

**ARTICLE**

Reinforcement Learning Model for Energy System Management to Ensure Energy Efficiency and Comfort in Buildings

Inna Bilous¹, Dmytro Biriukov¹, Dmytro Karpenko², Tatiana Eutukhova²,
Oleksandr Novoseltsev^{2,*} and Volodymyr Voloshchuk¹

¹Department of Heat and Alternative Power Engineering, National Technical University of Ukraine “KPI”, Kyiv, 03056, Ukraine

²Department of Structural Transformation of the Fuel and Energy Complex, General Energy Institute of NAS of Ukraine, Kyiv, 03150, Ukraine

*Corresponding Author: Oleksandr Novoseltsev. Email: anovos773@ukr.net

Received: 12 March 2024 Accepted: 24 April 2024 Published: 22 November 2024

ABSTRACT

This article focuses on the challenges of modeling energy supply systems for buildings, encompassing both methods and tools for simulating thermal regimes and engineering systems within buildings. Enhancing the comfort of living or working in buildings often necessitates increased consumption of energy and material, such as for thermal upgrades, which consequently incurs additional economic costs. It is crucial to acknowledge that such improvements do not always lead to a decrease in total pollutant emissions, considering emissions across all stages of production and usage of energy and materials aimed at boosting energy efficiency and comfort in buildings. In addition, it explores the methods and mechanisms for modeling the operating modes of electric boilers used to collectively improve energy efficiency and indoor climatic conditions. Using the developed mathematical models, the study examines the dynamic states of building energy supply systems and provides recommendations for improving their efficiency. These dynamic models are executed in software environments such as MATLAB/Sim-scape and Python, where the component detailing schemes for various types of controllers are demonstrated. Additionally, controllers based on reinforcement learning (RL) displayed more adaptive load level management. These RL-based controllers can lower instantaneous power usage by up to 35%, reduce absolute deviations from a comfortable temperature nearly by half, and cut down energy consumption by approximately 1% while maintaining comfort. When the energy source produces a constant energy amount, the RL-based heat controller more effectively maintains the temperature within the set range, preventing overheating. In conclusion, the introduced energy-dynamic building model and its software implementation offer a versatile tool for researchers, enabling the simulation of various energy supply systems to achieve optimal energy efficiency and indoor climate control in buildings.

KEYWORDS

Building energy management; building heating system; dynamic modeling; reinforcement learning; energy efficiency; comfortable temperature

Main Findings of the Research:

- Dynamic modelling is one of the most effective tools for making systemic management decisions to improve energy efficiency in concert with other internal and external influences.



- Energy supply sources specified from the libraries of existing software products must be cardinally transformed to allow analysis of processes in the structural components of these sources, especially in transient modes.
- The complexity and dynamic nature of modern buildings require advanced building control systems that are adaptable, optimized and self-learning.
- Dynamic reinforcement learning (RL) modelling of energy processes in buildings makes it possible to provide a comfortable indoor air temperature, taking into account the peculiarities of the effective operation of the energy supply sources.
- The results of the energy-dynamic RL grid modelling of a typical residential building show that it is possible to reduce the instantaneous output of the building's heating systems by up to 35%, to reduce the absolute deviations from a comfortable temperature by almost a factor of 2 and to reduce energy consumption by around 1% while maintaining a comfortable temperature.

1 Introduction

Globally, the building sector stands as the biggest energy consumer, responsible for about 35% of the total energy usage, outperforming the transportation and industrial sectors. In regions such as the European Union and the United States, buildings are particularly significant, consuming nearly 40% of all energy and contributing 36% of CO₂ emissions. Projections indicate that this proportion will rise by 50% by the year 2050, relative to 2013 levels [1]. As the principles of transitioning to a low-carbon, green economy gain traction, energy efficiency standards for buildings and their energy supply systems are expected to become even more stringent [2]. Apart from energy efficiency, several critical factors influencing indoor climate quality must be considered. Primary among these are indoor temperature and air pollution, where a failure to meet sanitary standards can severely impact the health of building occupants and employees. Common pollution sources include industrial processes, cooking activities, and the use of substandard construction materials [3]. The interplay of these various factors necessitates thorough investigation and subsequent strategic coordination for optimal results.

Reduction of energy consumption levels in buildings is achieved through various methods and depends on several elements, including the thermal insulation of structural components, the management of energy supply systems, and climatic conditions. To promote the efficient utilization of energy resources, it is necessary to adopt a comprehensive approach that examines the influence of these factors and identifies potential energy savings. The assessment of a building's energy efficiency relies on mathematical models, which are predominantly implemented as software suites. These software solutions often employ grid models that integrate physically interdependent elements.

Models formulated through equations of physical processes or simulation models are termed "white box" models. Simulation modeling involves framing the model as an algorithm within a computer program that replicates the object's behavior. These simulations are essentially computational experiments that model real-building dynamics through mathematical representations. The process entails simulating elementary phenomena that constitute the overall procedure while retaining their logical structure and temporal sequence. This approach provides information about the system's state at a specified time and facilitates the estimation of system characteristics.

Building Energy Modeling (BEM) is instrumental in forecasting a building's energy consumption and assessing energy savings relative to a standard baseline. BEM relies on typical meteorological year data and assumptions regarding building operations, which enables the calculation of various energy-conservation measures [4,5]. In Building Energy Management Systems (BEMS), a diverse array of

algorithms has been developed, especially for heating system control to maintain indoor comfort. These algorithms can be broadly categorized into conventional and advanced control strategies, including machine learning-based methods such as RL. These algorithms adjust the heating system's operation to efficiently reach desired temperature setpoints, responding adaptively to shifts in outdoor temperature, occupancy patterns, and other pertinent factors.

Conventional control methods in BEMS include “on/off” (hysteresis) and PID (proportional-integral-derivative) control strategies. The “on/off” method is straightforward, where the heating system is turned off once the room temperature exceeds a predetermined upper threshold and turned on when it drops below a lower threshold. Despite its simplicity, this method can cause significant temperature fluctuations and inefficiencies. Conversely, PID control is a more advanced strategy that computes an error value by comparing the desired setpoint to the actual temperature. The PID controller then adjusts the heating output based on proportional, integral, and derivative terms of this error, aiming to reduce the discrepancy smoothly over time. This method typically achieves higher energy efficiency and more stable temperature control than on/off control [6,7].

The use of RL approaches in building energy management systems represents a transformative potential that is increasingly becoming a cornerstone of contemporary research. The adoption of RL methodologies offers several key advantages that mark a significant departure from conventional energy management techniques. RL represents a significant leap forward in the quest for efficient, sustainable, and intelligent building operations. Its ability to learn from and adapt to real-time data creates a paradigm shift, enabling proactive rather than reactive management of energy resources. RL methods excel at dynamically adapting to varying conditions within the building environment. Traditional control systems often rely on predefined schedules and static algorithms that do not account for real-time changes. In contrast, RL can uninterruptedly learn and optimize energy consumption based on continuous feedback from building operational data and external inputs such as weather conditions and occupancy patterns. From an economic perspective, the use of RL in building energy management systems can result in significant cost savings. Energy optimization directly translates into lower utility bills and reduced operating costs. In addition, the predictive capabilities of RL can help predict maintenance needs and mitigate the risk of equipment failure, further reducing the costs associated with unplanned downtime and repairs [8,9].

