Received: 17 November 2022 | Revised: 25 May 2024 | Accepted: 5 June 2024

IET Generation, Transmission & Distribution

DOI: 10.1049/gtd2.13203

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Reinforcement learning layout-based optimal energy management in smart home: AI-based approach

Sajjad Afroosheh[1,2] | Khodakhast Esapour[3] | Reza Khorram-Nia[3] | Mazaher Karimi[4]

[1]Department of Electrical and Computer Engineering, University of Washington, Seattle, Washington, USA

[2]Department of Physics and Astronomy, Bowling Green State University, Bowling Green, Ohio, USA

[3]Department of Electrical Engineering, Safashahr Branch, Islamic Azad University, Safashahr, Iran

[4]School of Technology and Innovations, University of Vaasa, Vaasa, Finland

Correspondence
Khodakhast Esapour, Department of Electrical Engineering, Safashahr Branch, Islamic Azad University, Safashahr 71382786573, Iran.
Email: kh.isapour@gmail.com

Abstract

This research addresses the pressing need for enhanced energy management in smart homes, motivated by the inefficiencies of current methods in balancing power usage optimization with user comfort. By integrating reinforcement learning and a unique column-and-constraint generation strategy, the study aims to fill this gap and offer a comprehensive solution. Furthermore, the increasing adoption of renewable energy sources like solar panels underscores the importance of developing advanced energy management techniques, driving the exploration of innovative approaches such as the one proposed herein. The constraint coordination game (CCG) method is designed to efficiently manage the power usage of each appliance, including the charging and discharging of the energy storage system. Additionally, a deep learning model, specifically a deep neural network, is employed to forecast indoor temperatures, which significantly influence the energy demands of the air conditioning system. The synergistic combination of the CCG method with deep learning-based indoor temperature forecasting promises significant reductions in homeowner energy expenses while maintaining optimal appliance performance and user satisfaction. Testing conducted in simulated environments demonstrates promising results, showcasing a 12% reduction in energy costs compared to conventional energy management strategies.

1 | INTRODUCTION

In the United States, the energy used in homes is largely affected by various household devices like washers and air conditioners (ACs), contributing to nearly a third of the overall energy use [1]. It is essential to adopt effective and economical strategies for managing home energy to cut down on electricity costs and ensure appliances work at their best. This task gets more complex when adding different energy sources like electric cars, solar panels, and energy storage systems to the mix [2, 3]. These additions, along with modern metering tools and smart meters, are key to keeping the home energy grid stable. In this way, machine learning (ML) techniques, especially those based on data, are becoming more effective in managing home energy. For example, using neural networks to predict solar panel output for better energy management (EM) has been gaining traction. Recent literature on deep Q networks (DQN)-based home energy management systems (HEMSs) has explored the application of advanced reinforcement learning techniques to optimize energy usage in smart homes. These studies leverage DQN algorithms to directly learn policies for energy management tasks, bypassing the need for explicit value function estimation. Researchers have investigated the effectiveness of DQN-based approaches in various aspects of HEMSs, including load scheduling, demand response, and renewable energy integration [4]. By combining deep learning with reinforcement learning, DQN-based HEMSs aim to adaptively control home appliances and energy storage systems in real-time, considering dynamic environmental conditions and user preferences. Studies have demonstrated the potential of DQN-based methods to improve energy efficiency, reduce costs, and enhance overall system performance in smart home environments [5, 6].

Homeowners face the challenge of efficiently managing their appliances, which is where HEMSs come in, provided by utilities and other companies. These systems aim to reduce energy costs and meet user comfort by: (1) tracking energy use with smart meters, and (2) planning energy use for each appliance. The methods for optimizing these systems are a hot topic in

TABLE 1 Comparison of different methods.

Study case	Methodology	Key findings	Comparison
Su et al. (2022) [1]	Bi-level energy management	Improved cost savings through advanced pricing	Our method integrates deep learning (DL) for better forecasts
Qiu et al. (2022) [2]	Data-driven chance-constrained programming	Enhanced energy efficiency in smart grids	Our approach focuses on smart homes
Yang et al. (2023) [3]	Optimal dispatching model	Effective management of off-grid microgrids	Our work addresses on-grid smart homes
Hai et al. (2023) [4]	Renewable-based microgrid	Optimized energy use with electric vehicles	Similar approach but includes ESS
Liu et al. (2023) [5]	Reinforcement learning	Efficient household demand response	Our study combines RL with DL for HEMS

Abbreviations: ESS, energy storage system; HEMS, home energy management systems; RL, reinforcement learning.

