A BIMtoBEM Framework for Building Retrofit and HVAC Smart Control Assessment

(Article begins on next page)

A BIMtoBEM Framework for Building Retrofit and HVAC Smart Control Assessment

1st Maria Adelaide Loffa
Politecnico di Torino
Turin, Italy
maria.loffa@polito.it

2rd Angelo J. Donato
Politecnico di Torino
Turin, Italy
angelo.donato@polito.it

3rd Pietro Rando Mazzarino
Politecnico di Torino
Turin, Italy
pietro.randomazzarino@polito.it

4th Enrico Macii
Politecnico di Torino
Turin, Italy
enrico.macii@polito.it

5th Anna Osello
Politecnico di Torino
Turin, Italy
anna.osello@polito.it

6th Matteo Del Giudice
Politecnico di Torino
Turin, Italy
matteo.delgiudice@polito.it

7th Edoardo Patti
Politecnico di Torino
Turin, Italy
edoardo.patti@polito.it

8th Lorenzo Bottaccioli
Politecnico di Torino
Turin, Italy
lorenzo.bottaccioli@polito.it

***Abstract*—Achieving climate neutrality requires a fundamental transformation of production systems and energy use, driven by technological innovation. In the building sector, virtual representations of physical assets can accelerate this transition by enabling simulation-based evaluation of energy strategies. When combined with Reinforcement Learning (RL), these models support dynamic testing and real-time optimization of building operations. This study presents a simulation framework for assessing and comparing energy management strategies aimed at reducing energy consumption while maintaining thermal comfort. As a case study, the methodology is applied to an existing industrial facility using the BIMtoBEM modeling approach. The framework integrates detailed simulation models with RL-based control to optimize the performance of the Heating, Ventilation, and Air Conditioning (HVAC) system. Two digital models with increasing Levels of Detail, are developed to evaluate the impact of three structural and one mechanical refurbishment scenario, alongside two RL control strategies. By simulating different combinations of physical retrofits and control approaches, the framework enables users to identify the most impactful interventions and make informed decisions based on specific energy-saving goals. Results show that modifying the mechanics of the HVAC system alone leads to a 12% reduction in natural gas consumption, while combining retrofitting with RL can lead to 32% of savings, emphasizing the impact of both physical and control-based interventions.**

***Index Terms*—Dynamic systems and control, Modeling, Reinforcement Learning**

I. Introduction

Energy consumption in buildings accounts for 17.5% of global greenhouse gas emissions [1], and this is primarily driven by the demand for appliances and HVAC systems to maintain the occupants' desired comfort and needs. This is why improving their energy efficiency and optimizing their operation has become pivotal, making building and HVACs one of the most significant object of research [2]. Advanced control strategies, like RL, dynamically adjust control policies in response to evolving objectives, overcoming the limitations of conventional approaches, taking into account the unique characteristics of individual buildings and sites and incorporating available forecasts such as weather conditions and occupancy patterns [3]. In [4], the authors implemented a hybrid approach that combines a model-free algorithm [5] with a black-box model of the building, which provides domain knowledge to the agent, allowing it to freely interact with the environment during training. This hybrid method is both fast and scalable, making it easily replicable across different buildings. The approach discussed in the paper employs an offline problem formulation, where the agent learns the optimal policy from pre-collected batches of data. This contrasts with the online learning paradigm, where the agent is deployed directly in the building without prior knowledge of the environment, enabling real-time actions and immediate feedback [6]. Building on this, the authors extended their previous work by introducing an online RL approach for HVAC control in [7], where the agent is trained directly on a white-box model that acts as a stand-in for the actual building.

Combining Artificial Intelligence (AI)-driven control algorithms with traditional refurbishment strategies (e.g., envelope renovation) in a unique simulation framework allow to asses energy-efficient management under different energy performance configurations. Flexible frameworks that take into consideration modelling and operational aspects are fundamental for comprehensive benchmarking and informed policy making [8].