The objective of the study is to construct an energy dynamic model for buildings and conduct comparative simulations of various energy management systems and algorithms. This effort intends to enhance both energy efficiency and the internal climate of buildings in a balanced manner.

The hypothesis of the study claims that building occupant satisfaction can be maximized by optimizing dynamic operating modes and identifying the most effective boiler controller through dynamic modeling.

To fulfill this objective, the research employs a sequence of essential steps: forming an energy dynamic model of the building, validating and fine-tuning the model, conducting energy simulations using an electric boiler as the heating source, testing various controllers, analyzing the results, and formulating recommendations.

The novelty of the study lies in systematically addressing the challenge of achieving both energy efficiency and comfortable living conditions through the dynamic operation of energy supply sources. Specific contributions include:

- At the system-wide energy dynamic model level: by systematically coordinating mathematical models of different controllers used in building energy management systems, represented through structural diagrams of dynamic modeling.
- At the energy supply parameter control level: by selecting controllers using reinforcement learning techniques to efficiently regulate the power capacity of energy sources.
- At the modeling level: by enhancing the learning algorithm of the neural network model for reinforcement learning in managing the building's heating system and refining the reward function during both its training and operational phases.

The practical relevance of this study is to provide a methodological framework for designers and engineers, enabling them to enhance the efficiency and quality of building energy supply systems. This framework aids in making better-informed decisions regarding the measures required to improve outcomes.

2 Literature Review

Numerous advanced software tools are available for evaluating building energy consumption, including EnergyPlus, TRNSYS, eQuest, and MATLAB/Simulink [10–12]. One of the prevalent issues in the heating and cooling of residential buildings in regions with cold and temperate climates, such as Europe, is inadequate indoor temperature regulation [13]. The paper [14] provides an analysis of dynamic regulation mechanisms for indoor air temperature utilizing centralized control systems. When it comes to residential buildings with autonomous heating sources, selecting the appropriate heat source and understanding their operational characteristics involve consideration of multiple influential factors [15].

Most current building energy supply technologies employ heat sources with high exergy values, yet the energy required inside buildings to maintain thermal comfort is of relatively low exergy, typically within the temperature range of 18°C–20°C. This disparity leads to a significant loss of high-quality energy. Recently, an innovative approach has emerged in the heating and cooling sector, focusing on low-exergy buildings [16].

The review [17] focuses on four primary aspects of human comfort: thermal, visual, acoustic, and respiratory comfort. The increasing use of sensor technologies for data collection and the application of machine learning techniques, particularly in the realm of thermal comfort were highlighted. The thermal indoor comfort criterion is a critical aspect of BEMS that aims to optimize energy use while ensuring a comfortable indoor environment for occupants. The concept of indoor thermal comfort involves creating an environment where people feel neither too hot nor too cold and typically takes into account various factors such as air temperature, humidity, radiant temperature, air velocity, clothing insulation, and metabolic heat produced by human activities.

The study [18] shows that people have a significant willingness to pay for all aspects of comfort, with thermal comfort being the most valued of all. It should be noted that indoor thermal comfort temperatures can vary between 15.0°C and 33.8°C depending on factors such as climate, ventilation mode, building type, and age of occupants [19]. Common behavioral adaptations to regulate the indoor thermal environment include adjustment of clothing, use of fans, use of air conditioning, and opening of windows. These adaptations are influenced by environmental and demographic factors and highlight the importance of implementing building designs that take thermal comfort into account. Such factors must be considered in the implementation and operation of BEMS in heating, ventilation, and air conditioning (HVAC) systems related to thermal comfort control.

PID control and RL control are two different approaches with applications in energy systems and many other areas. Each approach offers specific advantages and challenges, and their applicability often depends on the nature of the control problem, the level of uncertainty and complexity of the environment, and the specific objectives of the control system. Research [20] compares the performance of RL and PID controllers in performing cold shutdown operations in nuclear power plants (NPPs). The RL-based controller shows superior performance in both reducing the operating errors and achieving the operating objectives faster than the PID-based controllers.

The study discusses the implications of incorporating RL controllers in NPPs, such as reduced operator burden and minimized component manipulation, while also noting the regulatory challenges posed by artificial intelligence (AI) based systems in critical safety environments. A hybrid approach can also be applied to control systems where RL algorithms provide continuous adaptation of the PID coefficients. For example, research [21] addresses the limitations of classical PID controllers, which require manual retuning under changing system dynamics that can disrupt industrial production. The adaptive Proportional Integral (PI) controller uses the Deep Deterministic Policy Gradient with a Twin Delayed reinforcement learning algorithm. This approach eliminates the need for a pre-defined system model, as the controller adjusts in real-time to optimize performance.

Researches [22–24] provide a comprehensive review of RL applications in the field of energy systems, addressing the increasing complexity caused by the integration of renewable energy sources and the transition to more sustainable energy practices. The authors present a classification of RL papers into categories based on their application areas, including building energy management, dispatch problems, energy markets, and grid operations. Key findings include a demonstrated 10%–20% performance improvement with RL over traditional models. The study also highlights the prevalence of a model-free RL algorithm (Q-learning technique) and the underutilization of batch RL algorithms.

The use of RL algorithms helps to improve the efficiency of building energy systems while taking into account the value of indoor comfort. For example, research [25] shows an average energy savings of 8% over seven months, with peak savings of 16% during the summer. The RL model also showed a 10.2% improvement in photovoltaic (PV) energy use and load shifting compared to rule-based methods, without compromising comfort, which was maintained within user-defined thresholds over 99% of the time. This autonomous approach highlights significant reductions in energy consumption and operating costs while maintaining high levels of occupant comfort. Another study [26] explores the use of multi-agent RL to optimize the control of thermostatically controlled loads without the need for exhaustive sensing or detailed models. Using a large-scale pilot in the Netherlands with over 50 houses, the study demonstrated significant energy savings (almost 200 kWh per household per year) through a collaborative multi-agent system. A case study [27] of an office building in Pennsylvania demonstrates the methodology, including steps such as modeling, calibration, and training. The RL-based control policy achieved thermal energy savings of 15% by controlling the heating system's feed water temperature. The framework promises improvements in both energy efficiency and indoor thermal comfort. This research [28] introduces a reinforcement learning control strategy, specifically using model-free Q-learning, to optimize the operation of HVAC and window systems to improve building energy efficiency and indoor comfort. Case studies in Miami and Los Angeles demonstrated reductions in HVAC system energy consumption (13% in Miami and 23% in Los Angeles).

RL controllers can adapt to changing conditions in real-time, ensuring optimal performance even as building occupancy or environmental conditions change. But while RL controllers offer theoretical benefits, comprehensive cost-benefit analyses are often lacking, making it difficult to justify the higher

initial investment. Ensuring that RL-based systems perform reliably and consistently in real-world scenarios remains a significant challenge, and existing research often lacks long-term deployment and validation studies. Integrating RL controllers into existing building management systems and ensuring compatibility with different building components is a practical challenge. In addition, reinforcement learning methods typically require high computing power and extensive training time, which can limit their practical use. Successful implementation of RL controllers requires a large amount of data for training, which can be difficult to obtain and maintain in real-time. Therefore, the development of effective approaches to training, neural network layer design, and reward functions of such models remains an important task.

