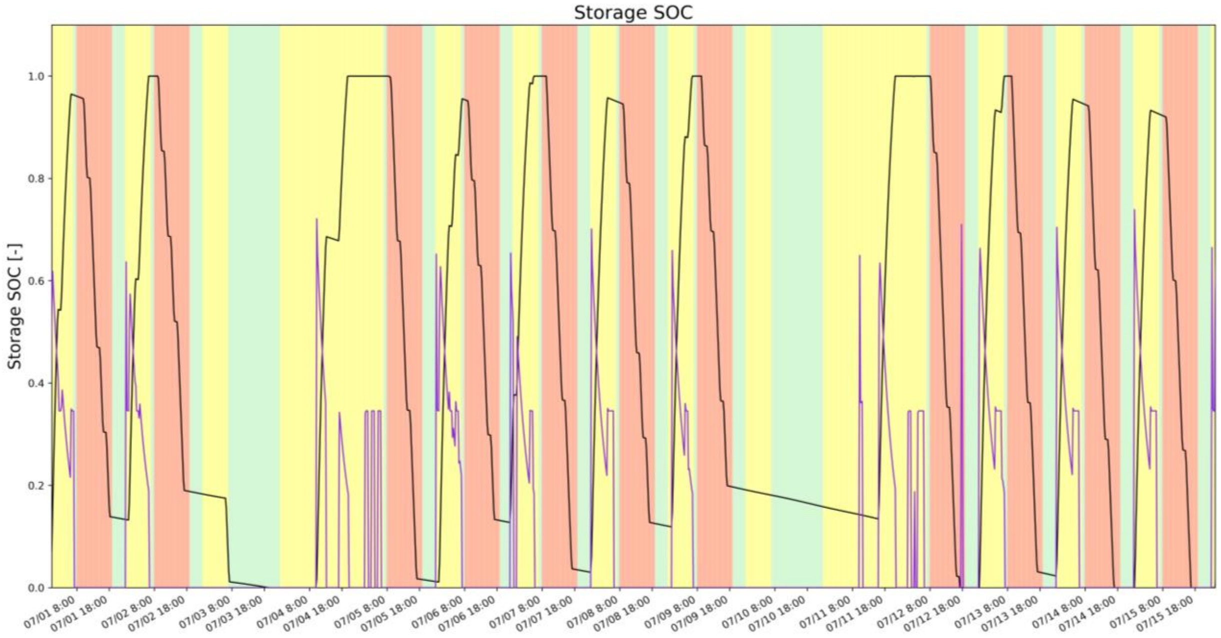
- TES SOC
- Chiller energy consumption

Electricity price

- Low
- Medium
- High

Results – DRL Controller





Results overview

Metrics	Baseline	DRL	Saving
Electricity costs [€]	67.2	60.1	-10.6%
Cumulative Temperature violations [°C]	250.5	137.8	-45.0%
Daily peak [kW]	24.3	19.4	-20.2%

Legend

- TES SOC
- Chiller energy consumption

Electricity price

- Low
- Medium
- High

The indoor temperature profile, the storage SOC and the energy consumption of the chiller [in kWh] are reported in these figures from DRL controller

Conclusions

- Minimizing contrasting objectives as the three considered in this work it is a complex task, since increasing one objective could degrates the other → in this case, the **DRL agent ensure a good tradeoff in the minimization of electricity/peak costs and thermal discomfort**

- DRL algorithm performances relies on the **choice of an appropriate set of hyperparameters.** Employing an **automated procedure** that explores the space of hyperparameters could **speed up the controller implementation process** in comparison to a manual approach:
    1. in a pareto set (i.e., set of all configurations which exhibit conflict among objectives), Optuna returns the information on the **pareto front**
    2. using the criterion of the minimum distance from the ideal point, it is obtained the **best solution** between non-dominated one

- The optimal solution may be improved by exploring more hyperparameters but at the cost of an increased computational time

Politecnico di Torino
Department of Energy "G.Ferraris"

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

THANK YOU FOR PAYING ATTENTION

References

1. S. Brandi, A. Gallo, A. Capozzoli, A predictive and adaptive control strategy to optimize the management of integrated energy systems in buildings, Energy Reports. 8 (2022) 1550–1567. https://doi.org/10.1016/j.egyr.2021.12.058.

2. T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: A Next-Generation Hyperparameter Optimization Framework, in: Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. & Data Min., Association for Computing Machinery, New York, NY, USA, 2019: pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.

3. J. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyper-parameter optimization, Adv. Neural Inf. Process. Syst. 24 25th Annu. Conf. Neural Inf. Process. Syst. 2011, NIPS 2011. (2011) 1–9.

4. Q. Xin, 3 - Optimization techniques in diesel engine system design, in: Q. Xin (Ed.), Diesel Engine Syst. Des., Woodhead Publishing, 2013: pp. 203–296. https://doi.org/https://doi.org/10.1533/9780857090836.1.203.

5. M. Zelany, A concept of compromise solutions and the method of the displaced ideal, Comput. Oper. Res. 1 (1974) 479–496. https://doi.org/https://doi.org/10.1016/0305-0548(74)90064-1.