This paper presents a novel simulation framework that combines the BIMtoBEM modeling methodology with an advanced control strategy. By extending a Building Information Model (BIM) into a Building Energy Model (BEM), the framework addresses the typical lack of operational data in existing buildings and enables performance assessment under different architectural and operational configurations. To this end, a digital model of an industrial facility is developed with increasing Levels of Detail (LOD), starting from a simplified "shoebox" model, Model 1.0, and progressing to a more refined version, Model 1.5, that includes detailed architectural and envelope features. These base models are further expanded into intervention scenarios: a mechanical upgrade of the HVAC system in Model 1.0, and three envelope refurbishment strategies in Model 1.5. This setup allows for the comparative evaluation of retrofit options, identifying those with the highest potential for energy savings. The models also serve as consistent and realistic environments for training and testing both offline and online RL algorithms, as in [4] and [7], focused on supply airflow control. This supports

This work is supported by MANAGE 5.0 project, an Italian National project funded by the Ministry of Economic Development - MISE (PNRR –Complementary Plan).

979-8-3315-9515-9/25/$31.00 ©2025 IEEE

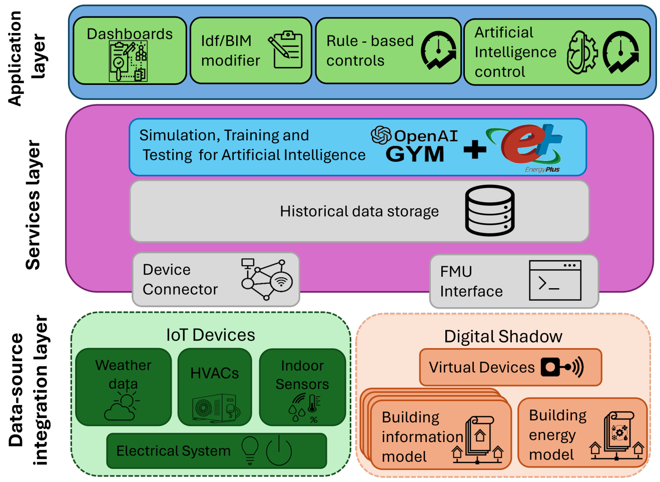

Fig. 1. Proposed methodology framework

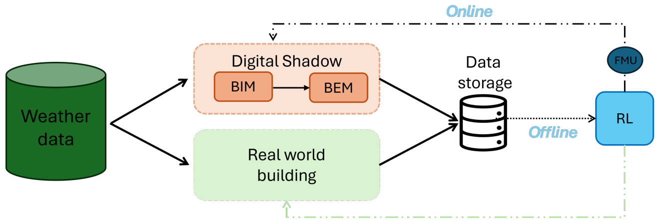

Fig. 2. Service layer implementation workflow

the development of effective, data-driven control policies. By comparing these control strategies, the study demonstrates the potential of simulation-driven RL to enhance building energy management.

The proposed methodology is directly transferable to real-world, IoT devices-equipped buildings by replacing the calibrated simulation with live data, offering a scalable and flexible solution for evaluating energy-saving measures and implementing smart building management strategies.

The rest of the paper is organized as follows: Section II describes the proposed methodology; Section III introduces our case study and the related outlined scenarios and Section IV presents the experimental results. Lastly, Section V reports our concluding remarks.

II. Methodology

The presented methodology can be recognized as a facility management platform, as illustrated in Fig. 1. It is organized into three conceptual layers that enable interoperability among the various actors: starting from the bottom, the *Data-Source Integration Layer* includes the elements that provide relevant data and facility features; the *Services Layer*, in the middle, connects these data sources and facility models to the platform's core functions, enabling data storage and facilitating AI-driven interventions; the *Application Layer* presents the results and insights in a user-accessible format, supporting the comparative analysis of different scenarios. The rest of this sections delves into the adopted framework in greater detail.

A. Data-Source Integration Layer

This layer can be seen as the combination of two main components: Internet of Things (IoT) devices and the Digital Shadow architecture. The IoT system collects data from multiple sources, including third-party weather services, building HVAC systems, and indoor sensors, with related actuators, that monitor variables such as air temperature and relative humidity. This setup allows for continuous monitoring of the building, providing insight into its energy performance in (near-) real-time. The Digital Shadow, which aims to realistically mirror the physical building, consists of BIM, which is subsequently translated into a BEM following the BIMtoBEM methodology [9]. The BIM, developed as federated model, comprises three distinct digital representations of the building: (i) the electrical system, which includes transformers, medium- and low-voltage switchboards, busbar trunking systems, and lighting devices; (ii) the mechanical system, primarily encompassing the HVAC infrastructure; (iii) the processes model, which represent the main production equipment; and (iv) the architectural model, which represents the building envelope and the internal room partitioning.