3 Study Area

3.1 Object of the Test Modelling

To analyze the energy characteristics of the building, an energy model representing a standard two-bedroom apartment typical of modern economy-class construction was developed using MATLAB/Simscape and Python programming environments.

This apartment is situated on the fourth floor of a five-story residential building, completed in 2016. It covers a total area of 49.44 m² and features a ceiling height of 2.7 m (refer to Fig. 1). The windows of the apartment face east and west, while a solid exterior wall is oriented towards the north. The fenestration consists of double-glazed, energy-efficient insulated glass units filled with argon. External walls consist of a load-bearing structure made from 0.4 m red hollow bricks, augmented with 0.05 m of mineral wool insulation. Natural ventilation in the apartment operates with an air exchange rate of 0.6 per hour. For climatic conditions, hourly data from a typical year, sourced from the International Weather for Energy Calculation (IWEC) weather file [29], were applied (refer to Fig. 2).



Figure 1: Apartment plan

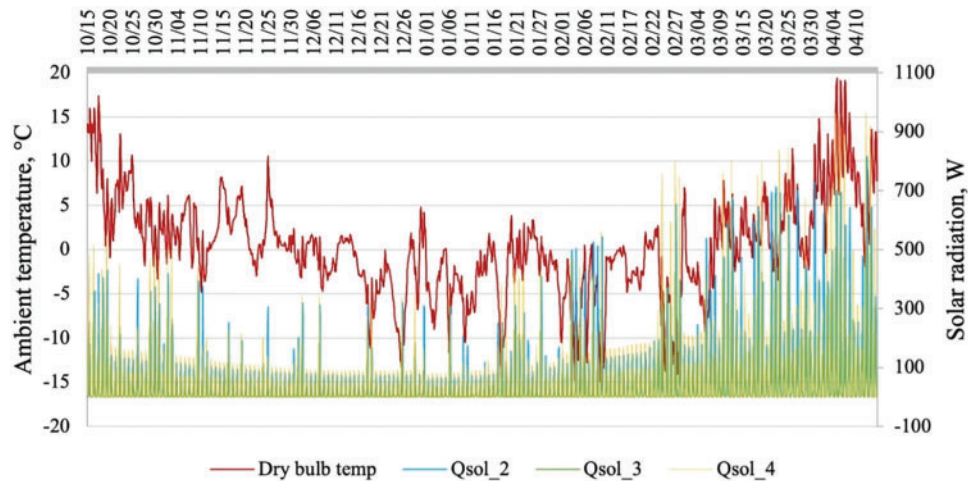


Figure 2: Climate data according to the IWEC weather file (Q_{sol_i} —incoming solar radiation into the i -th room, W/m^2)

3.2 General Features of the Model

To examine the energy performance of a building, a dynamic energy model of a typical modern economy-class two-bedroom apartment was developed using MATLAB/Simscape software and the Python programming environment. This model was constructed within MATLAB's environment by utilizing Simulink subsystems enhanced with Simscape tools. Simscape facilitates the swift creation of physical system models that can be seamlessly integrated with Simulink's block diagrams, focusing on the development of models based on physical connections of components.

The MATLAB environment supports the design of control systems that incorporate the physical attributes of systems modeled in Simulink. This approach accounts for the inertial properties of the building's materials and systems, alongside the dynamic variability of climate data. Consequently, the hybrid model provides outputs such as indoor air temperature, surface temperatures driven by radiation, and the demand on the HVAC system, among other performance metrics [12].

The interconnections of rooms/zones of the apartment in Simulink are shown in Fig. 3. The specific settings of the building's heating system are presented in [12].

3.3 A Model of a Heating System for the Building

This model provides an in-depth examination of a simulation for a heating system that incorporates different control mechanisms. This system integrates several components such as boiler, radiators, room, and building envelope as well as control strategies such as on/off, PID, and RL controllers. Such a system simulates the dynamic heating process in buildings and allows for expansion through different control methods.

Advanced building management systems require sophisticated heating controls to balance comfort and energy efficiency. The system initializes with key parameters such as target temperature, control type, and room-specific configurations. It incorporates iterative calculations to simulate heating dynamics over a predefined time series.

- The radiator component of the model is designed to simulate the thermal behavior of a radiator used in a heating system. It models the heat transfer properties of the radiator and calculates the heating power delivered to a room given specific parameters. This simulation is useful

for understanding thermal dynamics in rooms. The heat method calculates the heating power delivered to a room based on the room temperature and the temperatures of the supply and return water. Within this component, several steps are performed:

- Ensure the supply water temperature is not less than the return water temperature by adjusting it if necessary.
- Calculate the heat loss from the radiator to the room based on the room temperature and the surface area of the radiator. The heat loss is zero if the supply water temperature is less than the room temperature; otherwise, it is calculated using the convective heat transfer coefficient.
- Calculate the heat storage capacity of the system, considering both the water and steel components of the radiator.
- Update the return water temperature based on the amount of heat lost and the heat storage capacity.
- Calculate the heating power delivered by the radiator. This is based on the mass flow rate, specific heat capacity of water, and the temperature difference between the supply and return water.
- The component returns the heating power, supply and return water temperature, and mass flow rate.

The heat transfer coefficient of the radiator was determined in relation to its external surface:

$$k_R = \left(\frac{1}{\alpha_{air}} + \frac{1}{\alpha_{wf}} \cdot \frac{A_{air}}{A_{wf}} \right)^{-1}, \quad (1)$$

where:

k_R —heat transfer coefficient in the radiator, $W/(m^2 \cdot K)$;

A_{air} and A_{wf} —areas of heating surfaces of the radiator from the air side and from the waterside, respectively, m^2 ;

α_{air} and α_{wf} —corresponding heat transfer coefficients from the air side and from the waterside, respectively, $W/(m^2 \cdot K)$.

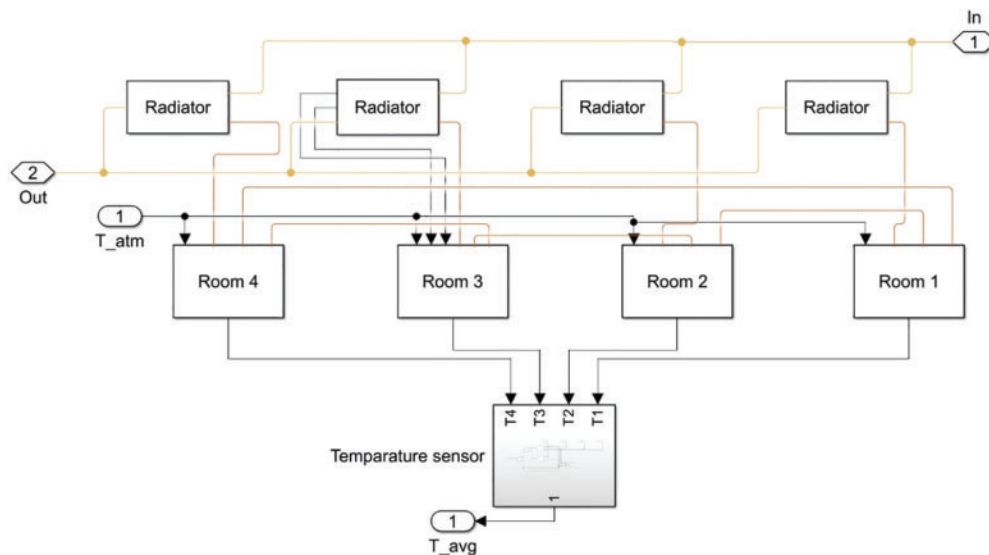


Figure 3: Block diagram of apartment thermal connections between building zones

The room component of the model enables simulation and calculation of thermal dynamics within a building's room by considering heat loss to the external environment and heat gain through solar radiation. Each room component calculates heat loss coefficient, volume, solar gain capacity, indoor temperature, and radiator components as heating devices of the room. Building component encapsulates certain characteristics and behaviors related to temperature management across multiple rooms within a building. As input, each room receives the data that characterize the radiators of the room, such as mass flow rate, mass of steel, and the surface area of the radiator.