research [7, 8]. Studies have investigated how to plan energy use for appliances and energy sources while keeping customers happy. This includes using various mathematical methods for load and appliance planning and exploring both broad and focused HEMS frameworks. Some research has focused on managing household energy with storage systems and electric vehicles [8] and using electric vehicle data to improve HEMSs [9]. Recent literature on DQN-based HEMSs has explored the application of advanced reinforcement learning techniques to optimize energy usage in smart homes. These studies leverage DQN algorithms to directly learn policies for energy management tasks, bypassing the need for explicit value function estimation. Researchers have investigated the effectiveness of DQN-based approaches in various aspects of HEMSs, including load scheduling, demand response, and renewable energy integration. By combining deep learning with reinforcement learning, DQN-based HEMSs aim to adaptively control home appliances and energy storage systems in real-time, considering dynamic environmental conditions and user preferences. Studies have demonstrated the potential of DQN-based methods to improve energy efficiency, reduce costs, and enhance overall system performance in smart home environments. However, challenges such as scalability, sample efficiency, and robustness to environmental uncertainties remain areas of active research in this field [10, 11].

However, the current methods for optimizing HEMSs have drawbacks, like not accurately representing how appliances are used or what consumers want, leading to unreliable energy plans and needing a lot of computing power. To overcome these issues, this study suggests using a reinforcement learning (RL) approach that does not rely on existing data [12, 13]. This research employs a novel method using deep learning (DL) to analyze energy management [14] systems in residential and commercial buildings. Unlike previous methods, this study incorporates the ongoing utilization of diverse home appliances (HAs) and distributed energy resources (DERs), focusing on user satisfaction and requirements. Earlier studies [15–17] adopted a basic quality of life (QL) methodology with a limited, unrealistic action range for scheduling HA and DER operations [18]. Furthermore, prior research [19, 20] focused exclusively on energy consumption (EC) planning in buildings, neglecting HA functionality. Table 1 compares some of the most significant works in this area.

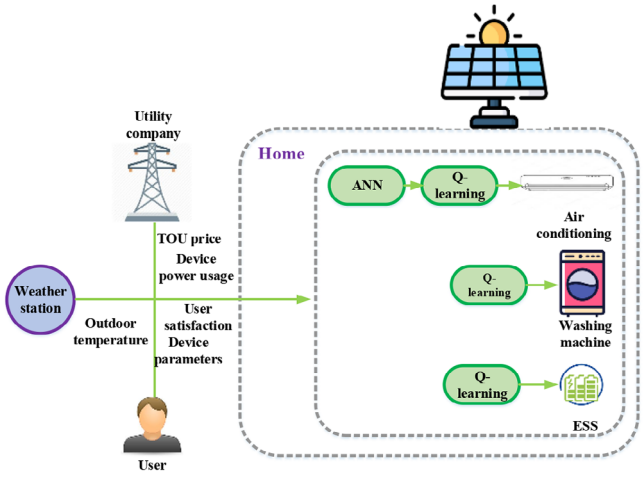

FIGURE 1 A diagram illustrating the suggested home energy management system (HEMS) architecture. ANN, artificial neural network, TOU, time-of-use.

This paper introduces a unique approach, combining column-and-constraint generation with a DL-based HEMS algorithm, diverging from conventional model-driven HEMS optimization techniques. The proposed constraint coordination game (CCG) and DL-enhanced HEMSs are illustrated in Figure 1, including data categorization related to utilities, meteorological stations, and consumer profiles. The main components of this study are as below:

- **Photovoltaic panels (PV panels)**: Monocrystalline silicon solar panels with 20% efficiency and 350 watts per panel power rating.
- **Inverter**: Grid-tied string inverter with 97% efficiency and a maximum power output of 3.5 kW.
- **Mounting and racking system**: Aluminium alloy material suitable for pitched roofs with asphalt shingles.

The study's primary contributions are:

- Development of a CCG method integrated with a DL-based HEMS optimizing energy consumption in a smart home with PV and energy storage system (ESS) devices.

– Implementation of a novel deep learning model for indoor temperature forecasting to enhance energy management efficiency.
– Demonstration of significant energy and cost savings through simulations, highlighting the advantages of the proposed method over traditional approaches.

The remainder of the paper is structured as follows: Section 2 defines various SH devices and introduces the traditional home energy management (HEM) optimization formula. Section 3 details the CCG-driven HEMS algorithm using the DL approach. Section 4 presents simulation results for the proposed HEMS algorithm, and Section 5 concludes the study.