Elements from (ii) and (iv), such as the thermophysical properties of the envelope, the thermal zone configuration, and detailed HVAC specifications, serve as critical inputs for generating the BEM and conducting accurate energy simulations. The Digital Shadow thus provides virtual sensor data derived from these models, enabling a realistic representation of building behavior. Our methodology includes two thermophysical models of the building envelope that differ in geometric representation. Specifically, Model 1.0 adopts a simple shoebox geometry, while Model 1.5 incorporates more detailed features, including complex roof shapes and inclinations. This modeling workflow is applicable to any building type, as long as that the necessary system information is available and sufficient calibration data can be obtained.

B. Services Layers

This layer enables interoperability and communication among various IoT devices, within the device connector, allowing them to operate cohesively with each other [10]. It also serves as the environment where AI simulations, training, and testing are conducted. These activities are supported by a specialized simulation container, the Functional Mock-up Unit (FMU), which facilitates bidirectional communication between the simulated environment and the AI framework. In this context, the digital representation can either act as a stand-in for the physical building or be used to generate synthetic historical data through BEM simulations, helping address the issue of limited operational data in building management.

The practical implementation of this setup, enabled by the Services Layer, is illustrated in Fig. 2. Meteorological data are provided as input to the Digital Shadow, allowing BEM simulations based on the weather data specified location and time period. This process facilitates the generation and storage of synthetic datasets. The AI controller, implemented as RL agent, can be trained in two ways: offline, using pre-sampled batches from the stored data; or online, by interacting directly with the Digital Shadow, executing control actions on the simulated system and receiving real-time feedback to iteratively improve its policy. The training and optimization process is supported by a set of integrated components, including OpenAI Gym [11] for RL environments and EnergyPlus for building energy simulation. To support this online training process, a custom learning environment was developed using the Functional Mock-up Interface - Machine Learning Center

(FMI-MLC) framework [12], which enables real-time co-simulation between Python-based RL agents and EnergyPlus model's FMU [13]. Effectively, the building's BIM serves as a virtual gym [14], enabling safe and effective agent training before deployment in a real-world scenario. This minimizes the risk of adverse outcomes, such as suboptimal decisions that could affect thermal comfort, energy efficiency, or system stability during early learning phases. Once the physical building is adequately equipped, the same training and evaluation process can be executed directly on the real-world building environment.

C. Application Layer

The top layer of the framework provides interfaces for visualization, monitoring, and user interaction. It includes dashboards for real-time data analytics and performance tracking, supporting human-in-the-loop supervision and informed decision-making on energy management, maintenance, and control strategies. Through these tools, users can not only compare and evaluate different control approaches, such as rule-based or AI-driven controllers, but also analyze the effects of physical interventions on the building's structural or mechanical systems, which can be simulated by modifying the BIM identification file (IDF). In such cases, the entire platform can be re-applied to the updated BIM model, enabling end-to-end testing of scenarios involving changes to the building envelope, layout, or HVAC configurations.

III. Case study

The case study illustrating the application of the proposed begins with a description of the BIM models, along with the corresponding refurbishment scenarios. This is followed by an overview of the implemented control strategies.

A. Digital Shadow Modeling

The objective of our study is an existing industrial plant located in Northwestern Italy, whose BIM model has been developed through the Autodesk Revit platform [15], following on-site inspections, complemented with available descriptive and contextual particulars. The BEM simulations have been carried out with the EnergyPlus [16] energy simulation tool.

To evaluate the impact of geometric elements, the architectural representation was developed with two progressively detailed configurations: *Model 1.0*, shown in Fig. 3a, features a simplified roof; *Model 1.5*, in Fig. 3b, on the other hand, reflects the actual building configuration, including a shed-type roof. Each model treats the entire volume as a single thermal zone, as there is no distinction in how the spaces are thermally managed.

For what concerns the mechanical system of the building, two different HVAC models were adopted: a Constant Air Volume (CAV) type, which allows control only over the supply air temperature, implemented using the EnergyPlus *AirTerminal:SingleDuct:ConstantVolume:NoReheat* template and a Variable Air Volume (VAV) system, enabling control over both the supply air temperature and the airflow rate. The corresponding EnergyPlus template is *AirTerminal:SingleDuct:VAV:Reheat*. Both HVAC configurations are integrated into the building envelope model via the DesignBuilder interface [17].