The boiler component provides a simulation model of the heating dynamics of an electric boiler. The model accepts parameters for power input and return water temperature and computes the resultant supply water temperature based on the specific heat capacity of water and the boiler's mass flow rate. Electric power input is calculated from the controller and used as input in this component.

The energy balance equation for the water heated in the boiler:

$$\frac{1}{2}c_w M_w \frac{d(T_{w,in} + T_{w,out})}{d\tau} = G_w (h_{w,in} - h_{w,out}) + k_R A_R \left(T_{w,R} - \frac{T_{w,in} + T_{w,out}}{2} \right), \quad (2)$$

where:

c_w —specific heat capacity of water in the electric boiler, J/(kg·K);

M_w —mass of water in the electric boiler, kg;

$T_{w,in}$, $T_{w,out}$ —temperature of water at the inlet and outlet of the electric boiler, respectively, K;

G_w —flow rate of water through the electric boiler, kg/s;

$h_{w,in}$, $h_{w,out}$ —specific enthalpy of water at the inlet and outlet of the electric boiler, respectively, J/kg;

A_R —heat transfer area in the radiator, m²;

$T_{w,R}$ —temperature of the water in the radiator, K.

The system, as a main component, encapsulates all described components and forms a dynamic simulation model of a building's heating system, incorporating multiple control strategies to optimize indoor thermal comfort and energy efficiency. Through detailed parameterization and modular component integration, it enables comprehensive analysis and performance tracking against varying environmental and operational conditions. The system serves as a model for simulating the thermal dynamics of a building with multiple rooms, each equipped with radiators and controlled by different types of controllers (on/off, PID, RL). Such an approach allows to simulation of the heating system under different conditions to determine the most efficient control strategy in BEMS.

In addition, the model allows to computation of energy variations within the building contingent on external conditions, ensures boiler operates within specified temperature limits, updates water temperatures in boiler and radiators considering heat exchange efficiency, adjusts room temperatures based on net energy flux, determines the required boiler power output via the chosen controller, track performance metrics such as fuel efficiency and occupant comfort ad log historical data on room temperatures, energy changes, comfort metrics, and fuel consumption.

The development of the model was carried out under the following simplifications:

It is assumed that the expansion of the water in the expansion valve is isenthalpic;

The mass flow rate of the working fluid is constant in all components of the heating system;

The mass of the water in the system remains constant during the transient process;

The temperatures of the heat carriers and the heat consumer change linearly;

Pressure losses in the components are insignificant and can be neglected;

Superheating of the water in the heating system is absent;

The amount of energy supply is equal to the amount of energy consumption in the heating system since the energy source is located directly in the building.

As the first option, the “on/off” controller is a typical implementation of a hysteresis-based control system used to regulate the boiler’s state based on temperature measurements, thereby aiming to maintain a desired temperature range and reduce the frequency of switching actions. This is commonly worthwhile in systems where is needed to avoid frequent on-off switching by having distinct thresholds for turning the system on and off. In this case, the controller is managing a boiler system based on temperature thresholds.

As the second option to control the indoor temperature of the building, the PID controller can be used. A PID controller is a widely used control loop feedback mechanism commonly used in industrial control systems.

The PID controller computes the control input $u(t)$ for a system, based on the error signal $e(t)$, which is the difference between the desired setpoint $r(t)$ and the measured process variable $y(t)$. The formula is as follows [20,21]:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt}, \quad (3)$$

where:

$u(t)$ —control input at time t ;

K_p —proportional gain parameter;

$e(t) = r(t) - y(t)$ —the error at time t ;

K_i —integral gain parameter;

$\int_0^t e(\tau)$ —integral of the error over time;

K_d —derivative gain parameter;

$\frac{de(t)}{dt}$ —derivative of the error concerning time.

Add to this, setpoint is the desired value that the system should achieve; K_p directly scales the error, and a higher K_p value increases the system’s responsiveness; K_i accumulates the error over time, aiming to eliminate the residual steady-state error; K_d is a prediction of future error, based on its rate of change, and a higher K_d value dampens oscillations.

As the third option, the RL controller is used in the model. RL controller uses TensorFlow and Keras libraries, specifically for managing a boiler’s power. The agent follows the Deep Q-Network (DQN) approach, utilizing an epsilon-greedy policy to learn the optimal control strategy. Sequential memory is used to store and replay experiences, allowing the agent to learn from historical data, generally with a fixed memory limit. The model initialization and weights loading suggest transfer learning, enabling rapid deployment of the controller by leveraging pre-trained knowledge. RL approach constructs the model, defines the policy and learning mechanism, and continuously updates the boiler power output based on state inputs, adhering to the pre-trained model behavior.

Before the RL model is trained, the heating system environment should be initialized to provide model-based learning. Based on the described model of the heating system, the environment uses the

simulation result at time t to return the action and reward values at each step of the learning process. RL agent maximizes the reward to achieve the most efficient performance of the controller. The environment is structured to emulate a scenario where the temperature inside a building is controlled by a heating system, and the setup adjusts the boiler power based on both external and internal factors to reach a target temperature. Each step during agent training updates the system with the new boiler power, outside temperature, and solar radiation, and computes the resultant indoor temperature. The action creator function determines the actual power supplied to the boiler based on the action taken. The penalty is calculated based on the absolute difference between the target temperature and the current temperature. This is capped at a maximum penalty:

$$P = |T_t - T_c|, \quad (4)$$

where:

P —penalty, K;

T_t —target temperature, K;

T_c —current temperature, K;

If the penalty P exceeds the maximum penalty P_{max} , it is capped at P_{max} . This ensures that the penalty value does not grow unbounded and influences the reward calculation in a controlled manner. The reward R is then calculated based on the capped penalty P . The formula used is:

$$R = \frac{P_{max} - P}{P_{max}}, \quad (5)$$

where:

R —reward;

P_{max} —maximum penalty, K;

This formula normalizes the reward within the range of 0 to 1, where:

A lower penalty value results in a higher reward.

A zero penalty (perfect match between current and target temperature) yields the maximum reward of 1.

If the penalty is at its maximum $P = P_{max}$, the reward will be zero.

The neural network model is built using the Keras library, specifically the sequential model. The architecture is designed to handle the state space and action space typical in reinforcement learning tasks.

Input Shape: The input to the model is a state-space tensor. This is flattened before feeding into the dense layers. The input tensor is flattened to ensure it is compatible with the subsequent dense layers. As inputs the current average indoor temperature and total gained solar radiation are used.