2 | SYSTEM OVERVIEW OF HOME ENERGY MANAGEMENT SYSTEM

2.1 | Preliminary

This section provides an overview of an automatic EM system designed for a single household with time-of-use (TOU) tariffs. The system's primary function is to control and schedule various types of HAs). These appliances can be categorized into:

- Controllable appliances (CAs)-($\mathcal{A}^c$): The HEMS actively manages the operation of CAs. CAs are further classified into two categories based on their functionalities: Shiftable appliances (SAs) ($\mathcal{A}_s^c$) and reducible appliances (RAs) ($\mathcal{A}_r^c$). SAs consist of two load types: (I) non-interruptible load (NIL) ($\mathcal{A}_s^{c,\mathrm{NI}}$), and (II) interruptible load (IL) ($\mathcal{A}_s^{c,I}$). The HEMS ensures uninterrupted operation of SAs with NILs during specific tasks but can interrupt SAs with ILs at any time.
- Uncontrollable appliances (UCAs) ($\mathcal{A}^{uc}$): The HEMS lacks control over UCAs, which include devices like televisions, personal computers, and lighting. Consequently, $\mathcal{A}^{uc}$ maintain a consistent energy consumption plan.

The photovoltaic (PV) system utilized in our research is an on-grid system, meaning it is connected to the main utility grid. This type of system allows for the seamless exchange of electricity between the PV system and the grid, enabling the homeowner to both consume electricity generated by the PV panels and sell excess electricity back to the grid. Components of the PV system in our research include:

1. Photovoltaic panels (PV panels):
   - Type: Monocrystalline silicon solar panels
   - Efficiency: 20%
   - Power rating: 350 W per panel
   - Dimensions: 1.7 m × 1.0 m
   - Number of panels: 10 panels installed in an array
2. Inverter:
   - Type: Grid-tied, string inverter
   - Efficiency: 97%
   - Maximum power output: 3.5 kW
   - Input voltage range: 200–450 V DC
   - Output voltage: 230 V AC, 50 Hz
3. Mounting and racking system:
   - Material: Aluminium alloy
   - Roof compatibility: Suitable for pitched roofs with asphalt shingles
   - Wind load resistance: Designed to withstand wind speeds of up to 150 km/h
   - Snow load capacity: Can support snow loads of up to 30 pounds per square foot
4. Monitoring system:
   - Data monitoring: Real-time monitoring of energy production
   - Communication: Wi-Fi or Ethernet connectivity for remote monitoring
   - Compatibility: Compatible with smartphone apps for easy access to system data
   - Data logging: Records historical energy production data for analysis and troubleshooting

The PV system utilized in our research is an on-grid system, meaning it is connected to the main utility grid. This type of system allows for the seamless exchange of electricity between the PV system and the grid, enabling the homeowner to both consume electricity generated by the PV panels and sell excess electricity back to the grid.

2.2 | Traditional HEMS optimization formula

The algorithm aims to determine the optimal operating plan for HAs and DERs by formulating a mixed-integer linear programming (MILP) optimization problem. This problem consists of an objective function (OF) and a set of constraints:

2.2.1 | Objective function

The OF for the HEMS optimization problem comprises two components involving different decision variables (E_t^{net}, T_t^{in}):

$$\min_{E_t^{\mathrm{net}},T_t^{\mathrm{in}}} \sum_{t\in T} \pi_t E_t^{\mathrm{net}} + \epsilon \sum_{t\in T} \left| T_t^{\mathrm{in}} - T^{\mathrm{set}} \right| \tag{1}$$

In this context, $J_1(E_t^{\mathrm{net}})$ represents the overall energy cost, which is calculated based on the TOU tariffs denoted as π_t and the net energy consumption (E_t^{net}) at each period, t. Furthermore, E_t^{net} can be determined by considering the forecasted PV production output and the energy consumption patterns of CAs and UCAs. $J_2(T_t^{\mathrm{in}})$ denotes the cumulative penalty cost associated with user dissatisfaction. The degree of dissatisfaction is evaluated through the difference between the user's desired indoor temperature, denoted as T^{set}, and the actual indoor temperature, T_t^{in}. Here, ϵ represents the penalty factor linked to user dissatisfaction. It is worth noting that a higher value of ϵ results in a lower $J_2(T_t^{\mathrm{in}})$, reducing user dissatisfaction and,

consequently, lowering power consumption. An operator of the HEMS has the flexibility to adjust ϵ to meet the user's desired satisfaction level without causing a significant increase in the user's energy expenses. The subsequent sections elaborate on the specific inequality and equality constraints inherent in the optimization problem of the HEMS.