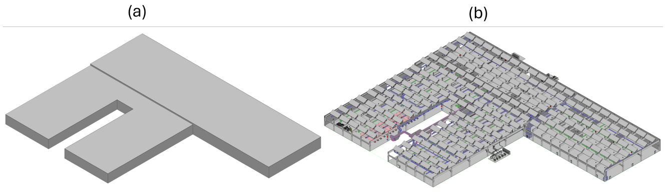

Fig. 3. Comparison between Architectural Model 1.0 (a) and 1.5 (b)

B. Scenarios definition

Five different scenarios were defined by combining the two architectural models, Model 1.0 and Model 1.5, with the two HVAC configurations, CAV and VAV, introduced in the previous paragraph.

a) Scenario 1.0-Baseline: couples the simplified Model 1.0 with the CAV system;

b) Scenario 1.0-A: combines Model 1.0 with the more advanced VAV HVAC template;

c) Scenario 1.5-Baseline: features the Model 1.5 architectural representation and the CAV system.

To further investigate the impact of additional envelope features on energy savings, Scenario 1.5-Baseline was extended with three refurbishment strategies; all of them conceived and implemented on Model 1.5 coupled with the CAV template.

d) Scenario 1.5-A: focuses on the glazed areas of the shed roof, which have been categorized into three types based on their structural openings for glazing elements. The intervention involves replacing the existing traditional single-pane glass with a more suitable and technologically advanced glazing system designed for industrial use. The selected solution has a U-value above 1.1 W/m²K and complies with current Italian regulations [18];

e) Scenario 1.5-B: targets the plant's opaque surfaces, particularly the external walls. Previous inspections revealed the absence of any insulation layer; therefore, the proposed scenario evaluates the impact of installing polyurethane panels, paired with a metal carter capable of withstanding mechanical stress, fire, and low temperatures. The resulting assembly achieves an estimated U-value of 0.12 W/m²K;

f) Scenario 1.5-C: addresses the opaque portions of the roof, specifically the horizontal surfaces between each row of sheds, as well as the non-glazed sections of the shed structures themselves. For the sloped shed surfaces, a suitable insulated construction finished with metal cladding has been selected. In contrast, the horizontal areas are designed with an insulated build-up that ensures adequate slope for proper rainwater drainage. The two setups have an U-value of 0.17 and 0.15 W/m²K respectively.

Each of the defined simulation scenarios was evaluated using two control strategies: a basic PID controller serving as a benchmark, and a more advanced RL approach, which will be detailed in the following section.

C. Reinforcement Learning Algorithms

In a RL framework, the agent interacts with the environment to learn its dynamics through a reward-based algorithm. It observes the environment's state s, either fully or partially, and selects actions a that influence future states. Based on these actions, the agent, a decision-making entity, receives a reward signal r that reflects its performance relative to the control objectives, guiding its learning process. The decisions made by the agent are governed by a policy π, and the ultimate goal is to learn the optimal policy π^* that maximizes cumulative rewards.

In RL, the state-value function v_π and action-value function q_π estimate expected returns under a policy [19]. Different RL methods target different goals: value-based approaches like Q-learning learn the optimal action-value function without relying on a fixed policy (off-policy), while policy-based methods directly optimize the decision-making strategy and typically follow an on-policy framework [20].

We implement and compare two different learning frameworks for the HVAC system control. While both share the state representation and reward function, they differ in the state-transition function, action space and algorithm.

For the first implemented framework, applied to Scenario 1.0-Baseline, Scenario 1.5-Baseline, with its related Scenarios 1.5-A, 1.5-B and 1.5-C, an off-policy DDQN algorithm is coupled with a fully connected feedforward neural network that serves as a system identification model and acts as a data-driven state transition function, learning to predict the next indoor temperature based on selected disturbances. The input features include normalized values of indoor and outdoor temperature, HVAC temperature differential, and occupancy status. The network comprises two hidden layers with ReLU activations and is trained to minimize the mean squared error between predicted and actual temperatures. This model enables the agent to interact with a fast and reliable representation of the building during offline training, supporting efficient policy learning across different scenarios. The agent is trained offline on a three-month dataset (from January 1st to March 31st in 2021), generated through EnergyPlus weather files and simulations of the building models. The action space is discrete and represents the increase in supply air temperature provided by the HVAC heating coil, with available actions being: [0, 0.5, 1, 1.5, 2.5, 3.5, 4.5, 5] °C.