Hidden Layers:

Layer 1: Dense, number of units: 128, activation function: ReLU.

Layer 2: Dense, number of units: 128, activation function: ReLU.

Layer 3: Dense, number of units: 128, activation function: ReLU.

Three dense layers, each with 128 units, are used to capture complex patterns in the input data. Each dense layer is followed by a rectified linear unit (ReLU) activation function, which introduces non-linearity and enables the model to learn more complex representations.

Output Layer: Dense, number of units: action space length, activation function: Linear. The final dense layer's units correspond to the action space length, representing the possible actions in the environment. A linear activation function is used in the output layer, commonly used in Q-learning tasks and other scenarios where the output is continuous. As output, the model returns the index of the action where action space consists of 12 actions that represent the state of the power of the electric boiler. Action with value -1 means that the boiler is disabled, 0 means the boiler remains in the same state from the previous iteration, and the actions from 1 to 5 define the level of the boiler power based on the maximum boiler power in the heating system.

Using the deep learning framework for building and training the neural network, the agent utilizes DQN with sequential memory and an epsilon-greedy policy, trained and evaluated within a developed custom environment with a heating system model of the building.

4 Engineering Calculations

An individual electric boiler was selected to evaluate various heat supply controllers for the building. The room air temperature control utilizes the average values from rooms 3 and 4 (living rooms), aiming to maintain a setpoint temperature of 20°C . The justification for sensor placement and quantity is provided in Reference [30]. The heating system's heat carrier flow rate is 0.998 kg/s .

For the design conditions, the heating devices were selected based on a mathematical model developed in the MATLAB software environment, considering the type, area, mass, and thermal inertia of the devices. The design conditions assumed an exterior temperature of -22°C without solar heat gains, with an indoor air temperature of 20°C . The heating power required for each room is as follows: bedroom 1/room 4– 500 W , bedroom 2/room 3– 1200 W , kitchen/room 2– 880 W , and common areas (corridor, bathroom)/room 1– 320 W , which may vary with external conditions and boiler regimes.

To validate the apartment's mathematical model in the MATLAB environment, results were compared using the quasi-steady-state method, detailed in Reference [30]. Additionally, a comparison was made with simulation outcomes from the EnergyPlus software environment [31]. Both the MATLAB and EnergyPlus models consider the thermal inertia properties of each building enclosure individually. Research [12] indicates that quantitative regulation of heating, without the inertia of heating devices, based on the EnergyPlus model, results in the lowest heat consumption for the building. The deviation of MATLAB simulation results compared to EnergyPlus ranges from 4% – 10% , depending on the heating month [12]. Model validation was conducted using an independent electric boiler.

The study evaluates the use of “on/off”, PID, and RL controllers to calculate and compare thermal comfort conditions and electric boiler energy efficiency by using each of the listed controllers. Fig. 4 shows the hourly load on the heating system (Watts).

Fig. 4 shows that the boiler, when switched on in hysteresis mode, delivers close to the maximum installed power of 3.5 kW . The PID and RL controllers have more flexible regulation, and the load level on the heating system depends on the required amount of thermal energy at a given moment. The use of a PID and RL allows reducing the instantaneous power value to 35% , but the frequency of boiler activation is significantly higher. Fig. 5 shows the monthly energy consumption for heating purposes.

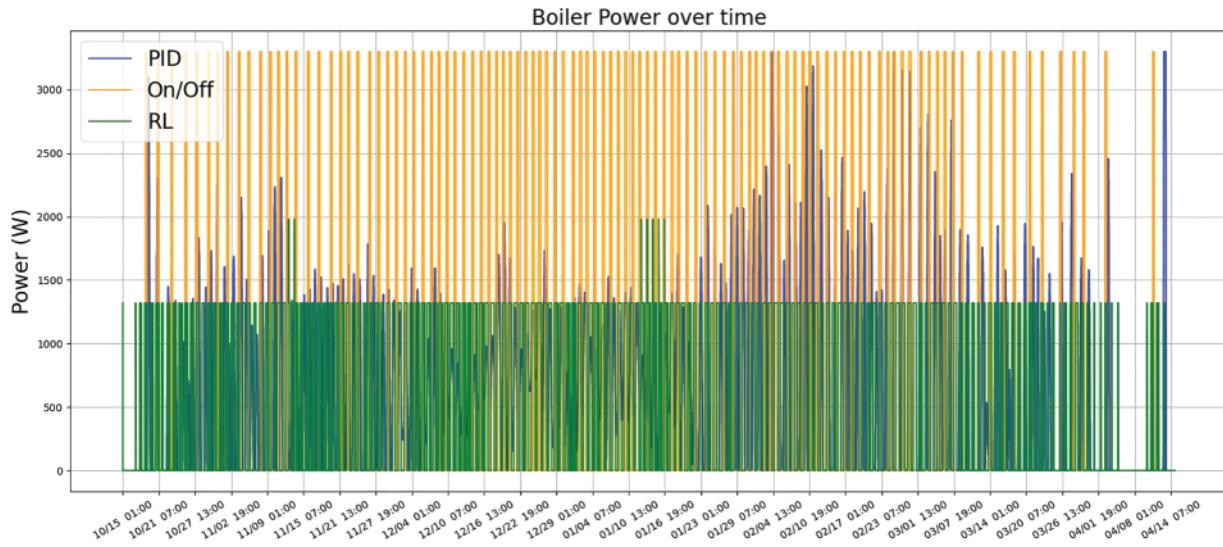


Figure 4: Dynamics of changes in the load on the heating system

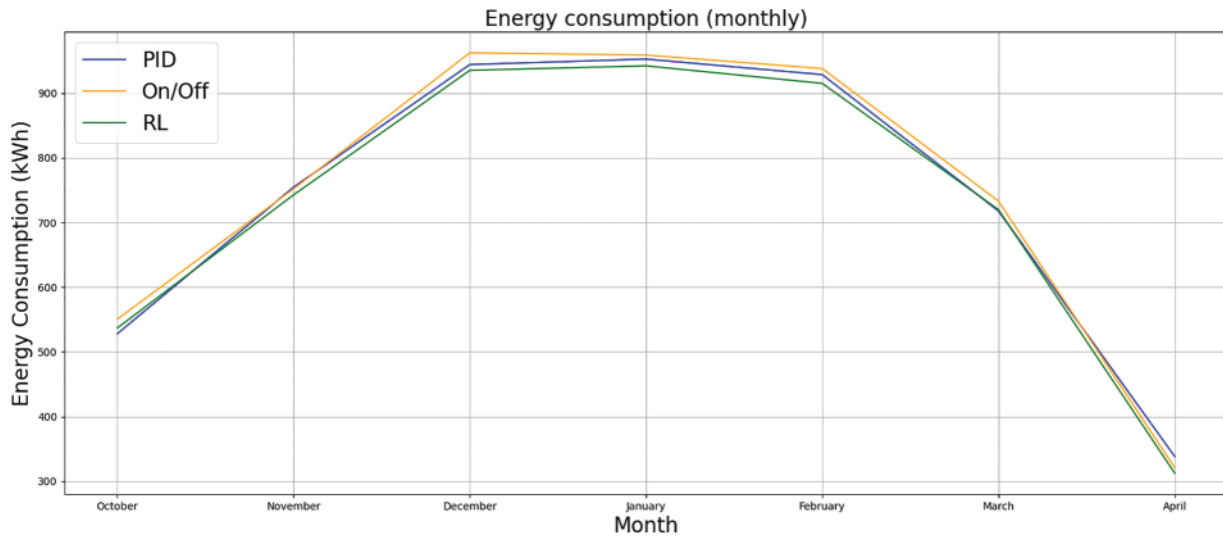


Figure 5: Monthly comparison of heating system energy consumption using different controllers

It appears that the RL controller exhibits superior performance in regulating room air temperature within specified areas compared to the electric boiler. Notable variations in room air temperature occur at the onset and conclusion of the heating season. These fluctuations are likely caused by increased solar activity during the day in these transitional periods, resulting in the temporary deactivation of the heating system. Consequently, due to passive solar heating, the room air temperature can temporarily rise to 21°C. Fig. 6 illustrates the average monthly air temperatures in representative rooms for the controller types.