2.2.2 | Net energy usage

Equation (2) illustrates the restriction on the net EC, which essentially quantifies the difference between the total energy usage of all devices, denoted as $\sum_{a\in\mathcal{A}} E_{a,t}$, and the estimated PV production output, represented as $\hat{E}_t^{\mathrm{PV}}$. Equation (3) further dissects the overall energy consumption expressed in Equation (2) into four distinct categories of appliances: RAs where $a \in \mathcal{A}_r^c$, SAs using a NIL where ($a \in \mathcal{A}_s^{c,\mathrm{NI}}$), SAs using an IL where $a \in \mathcal{A}_s^{c,I}$, and UCAs where $a \in \mathcal{A}^{uc}$. This decomposition helps in understanding how energy is utilized and distributed within the home.

$$E_t^{\mathrm{net}} = \sum_{a\in\mathcal{A}} E_{a,t} - \hat{E}_t^{\mathrm{PV}} \tag{2}$$

$$\sum_{a\in\mathcal{A}} E_{a,t} = \sum_{a\in\mathcal{A}_r^c} E_{a,t} + \sum_{a\in\mathcal{A}_s^{c,\mathrm{NI}}} E_{a,t} + \sum_{a\in\mathcal{A}_s^{c,I}} \left(E_{a,t}^{\mathrm{ch}} - E_{a,t}^{\mathrm{dch}}\right) + \sum_{a\in\mathcal{A}^{uc}} E_{a,t} \tag{3}$$

2.2.3 | Operation characteristics for CAs

For RAs $a \in \mathcal{A}_r^c$, Equation (4) outlines the constraints governing the temperature dynamics of these appliances, such as ACs, at time t (T_t^{in}). This constraint is defined based on the previous indoor temperature T_{t-1}^{in} at time $t-1$, the anticipated outdoor temperature at time $t-1$ ($\hat{T}_{t-1}^{\mathrm{out}}$), the energy consumption of the RAs ($E_{a,t}$), and various environmental variables denoted as (a, b), which collectively influence the internal thermal conditions.

Additionally, Equation (5) establishes the allowable range for the desired indoor temperature, bounded by $T^{\min}$ and $T^{\max}$, ensuring user comfort and satisfaction. The parameters $E_a^{\min}$ and $E_a^{\max}$ indicate the limits on the energy consumption capacity for the specific RA, which are further detailed in Equation (6):

$$T_t^{\mathrm{in}} = T_{t-1}^{\mathrm{in}} + \alpha\left(\hat{T}_{t-1}^{\mathrm{out}} - T_{t-1}^{\mathrm{in}}\right) + \beta E_{a,t} \tag{4}$$

$$T^{\min} \le T_t^{\mathrm{in}} \le T^{\max} \tag{5}$$

$$E_a^{\min} \le E_{a,t} \le E_a^{\max} \tag{6}$$

Equations (7) through (9) are designed to ensure the desired operation of SAs with a NIL $a \in \mathcal{A}_s^{c,\mathrm{NI}}$, which can include appliances like washing machines (WMs). These equations rely on a binary decision parameter, namely $b_{a,t}^{c,\mathrm{NI}}$, and address three key aspects:

(i) First, they consider the stopping time, where ω_s^{pref} and ω_f^{pref} represent the user's preferred start and finish times, as described in Equation (7).
(ii) Second, they account for the duration of daily operation, characterized by L_a periods, as stated in Equation (8).
(iii) Third, they address consecutive operations for L_a periods, as outlined in Equation (9). Furthermore, Equation (10) provides information regarding the EC capacity, denoted as $E_a^{\max}$, specifically for SAs that fall within this category:

$$b_{a,t}^{c,\mathrm{NI}} = 0,\ t \in \left[1, \omega_s^{\mathrm{pref}}\right) \cup \left(\omega_f^{\mathrm{pref}}, T\right] \tag{7}$$

$$\sum_{t=\omega_s^{\mathrm{pref}}}^{\omega_f^{\mathrm{pref}}} b_{a,t}^{c,\mathrm{NI}} = L_a \tag{8}$$

$$\sum_{t=p}^{p+L_a-1} b_{a,t}^{c,\mathrm{NI}} \ge \left(b_p^{c,\mathrm{NI}} - b_{p-1}^{c,\mathrm{NI}}\right) L_a,\quad \forall p \in \left(\omega_s^{\mathrm{pref}}, \omega_f^{\mathrm{pref}} - L_a + 1\right) \tag{9}$$

$$E_{a,t} = b_{a,t}^{c,\mathrm{NI}}\, E_a^{\max} \tag{10}$$

Equation (11) outlines the operational dynamics governing the state of energy (SOE) for ESSs ($a \in \mathcal{A}_s^{c,I}$) at a specific time, denoted as t. This equation takes into consideration various factors, including the SOE at $t-1$, the charging and discharging performance characteristics denoted as η_a^{ch} and η_a^{dch}, and the corresponding charging and discharging power levels, $E_{a,t}^{\mathrm{ch}}$ and $E_{a,t}^{\mathrm{dch}}$, respectively.