For the second adopted framework, applied to Scenario 1.0-A, the simulated building serves as a test-bed for the real one, providing ground-truth values for the state-transition function [7]. In this case, a policy-based DDPG agent is trained in an online fashion, directly interacting with the environment by executing actions on the simulated system and receiving real-time feedback. The control actions involve adjusting the supply air temperature and mass flow rate. The action space is continuous, allowing the agent to select and apply any value within a defined range: temperature values range from 0 to 4.5 °C, and mass flow rates from 0 to 60 kg/s.

In both cases, the state representation includes: the difference between the setpoint and the indoor temperature over the past four time steps, the difference between the setpoint and the current outdoor temperature, outdoor relative humidity, wind speed and direction, diffuse and direct solar radiation, and the number of hours remaining until the start and end of the next occupancy period [4]. The reward function takes as input the current state and control action, and outputs the expected reward, which the agent aims to maximize. In both cases, the objective is to optimize HVAC energy consumption while ensuring thermal comfort for occupants when present. To achieve this, the reward function, shown in Eq. (1), consists of three main components: (1) the squared deviation between the resulting indoor temperature (x_t) and the air setpoint ($x_{setpoint}$), (2) the cost of the selected control action (u_t), (3) a term ($var\_penalty$) that penalizes the cumulative error between predicted and actual next states over the course of an episode, encouraging stable and reliable policy behavior.

$$\mathrm{R}(s,u) = -(\beta\cdot(x_t - x_{\text{setpoint}})^2 + \rho\cdot u + 0.5\cdot \mathit{var\_penalty}) \quad (1)$$

The weighting parameters β and ρ balance comfort and energy efficiency: a higher β prioritizes maintaining the indoor temperature close to the setpoint, while a higher ρ places greater emphasis on minimizing energy consumption. β is set to 2 when the building is occupied, to prioritize thermal comfort, and reduced to 0.05 when unoccupied, reflecting a lower need for strict temperature regulation. On the other hand, ρ is kept constant at 1, ensuring that energy saving is always taken into account.

Within the RL frameworks, the building models serve distinct roles: in the online approach, they act as the training environment, while in the offline setting, they are used to generate the training datasets. In both cases, the resulting RL agents are tested on simulations from the heating season of 2022 using the presented mock-ups.

IV. Experimental Results

This section presents the results obtained using the proposed methodology, comparing the different scenarios analyzed, with and without the implementation of the RL control strategy. The focus is placed on demonstrating the capabilities of the simulation framework, which enables efficient evaluation of diverse configurations and interventions with minimal setup effort. The comparison is based on natural gas consumption for the heating system, and the Predicted Percentage of Dissatisfied (PPD) index, which quantifies the occupants' thermal comfort.

A. Refurbishment measures and energy savings

To ensure the validity of the implemented methodology, it is essential that the building models are calibrated against real data. For the calibration to be considered acceptable, the discrepancy between simulated and actual energy consumption should not exceed 8%. For this purpose, natural gas consumption bills from the year 2022 were used, focusing exclusively on the winter and autumn months (January–March and October–December), as these correspond to the heating season. Scenario 1.0-Baseline showed a 4% deviation, from the actual natural gas consumption, while Scenario 1.5-Baseline achieved a reduced error of 1.3%. This improvement can be justified by the increased geometric and envelope detail, which enhances the accuracy of thermal performance simulation. As shown in the bar plot in Fig. 4, which compares the monthly natural gas consumption extracted from utility bills with that of the two models prior to the RL implementation, the improved calibration is clearly visible.

Fig. 4 also reports the energy consumption across all evaluated scenarios. Among them, the intervention that led to the greatest energy savings is Scenario 1.5-C, which involves enhancing the opaque surfaces of the roof. This upgrade resulted in a 21% reduction in consumption compared to the original calibrated Scenario 1.5-Baseline: its effectiveness is due to the focus on the opaque roof surfaces, which also surround the roof sheds. These areas are particularly vulnerable to heat loss, especially where the glazed areas of the sheds meet the opaque surfaces, creating significant thermal bridges that contribute to higher heat loss. Hence insulating the roof construction reduces overall energy consumption.