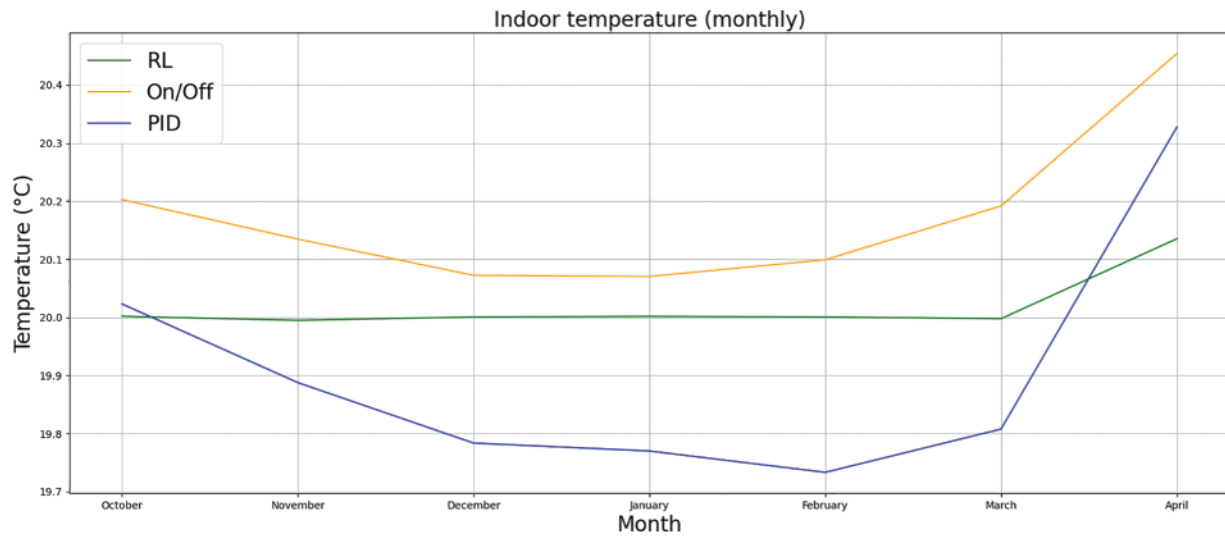


Figure 6: Average monthly indoor air temperature controlled by different types of controller

The RL and PID controllers better maintain the set level of air temperature in the room and prevent overheating. Fig. 7 shows the hourly average air temperature across all rooms in the apartment.

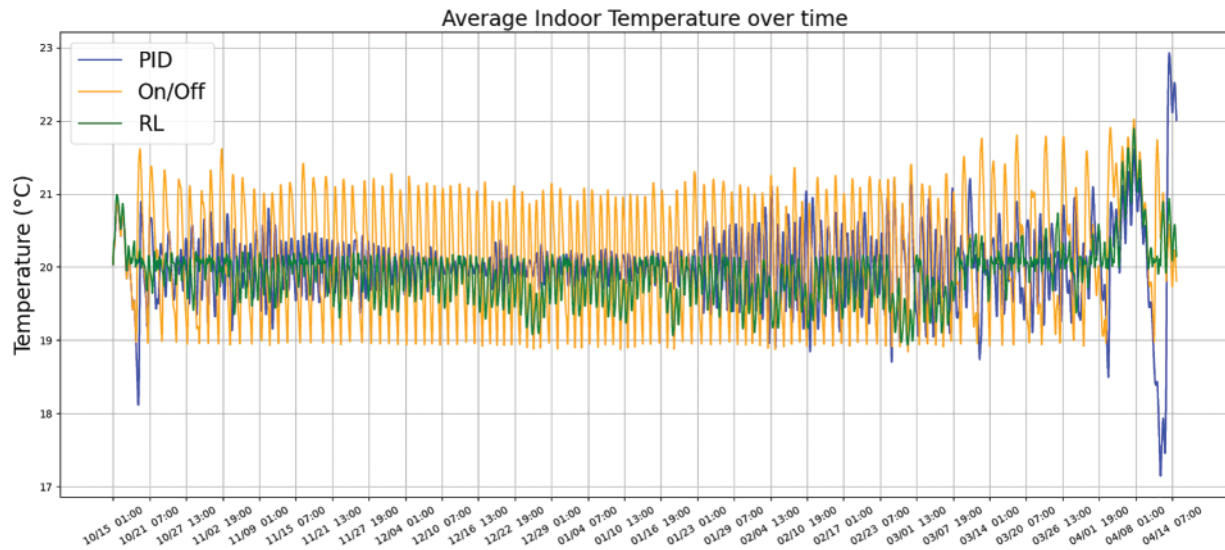


Figure 7: Hourly average air temperature across rooms in the apartment

Due to different room occupancy schedules, heat flows between rooms, and various orientations, different temperatures are observed across all rooms, and the average temperature fluctuation across all rooms exceeds the setpoint on the controller ($19.5^{\circ}\text{C} \dots 20^{\circ}\text{C}$), which should correspond to the temperature in representative rooms (bedrooms 1 and 2). The simulation result shows that the deviation of the temperature from the set value during the operation of the controllers reaches different values, in particular, during the “on/off” operation of the controller, the deviation is 0.66°C , the PID of the controller is 0.35°C , and the RL of the controller is 0.27°C . At the same time, an overconsumption

of electrical energy in the amount of 60 kWh during PID controller operation and 100 kWh during on/off controller operation was observed relative to the RL controller.

5 Discussion

Dynamic modeling stands out as a highly effective approach for addressing this issue because it is in the analysis of system dynamics that the strengths and weaknesses of implementing specific management measures for the concurrent improvement of energy efficiency and indoor climatic conditions in buildings become evident. Existing software solutions typically specify energy sources from pre-defined libraries, limiting the capacity to analyze processes within the structural components of these sources, particularly in transient states. This work presents an energy-dynamic model of an energy supply system designed within the MATLAB/Simscape environment, which enables detailed examination of the components of various types of energy sources.