Furthermore, Equation (12) establishes the limitations on the capacity of the SOE within the ESS, which are defined by the parameters $SOE_a^{\min}$ and $SOE_a^{\max}$. Equations (13) and (14) introduce constraints pertaining to the charging ($E_{a,t}^{\mathrm{ch}}$) and discharging ($E_{a,t}^{\mathrm{dch}}$) power levels of the ESS. These constraints are influenced by the binary decision variable denoted as $b_{a,t}^{c,I}$, which determines whether the ESS is in an "off" or "on" state.

$$SOE_{a,t} = SOE_{a,t-1} + \eta_a^{\mathrm{ch}} E_{a,t}^{\mathrm{ch}} - \frac{E_{a,t}^{\mathrm{dch}}}{\eta_a^{\mathrm{dch}}} \tag{11}$$

$$SOE_a^{\min} \le SOE_{a,t} \le SOE_a^{\max} \tag{12}$$

$$E_a^{\mathrm{ch,min}}\, b_{a,t}^{c,I} \le E_{a,t}^{\mathrm{ch}} \le E_a^{\mathrm{ch,max}}\, b_{a,t}^{c,I} \tag{13}$$

$$E_a^{\mathrm{dch,min}}\left(1 - b_{a,t}^{c,I}\right) \le E_{a,t}^{\mathrm{dch}} \le E_a^{\mathrm{dch,max}}\left(1 - b_{a,t}^{c,I}\right) \tag{14}$$

Finally, to handle the non-linearity of the objective function $J_2(T_t^{\mathrm{in}})$, the problem is linearized, transforming the mixed-integer non-linear programming (MINLP) into a MILP optimization problem.

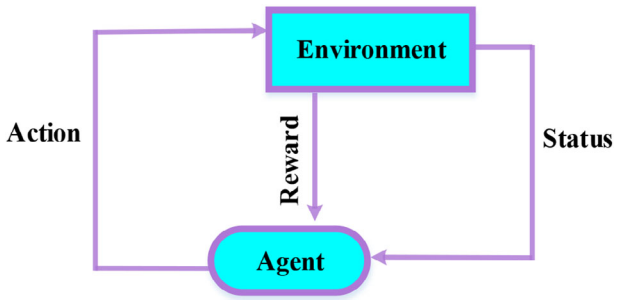

FIGURE 2 Conceptual model of reinforcement learning.

$$\Delta T_t = \left| T_t^{\text{in}} - T^{\text{set}} \right| \tag{15}$$

$$\Delta T_t \geq T_t^{\text{in}} - T^{\text{set}} \tag{16}$$

$$\Delta T_t \geq T^{\text{set}} - T_t^{\text{in}} \tag{17}$$

3 | RL- AND ANN-DRIVEN HEM FORMULA

3.1 | HEM through QL

RL is a pivotal technique in ML for optimal decision-making in unpredictable settings. As depicted in Figure 2, an agent learns to select actions based on the current state of the environment and then applies these actions to the environment.

Q-Learning (QL) is a prevalent RL method used to define the best policy v^* in decision-making scenarios. The fundamental approach of QL involves calculating Q-values $Q(s_t,\ a_t)$ for state-action pairs $(s_t,\ a_t)$ at discrete time intervals t, and refining these values toward the highest cumulative reward using the Bellman equation:

$$Q_{v^*}^*(s_t, a_t) = r(s_t, a_t) + \gamma \max Q\left(s_{t+1}, a_{t+1}\right) \tag{18}$$

Equation (18) determines the optimal Q-value $Q_{v^*}(s_t, a_t)$ by summing the immediate reward $r(s_t,\ a_t)$ and the best future discounted reward fl max $Q(s_{t+1},\ a_{t+1})$, where γ (ranging from 0 to 1) is the discount factor. A lower γ makes the agent prioritize immediate rewards, while a higher γ emphasizes future rewards. The QL algorithm allows the system operator to balance immediate and future rewards by adjusting γ.