Meanwhile, modifying the HVAC system from CAV to VAV in Scenario 1.0-Baseline and Scenario 1.0-A delivers a 12% saving and this is largely due to the ability to modulate the supply airflow, allowing the system to operate at or near minimum mass flow rates when the building is empty or already heated, thus significantly lowering energy demand.

Table IV-A also mentions the consumption for each model and scenario when the DDQN algorithm, combined with the black-box system identification model, is applied. The optimized Scenario 1.0-Baseline shows a 41% reduction in natural gas consumption for heating, compared to a 14% reduction in optimized Scenario 1.5-Baseline. While both models meet calibration criteria and reflect the real building's overall behavior, these results highlight that the simpler Model 1.0, despite being faithful to the available data, is less accurate and reliable. The more detailed one provides a closer representation of the building's thermal dynamics, leading to more robust and realistic predictions. However, this comes at the cost of significantly increased modeling effort and simulation time. As such, Model 1.0 may be more suitable for applications where rapid analysis or iterative control strategy testing is prioritized.

With respect to Scenario 1.5-Baseline and its three refurbishment scenarios; out of all three options, the introduction of the RL algorithm provides the greatest benefit in the scenario 1.5-B, where it enables savings of up to 32 % of the natural gas in the heating season. The refurbishment intervention on the external walls increases the building's thermal mass and capacitance, allowing the envelope to store and gradually release heat. This characteristic works particularly well with the RL agent, which learns to anticipate heating needs and adjusts control actions accordingly. By exploiting the delayed heat transfer of the improved envelope, the agent can optimize energy use more effectively, reducing consumption. Aside from the Scenario 1.5-B, the remaining three configurations also achieve significant control savings, ranging from 9% in Scenario 1.5-C, where the control reduces consumption that was already minimized by structural improvements, to 14% in the unmodified building, and 28% in the case focusing on the glazed surfaces. Additionally, it is worth noting that the mean PPD index, shown in the last column of Table IV-A, remains below the acceptable limit of 10% (ISO 7730) in all cases, ensuring thermal comfort is not compromised.

To better illustrate the effectiveness of the RL control, Fig. 5 presents the supply air temperature results for Scenario 1.5-B. This is shown as a representative case to avoid confusion from displaying multiple similar results for each scenario. The trend observed in this graph highlights that the control agent is able to reduce the supply air temperature when the occupancy flag is zero and, when the building is unoccupied, it does not continuously heat at its maximum level.

B. RL Algorithm Comparison: DDQN vs. DDPG

The RL algorithms, offline DDQN and online DDPG, were applied to Scenario 1.0-Baseline and to Scenario 1.0-A respectively. Despite its simplified nature, Model 1.0 was selected for testing the online methodology, as it provides a more straightforward and manageable environment for evaluating the performance of the online RL strategy. Although this choice may yield less reliable results compared to more detailed models, it offers clearer insights into the online approach's behavior.