While traditional on/off and PID controllers have long been the backbone of building energy management, the increased complexity and dynamic behavior of modern buildings necessitate more sophisticated control techniques. Reinforcement Learning (RL) controllers offer substantial promise due to their adaptive nature, optimization potential, and self-learning capabilities. Once adequately trained, RL controllers require minimal human supervision, potentially lowering operational costs and reducing the need for specialized personnel. Nonetheless, considerable challenges persist, such as computational demands, data requirements, integration difficulties, and the necessity for extensive validation.

Overall, the dynamic modeling of energy processes in buildings employed in this research enabled the evaluation of comfortable indoor air temperature maintenance, considering the effective operation characteristics of the energy supply source. It has been demonstrated that an energy management system based on an RL controller more effectively sustains the desired indoor air temperature and mitigates overheating with reduced energy production.

The practical implications of the findings are significant, providing a thorough assessment of the building's dynamic characteristics, including energy consumption and indoor temperature levels. The insights gained from this study can inform future research on large-scale energy systems, where reinforcement learning algorithms are increasingly utilized to enhance energy efficiency and the reliability of energy supply to consumers. However, these studies often overlook the specificities of energy management within buildings, which are the highest energy consumers in most countries worldwide. There is a critical need to develop a model that addresses the contemporary energy usage areas in buildings, such as prosumers, energy cooperatives, energy clusters, and smart grids.

6 Conclusions

Enhancing energy efficiency while simultaneously improving the comfort levels of climatic conditions in residential or occupational buildings is a multifaceted issue without a straightforward solution. This complexity arises because improving comfort often necessitates increased energy and/or material consumption, leading to additional economic costs. Consequently, these improvements do not always result in a net reduction of total emissions, considering all phases of energy and material production and usage aimed at boosting energy efficiency and comfort in buildings.

The outcomes of addressing this challenge are illustrated through an energy-dynamic grid model of a standard residential building. The study explores the potential of utilizing an electric boiler controlled by a Reinforcement Learning RL-based controller. From the perspective of the end user,

various types of controllers offer nearly equivalent value over annual intervals. However, the RL-based controller offers more adaptable system load-level adjustment, responding to internal and external condition variations, thereby enabling up to a 35% reduction in instantaneous power usage.

The study's findings highlight the RL controller's effectiveness. Utilizing input data that reflect heating season parameters, the RL controller reduced absolute deviations from comfortable temperatures by nearly half compared to traditional energy resource management methods. Additionally, energy consumption was reduced by approximately 1% while maintaining comfort. These results indicate a synergistic effect when implementing the RL controller in the heating system, underscoring its significant value.

Acknowledgement: The authors thank the reviewers and editors for their valuable contributions that significantly improved this manuscript.

Funding Statement: The authors received no specific funding for this study.

Author Contributions: Conceptualization, Inna Bilous and Oleksandr Novoseltsev; methodology, Dmytro Biriukov and Dmytro Karpenko; software, Dmytro Biriukov, Dmytro Karpenko and Volodymyr Voloshchuk; validation, Dmytro Karpenko, Tatiana Eutukhova and Oleksandr Novoseltsev; formal analysis, Inna Bilous, Dmytro Biriukov, Dmytro Karpenko, Tatiana Eutukhova, Oleksandr Novoseltsev and Volodymyr Voloshchuk; investigation, Dmytro Biriukov and Volodymyr Voloshchuk; resources, Dmytro Biriukov and Dmytro Karpenko; data curation, Inna Bilous, Tatiana Eutukhova and Oleksandr Novoseltsev; writing—original draft preparation, Inna Bilous, Dmytro Biriukov, Dmytro Karpenko, Tatiana Eutukhova, Oleksandr Novoseltsev and Volodymyr Voloshchuk; writing—review and editing, Inna Bilous, Dmytro Biriukov, Dmytro Karpenko, Tatiana Eutukhova, Oleksandr Novoseltsev and Volodymyr Voloshchuk; visualization, Dmytro Biriukov, Dmytro Karpenko and Volodymyr Voloshchuk; supervision, Inna Bilous, Tatiana Eutukhova and Oleksandr Novoseltsev. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Not applicable.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] M. Jradi, "Dynamic energy modelling as an alternative approach for reducing performance gaps in retrofitted schools in Denmark," *Appl. Sci.*, vol. 10, no. 21, Nov. 2020, Art. no. 7862. doi: [10.3390/app10217862](https://doi.org/10.3390/app10217862).
- [2] O. Kovalko, T. Eutukhova, and O. Novoseltsev, "Energy-related services as a business: Eco-transformation logic to support the low-carbon transition," *Energy Eng.*, vol. 119, no. 1, pp. 103–121, 2022. doi: [10.32604/EE.2022.017709](https://doi.org/10.32604/EE.2022.017709).
- [3] A. Zaporozhets, V. Babak, G. Kostenko, A. Sverdlova, O. Dekusha and S. Kornienko, "Some features of air pollution monitoring as a component of the microclimate of the premises," *Syst. Res. Energy*, vol. 2023, no. 4, pp. 65–73, Nov. 2023. doi: [10.15407/srenergy2023.04.065](https://doi.org/10.15407/srenergy2023.04.065).