When the Q-value $Q(s_t,\ a_t)$ is updated for a specific state-action pair at time t, it is added to the Q-value table (QVT). The agent selects actions based on the QVT at each time t, and the corresponding QVT entry is updated using Bellman's equation:

$$Q(s_t,\ a_t) \leftarrow (1-\theta) Q(s_t, a_t) + \theta r(s_t, a_t) + \gamma \max Q\left(s_{t+1}, a_{t+1}\right) \tag{19}$$

Here, $\theta \in [0,\ 1]$ represents the learning rate, influencing the extent to which the new Q-value supersedes the previous one. A balance between exploitation and exploration is achieved by setting θ within this range. Iterative updates of $Q(s_t,\ a_t)$ using Equation (19) lead to the optimal policy v^*, defined as:

$$v^* = argmax Q(s_t,\ a_t) \tag{20}$$

3.1.1 | State space

This paper considers the case of executing the suggested QL algorithm for one day with a planning resolution of 1 h. For $\forall t = 1, \ldots, 24$, state spaces for WM, AC, and ESS appear like this:

$$S^{\text{WM}} = \left\{E_t^{\text{WM}}\right\},\ S^{\text{AC}} = \left\{E_t^{\text{AC}}\right\},\ S^{\text{ESS}} = \left\{SOE_t^{\text{ESS}}\right\}, \tag{21}$$

in which E_t^{WM} shows the EC of the WM, E_t^{AC} shows the EC of the AC, and SOC_t^{ESS} represents SOE of the ESS for t.

3.1.2 | Action space

According to Section 3.1.1, the optimum action of an appliance is determined by its environment, such as its current status. WM, AC, and ESS action spaces appear in the following way:

$$A^{\text{WM}} = \{\text{On, Off}\} \tag{22}$$

$$A^{\text{AC}} = \left\{0, \Delta E^{\text{AC}}, 2\Delta E^{\text{AC}}, \ldots, 9\Delta E^{\text{AC}}\right\} \tag{23}$$

$$A^{\text{ESS}} = \left\{-4\Delta E^{\text{ESS}}, -3\Delta E^{\text{ESS}}, \ldots, 4\Delta E^{\text{ESS}}\right\} \tag{24}$$

[On, Off] is the binary action performed by the WM agent in Equation (22). When the WM agent is configured to use the "On" action, the WM uses fixed amounts of power ($E^{\text{WM,max}}$), but when WM agent is configured to use the "Off" action, the WM is turned off. Equation (23) discretizes the AC unit's action in 10 degrees of AC–EC, in which ΔE^{AC} indicates the AC's EC unit. In a similar way to the AC agent, Equation (24) defines the discrete group for actions of an ESS unit using the ESS ΔE^{ESS} energy unit. The discretized actions can be divided into discharge and charge actions, which are $\{-4\Delta E^{\text{ESS}}, \ldots, -1\Delta E^{\text{ESS}}\}$ and $\{1\Delta E^{\text{ESS}}, \ldots, 4\Delta E^{\text{ESS}}\}$, respectively. The suggested algorithm computes the devices' ECs per hour for the following 24 h. Based on the state and action sets described previously, the QVTs for the WM, AC, and ESS agents have been presented by applying the $|\mathcal{T}| \times |A^{\text{WM}}|$, $|\mathcal{T}| \times |A^{\text{AC}}|$, and $|\mathcal{T}| \times |A^{\text{ESS}}|$ matrices, having $|\mathcal{T}| = 24$, $|A^{\text{WM}}| = 2$, $|A^{\text{AC}}| = 10$, and $|A^{\text{ESS}}| = 9$, respectively. Here, $|A|$ represents the cardinality for the group A (or the total count of elements of A).

3.1.3 | Rewards

A device agent's reward function can be determined by adding the negative ElC and negative discomfort costs corresponding to the satisfaction and AO features the user prefers. In the case

of the HEMS, the overall reward r^{Total} would be:

$$r^{\text{Total}} = r_t^{\text{WM}} + r_t^{\text{AC}} + r_t^{\text{ESS}} \tag{25}$$

where, the three reward functions r_t^{WM}, r_t^{AC}, and r_t^{ESS} are aimed at evaluating the HEMS efficiency based on: (i) the ElC and user undesirable function for the WM, (ii) the ElC and user heating dissatisfaction for the AC, and (iii) the ElC and power inefficiency usage of the ESS because it is over-charged and under-charged. WM agent's reward function looks like this:

$$r_t^{\text{WM}} = \begin{cases} -\left[\pi_t E_t^{\text{WM}} + \bar{\delta}\left(\omega_s^{\text{pref}} - t\right)\right], & \text{if } t < \omega_s^{\text{pref}} \\ -\left[\pi_t E_t^{\text{WM}} + \underline{\delta}\left(t - \omega_f^{\text{pref}}\right)\right], & \text{if } t > \omega_f^{\text{pref}} \\ -\pi_t E_t^{\text{WM}} & \text{otherwise} \end{cases}, \tag{26}$$

in which, ω_s^{pref} shows the user's preferable start time and ω_f^{pref} shows the user's preferable finish time of the WM; however, $\bar{\delta}$ and $\underline{\delta}$ show the penalties for early and delayed function, respectively, in comparison with the user's preferable function period. A negative discomfort price will be added to the reward functions when a WM agent plans the WM-EC prior to ω_s^{pref} or following ω_f^{pref}; if else, just a negative ElC is included in the reward functions. This reward function of the AC unit would be:

$$r_t^{\text{AC}} = \begin{cases} -\left[\pi_t E_t^{\text{AC}} + k\left(T^{\min} - T_t^{\text{in}}\right)\right], & \text{if } T_t^{\text{in}} < T^{\min} \\ -\left[\pi_t E_t^{\text{AC}} + k\left(T_t^{\text{in}} - T^{\max}\right)\right], & \text{if } T_t^{\text{in}} > T^{\max} \\ -\pi_t E_t^{\text{AC}} & \text{otherwise} \end{cases}, \tag{27}$$

in which, k shows the penalty for the user's thermal dissatisfaction. As a result of the deviation of T_t^{in} from $T^{\min}$ and $T^{\max}$, the discomfort price has been calculated, and it can be regarded as the reward that is negative just when T_t^{in} has deviated from predetermined levels of $[T^{\min}, T^{\max}]$.

Last, the ESS agent's reward function includes a negative ElC and negative power insufficient usage price:

$$r_t^{\text{ESS}} = \begin{cases} -\left[\pi_t E_t^{\text{ESS}} + \bar{\tau}\left(SOE_t - SOE^{\max}\right)\right], & \text{if } SOE_t < SOE^{\max} \\ -\left[\pi_t E_t^{\text{ESS}} + \underline{\tau}\left(SOE^{\min} - SOE_t\right)\right], & \text{if } SOE_t > SOE^{\min} \\ -\pi_t E_t^{\text{ESS}} & \text{otherwise} \end{cases}, \tag{28}$$

in which, $\bar{\tau}$ and $\underline{\tau}$ show the penalties for the ESS overcharge and undercharge, respectively. An ESS is considered underutilized when the SOE falls below $SOE^{\min}$ (undercharge) or exceeds $SOE^{\max}$ (overcharge), and it would be a reward term, alongside the ElC when the ESS is underutilized.

ALGORITHM 1 QL-based EM of SH using ESS, PV system, and HAs.

1. Initializing every power requirement of the appliance, discomfort variables, and QL variables
2. %% Learning using ANN to predict the AC's temperature
3. Indoor temperatures for epoch $t-1 \rightarrow T_{t-1}^{\text{in}}$
4. Maximum and minimum amount of consumer's comfort temperature domain $\rightarrow T^{\min}$, $T^{\max}$
5. Predicted outdoor temperature at epoch $t \rightarrow \hat{T}_t^{\text{out}}$
6. EC of AC unit at epoch $t \rightarrow E_t^{\text{AC}}$
7. Forecasted indoor temperature at epoch $t \rightarrow T_t^{\text{in}}$
8. Learn procedure using ANN and forecast the temperature predicting layout $\hat{f}$
9. $T_t^{\text{in}} = \hat{f}\left(T_t^{\text{in}}, T^{\max}, T^{\min}, \hat{T}_t^{\text{out}}, E_t^{\text{AC}}\right)$
10. Initialize QV of every unit
11. **for** episode = 1, MaxEpisode **do**
12. Initializing action, time, and status epoch
13. **for** time stage = 1:24 **do**
14. Choose a_t from existing s_t applying e-greedy policy
15. Take action a_t; observe $r(s_t, a_t)$ and s_{t+1}
16. $Q(s_t, a_t) \leftarrow (1-\theta)Q(s_t, a_t) + \theta[r(s_t, a_t) + \gamma \max Q(s_{t+1}, a_{t+1})$
17. **end**
18. **end**
19. Determine optimum policies using greatest QV

3.2 | Forecasting indoor temperature through ANN

According to the suggested ANN scheme, the AC agent learns the relation between the AC–EC and present indoor temperature by estimating the function f, which represents the relationship in the following way:

$$\hat{f} = \left(T_{t-1}^{\text{in}}, T^{\min}, T^{\max}, \hat{T}_t^{\text{out}}, E_t^{\text{AC}}\right) = T_t^{\text{in}}, \tag{29}$$

in which, $\hat{f}$ shows the approximated function explaining the relation among the input information from the empirical transfer pricing (ETP) scheme in Section 3.1.2, including the prior indoor temperature (T_{t-1}^{in}), user's desired indoor thermal levels ($T^{\min}$, $T^{\max}$), weather predicting ($\hat{T}_t^{\text{out}}$), and AC-EC (E_t^{AC}) and the output for the forecasted present indoor temperature.