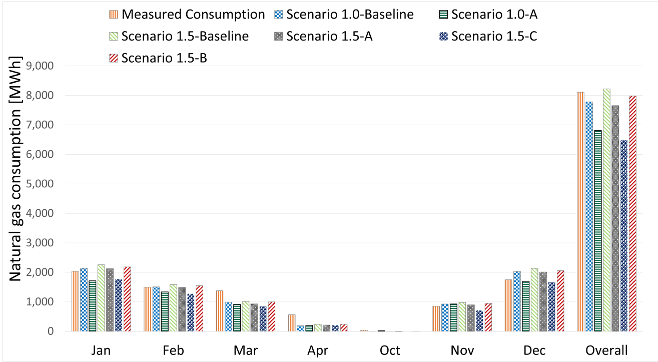

Fig. 4. Monthly natural gas consumption models and scenarios

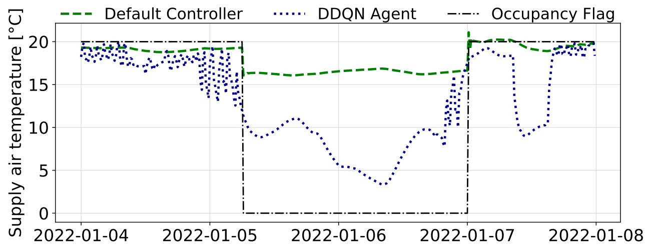

Fig. 5. DDQN agent and default controller supply air temperature comparison

The trained DDPG agent, operating in a continuous action space, effectively adjusts the indoor air temperature to closely match the desired setpoint, except during unoccupied periods when the heating system is completely turned off to optimize energy use, as shown by the dotted line in Fig. 6. In contrast, the DDQN agent, during occupancy, follows a temperature profile that mirrors the daily and nightly outdoor air temperature fluctuations, albeit much more smoothly (dashed line in Fig. 6).

Overall, the DDPG agent's control for the Scenario 1.0-A resulted in a 45% reduction in natural gas consumption during the considered heating period, compared to a 41% reduction achieved by the DDQN agent. The superior performance of the DDPG agent is likely attributed to its ability to control both the supply air temperature and the HVAC airflow, the latter being a key factor in regulating indoor conditions. In contrast, the DDQN agent operates in a CAV configuration and is limited to discrete temperature control. The DDPG agent dynamically modulates the mass flow rate and, consequently, the heating coil heating power: unlike the default controller and the DDQN for the CAV system in Scenario 1.0-Baseline, which operates at maximum power, the agent is able to reduce it to zero, as shown in Fig. 7. Another key difference is that during training, the DDPG agent receives real-time feedback directly from the actual environment. This results in more accurate responses, allowing the agent to make more informed and effective decisions. In contrast, the DDQN agent relies on responses generated by the system identification neural network, which may not capture the full complexity of the real-world environment.

As shown in Fig. 7, the DDPG agent dynamically lowers the mass flow rate, allowing the heating coil power to drop

TABLE I
ENERGY PERFORMANCE COMPARISON WITH REFURBISHMENT AND OPTIMIZED CONTROL STRATEGIES

	Default control [MWh]	Refurbishment Savings [%]	Optimized Control [MWh]	Control savings [%]	PPD [%]
Scenario 1.0-Baseline	7786	-	4541	41%	8%
Scenario 1.0-A	6813	12%	4771	45%	6%
Scenario 1.5-Baseline	8220	-	7041	14%	6.6%
Scenario 1.5-A	7653	13%	5517	28%	8%
Scenario 1.5-B	7980	10%	5362	32%	8.4%
Scenario 1.5-C	6479	21%	5900	9%	6.4%

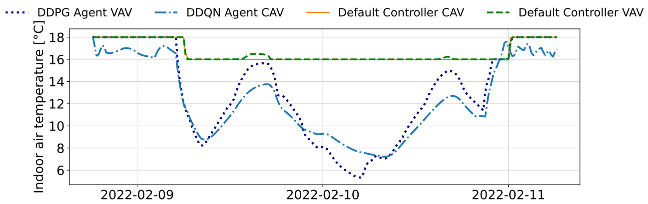

Fig. 6. Offline DDQN and online DDPG agents indoor air temperature comparison

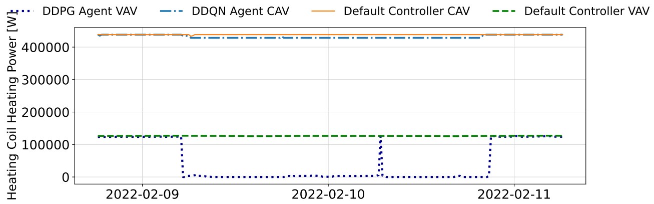

Fig. 7. Offline DDQN and online DDPG heating coil heating rate comparison

significantly, even to zero, during low-demand periods, unlike the more rigid operation of the default and DDQN controllers.

V. CONCLUSIONS

This study introduces a simulation framework for facility management that allows users to test and compare various building configuration scenarios to identify the most effective energy-saving interventions. The framework combines the BIMtoBEM methodology with advanced RL control, applied to an existing industrial facility. A Digital Shadow of the building is created, incorporating multiple refurbishment strategies and two RL algorithms, trained in an offline and online fashion. The approach provides a reliable yet simple way to estimate energy savings and is easily adaptable to other buildings.