- [4] O. Kon, B. Yuksel, and A. D. Karaoglan, "Modeling the relationship between outdoor meteorological data and energy consumptions at heating and cooling periods: Application in a university building," *Numer. Algebra, Control Optim.*, vol. 14, no. 1, pp. 59–79, 2022. doi: [10.3934/naco.2022021](https://doi.org/10.3934/naco.2022021).
- [5] D. Han, M. Kalantari, and A. Rajabifard, "The development of an integrated BIM-based visual demolition waste management planning system for sustainability-oriented decision-making," *J. Environ. Manag.*, vol. 351, Feb. 2024, Art. no. 119856. doi: [10.1016/j.jenvman.2023.119856](https://doi.org/10.1016/j.jenvman.2023.119856).
- [6] G. Ulpiani, M. Borgognoni, A. Romagnoli, and C. Di Perna, "Comparing the performance of on/off, PID and fuzzy controllers applied to the heating system of an energy-efficient building," *Energy Build*, vol. 116, pp. 1–17, Mar. 2016. doi: [10.1016/j.enbuild.2015.12.027](https://doi.org/10.1016/j.enbuild.2015.12.027).
- [7] M. Dulău, "Experiments on temperature control using on-off algorithm combined with PID algorithm," *Acta Marisiensis. Ser. Technol.*, vol. 16, no. 1, pp. 5–9, Jun. 2019. doi: [10.2478/amset-2019-0001](https://doi.org/10.2478/amset-2019-0001).
- [8] L. Yu, S. Qin, M. Zhang, C. Shen, T. Jiang and X. Guan, "A review of deep reinforcement learning for smart building energy management," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12046–12063, Aug. 2021. doi: [10.1109/JIOT.2021.3078462](https://doi.org/10.1109/JIOT.2021.3078462).
- [9] A. Shaqour and A. Hagishima, "Systematic review on deep reinforcement learning-based energy management for different building types," *Energies*, vol. 15, no. 22, Nov. 2022, Art. no. 8663. doi: [10.3390/en15228663](https://doi.org/10.3390/en15228663).
- [10] Y. Bae *et al.*, "Sensor impacts on building and HVAC controls: A critical review for building energy performance," *Adv. Appl. Energy.*, vol. 4, no. 3, Nov. 2021, Art. no. 100068. doi: [10.1016/j.adapen.2021.100068](https://doi.org/10.1016/j.adapen.2021.100068).
- [11] G. Liu, X. Zhou, J. Yan, and G. Yan, "A temperature and time-sharing dynamic control approach for space heating of buildings in district heating system," *Energy*, vol. 221, no. 9, Apr. 2021, Art. no. 119835. doi: [10.1016/j.energy.2021.119835](https://doi.org/10.1016/j.energy.2021.119835).
- [12] V. Deshko, I. Bilous, D. Biriukov, and O. Yatsenko, "Transient energy models of housing facilities operation," *Rocz. Ochr. Srodowiska*, vol. 23, pp. 539–551, 2021. doi: [10.54740/ros.2021.038](https://doi.org/10.54740/ros.2021.038).
- [13] V. Deshko, I. Bilous, and O. Maksymenko, "Impact of heating system local control on energy consumption in apartment buildings," *Rocz. Ochr. Srodowiska*, vol. 25, pp. 77–85, Jun. 2023. doi: [10.54740/ros.2023.009](https://doi.org/10.54740/ros.2023.009).
- [14] N. Alibabaei, A. S. Fung, K. Raahemifar, and A. Moghimi, "Effects of intelligent strategy planning models on residential HVAC system energy demand and cost during the heating and cooling seasons," *Appl. Energy*, vol. 185, no. 6, pp. 29–43, Jan. 2017. doi: [10.1016/j.apenergy.2016.10.062](https://doi.org/10.1016/j.apenergy.2016.10.062).
- [15] V. Deshko, I. Bilous, I. Sukhodub, and O. Yatsenko, "Evaluation of energy use for heating in residential building under the influence of air exchange modes," *J. Building Eng.*, vol. 42, no. 3, Oct. 2021, Art. no. 103020. doi: [10.1016/j.jobe.2021.103020](https://doi.org/10.1016/j.jobe.2021.103020).
- [16] K. Hinkelman, S. Anbarasu, and W. Zuo, "Exergy-based ecological network analysis for building and community energy systems," *Energy Build*, vol. 303, no. 12, Nov. 2023, Art. no. 113807. doi: [10.1016/j.enbuild.2023.113807](https://doi.org/10.1016/j.enbuild.2023.113807).
- [17] Y. Song, F. Mao, and Q. Liu, "Human comfort in indoor environment: A review on assessment criteria, data collection and data analysis methods," *IEEE Access*, vol. 7, pp. 119774–119786, 2019. doi: [10.1109/ACCESS.2019.2937320](https://doi.org/10.1109/ACCESS.2019.2937320).
- [18] R. Berto, F. Tintaglia, and P. Rosato, "How much is the indoor comfort of a residential building worth? A discrete choice experiment," *Building Environ.*, vol. 245, Oct. 2023, Art. no. 110911. doi: [10.1016/j.buildenv.2023.110911](https://doi.org/10.1016/j.buildenv.2023.110911).
- [19] F. S. Arsad, R. Hod, N. Ahmad, M. Baharom, and M. H. Ja'afar, "Assessment of indoor thermal comfort temperature and related behavioural adaptations: A systematic review," *Environ. Sci. Pollut. Res.*, vol. 30, no. 29, pp. 73137–73149, May 2023. doi: [10.1007/s11356-023-27089-9](https://doi.org/10.1007/s11356-023-27089-9).
- [20] D. Lee, S. Koo, I. Jang, and J. Kim, "Comparison of deep reinforcement learning and PID controllers for automatic cold shutdown operation," *Energies*, vol. 15, no. 8, Apr. 2022, Art. no. 2834. doi: [10.3390/en15082834](https://doi.org/10.3390/en15082834).

- [21] U. Alejandro-Sanjines, A. Maisincho-Jivaja, V. Asanza, L. L. Lorente-Leyva, and D. H. Peluffo-Ordóñez, “Adaptive PI controller based on a reinforcement learning algorithm for speed control of a DC motor,” *Biomimetics*, vol. 8, no. 5, Sep. 2023, Art. no. 434. doi: [10.3390/biomimetics8050434](https://doi.org/10.3390/biomimetics8050434).
- [22] A. T. D. Perera and P. Kamalaruban, “Applications of reinforcement learning in energy systems,” *Renew. Sustain. Energy Rev.*, vol. 137, no. 4, Mar. 2021, Art. no. 110618. doi: [10.1016/j.rser.2020.110618](https://doi.org/10.1016/j.rser.2020.110618).
- [23] D. Cao *et al.*, “Reinforcement learning and its applications in modern power and energy systems: A review,” *J. Modern Power Syst. Clean Energy*, vol. 8, no. 6, pp. 1029–1042, 2020. doi: [10.35833/MPCE.2020.000552](https://doi.org/10.35833/MPCE.2020.000552).
- [24] G. Pesántez, W. Guamán, J. Córdova, M. Torres, and P. Benalcazar, “Reinforcement learning for efficient power systems planning: A review of operational and expansion strategies,” *Energies*, vol. 17, no. 9, May 2024, Art. no. 2167. doi: [10.3390/en17092167](https://doi.org/10.3390/en17092167).
- [25] P. Lissa, C. Deane, M. Schukat, F. Seri, M. Keane and E. Barrett, “Deep reinforcement learning for home energy management system control,” *Energy AI*, vol. 3, no. 1, Mar. 2021, Art. no. 100043. doi: [10.1016/j.egyai.2020.100043](https://doi.org/10.1016/j.egyai.2020.100043).
- [26] H. Kazmi, J. Suykens, A. Balint, and J. Driesen, “Multi-agent reinforcement learning for modeling and control of thermostatically controlled loads,” *Appl. Energy*, vol. 238, no. 3, pp. 1022–1035, Mar. 2019. doi: [10.1016/j.apenergy.2019.01.140](https://doi.org/10.1016/j.apenergy.2019.01.140).
- [27] Z. Zhang, A. Chong, Y. Pan, C. Zhang, and K. P. Lam, “Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning,” *Energy Build*, vol. 199, no. 6, pp. 472–490, Sep. 2019. doi: [10.1016/j.enbuild.2019.07.029](https://doi.org/10.1016/j.enbuild.2019.07.029).
- [28] Y. Chen, L. K. Norford, H. W. Samuelson, and A. Malkawi, “Optimal control of HVAC and window systems for natural ventilation through reinforcement learning,” *Energy Build*, vol. 169, no. 1, pp. 195–205, Jun. 2018. doi: [10.1016/j.enbuild.2018.03.051](https://doi.org/10.1016/j.enbuild.2018.03.051).
- [29] EnergyPlus, “Weather data by location,” Sep. 2001. Accessed: Sep. 19, 2024. [Online]. Available: https://energyplus.net/weather-location/europe_wmo_region_6/UKR/UKR_Kiev.333450_IWEC
- [30] V. Deshko, I. Bilous, and D. Biryukov, “Modeling of unsteady temperature regimes of autonomous heating systems,” *J. New Technol. Environmental Sci.*, vol. 3, pp. 91–99, 2021. doi: [10.53412/jntes-2021-3-1](https://doi.org/10.53412/jntes-2021-3-1).
- [31] EnergyPlus, “EnergyPlus Version 23.2.0,” Sep. 2023. Accessed: Sep. 19, 2024. [Online]. Available: <https://energyplus.net/>