To further validate and expand the applicability of these strategies, future works will explore online training on higher-detail building models, potentially combined with multi-agent RL paradigm. The transition from a Digital Shadow to a full Digital Twin will be pursued enabling two-way communication between the physical system and its digital counterpart.

REFERENCES

[1] H. Ritchie, P. Rosado, and M. Roser, "Emissions by sector," *Our World in Data*, 2020, https://ourworldindata.org/emissions-by-sector.

[2] I. Ahmed, M. Asif, H. H. Alhelou, M. Khalid *et al.*, "A review on enhancing energy efficiency and adaptability through system integration for smart buildings," *Journal of Building Engineering*, p. 109354, 2024.

[3] Z. Nagy, G. Henze, S. Dey, J. Arroyo, L. Helsen, X. Zhang, B. Chen, K. Amasyali, K. Kurte, A. Zamzam *et al.*, "Ten questions concerning reinforcement learning for building energy management," *Building and Environment*, vol. 241, p. 110435, 2023.

[4] F. M. Solinas, A. Bellagarda, E. Macii, E. Patti, and L. Bottaccioli, "An hybrid model-free reinforcement learning approach for hvac control," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*. IEEE, 2021, pp. 1–6.

[5] X. Wang, W. Xiong, H. Wang, and W. Y. Wang, "Look before you leap: Bridging model-free and model-based reinforcement learning for planned-ahead vision-and-language navigation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.

[6] S. Brandi, M. Fiorentini, and A. Capozzoli, "Comparison of online and offline deep reinforcement learning with model predictive control for thermal energy management," *Automation in Construction*, vol. 135, p. 104128, 2022.

[7] F. M. Solinas, A. Macii, E. Patti, and L. Bottaccioli, "An online reinforcement learning approach for hvac control," *Expert Systems with Applications*, vol. 238, p. 121749, 2024.

[8] P. Rando Mazzarino, A. Macii, L. Bottaccioli, and E. Patti, "A multi-agent framework for smart grid simulations: Strategies for power-to-heat flexibility management in residential context," *Sustainable Energy, Grids and Networks*, vol. 34, p. 101072, 2023.

[9] A. Osello, M. Del Giudice, A. J. Donato, and A. Fratto, "Towards climate neutrality – the key role of the digital twin in industry 5.0," *AGATHÓN — International Journal of Architecture, Art and Design*, vol. 15, p. 276–285, Jun. 2024. [Online]. Available: https://www.agathon.it/agathon/article/view/418

[10] E. Patti, A. Acquaviva, and E. Macii, "Enable sensor networks interoperability in smart public spaces through a service oriented approach," in *5th IEEE International Workshop on Advances in Sensors and Interfaces IWASI*, 2013, pp. 2–7.

[11] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai gym," 2016.

[12] R. A. Gehbauer, Christoph and L. Eleanor, "Advanced control of dynamic facades and hvac with reinforcement learning based on standardized co-simulation," *Building Simulation*, 2021.

[13] "Functional mock-up interface." [Online]. Available: https://fmi-standard.org/

[14] A. Shahcheraghian, H. Madani, and A. Ilinca, "From white to black-box models: A review of simulation tools for building energy management and their application in consulting practices," *Energies*, vol. 17, no. 2, 2024. [Online]. Available: https://www.mdpi.com/1996-1073/17/2/376

[15] Autodesk, "Revit learning," 2024. [Online]. Available: https://help.autodesk.com/view/RVT/2024/ENU/

[16] EnergyPlus, "Energyplus documentation," 2021. [Online]. Available: https://energyplus.net/documentation

[17] "Design builder - design builder support." [Online]. Available: https://designbuilder.co.uk/helpv7.3/

[18] "Ministero dello sviluppo economico , decreto requisiti minimi - appendice a," 2015. [Online]. Available: https://www.gazzettaufficiale.it/eli/id/2015/07/15/15A05198/sg

[19] S. Brandi, M. S. Piscitelli, M. Martellacci, and A. Capozzoli, "Deep reinforcement learning to optimise indoor temperature control and heating energy consumption in buildings," *Energy and Buildings*, vol. 224, p. 110225, 2020.

[20] D. Azuatalam, W.-L. Lee, F. De Nijs, and A. Liebman, "Reinforcement learning for whole-building hvac control and demand response," *Energy and AI*, vol. 2, p. 100020, 2020.