According to Figure 3, in the suggested ANN scheme, there are 5 neurons in the input data layer, 17 neurons in the latent layers, and a neuron in the output layer. Layers calculate the weighted total of the input vector and the fixed bias b_i, having one weight W_i, and transfer the weighted total to the next layer using the transfer functions. The paper uses the rectified linear unit (ReLu) function to be one transfer function [21]. Additionally, the developed ANN scheme is trained using the Adam optimization algorithm [22], with a learning rate of 0.005.

Based on Algorithm 1, EM policies are learned by HEMSs using PV, ESS, and HAs, to optimize energy bills and user satisfaction levels. The HEMS can receive the hour-ahead indoor temperature, user-desired range of indoor temperatures,

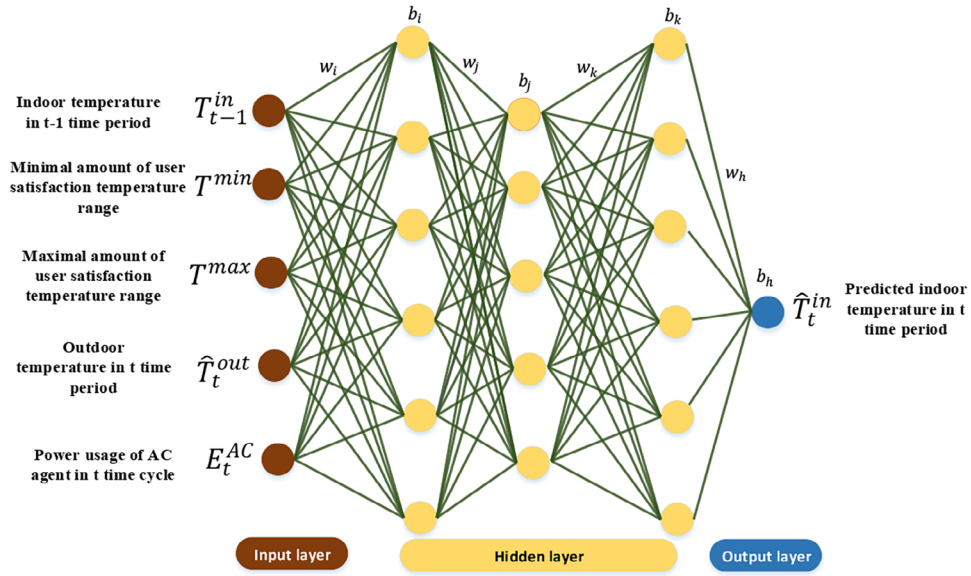

FIGURE 3 Structure for this suggested artificial neural network (ANN) scheme.

forecast external temperature, and AC-EC (E_t^{AC}), and apply the ANN to predict the present inner temperature.

4 | INSTANCES OF NUMERICAL DATA

4.1 | Setting up the simulation

Figure 4a shows a residential scenario equipped with two primary HAs—an AC and a WM, alongside an ESS, all managed by a HEMS operating on a TOU tariff scheme. The simulation adopted an hourly planning interval. The predicted PV energy generation $\hat{E}_t^{\mathrm{PV}}$ shown in Figure 4a and the anticipated external temperature $\hat{T}_{t-1}^{\mathrm{out}}$ in Figure 4b were presumed precise. The AC units were capped at an EC limit of 3000 Wh, while the WM and the combined uninterruptible critical appliance array had maximum ECs of 500 and 1700 Wh, respectively. The temperature range satisfying consumer comfort was set between 23°C and 25°C, with a targeted set temperature T^{set} of 24°C. Discomfort penalties in the MILP-based and DL-enhanced constraint coordination game HEMS were both set at 100. AC thermal characteristics were defined with ff = 0.8 and fi = −0.02. The operational window for the WM was scheduled from 6 AM to 10 PM with a maximum continuous operation duration of 2 h. The ESS's maximum charging and discharging capacities were both set at 4000 Wh, with an initial SOE of 2400 Wh, and minimum and maximum SOE thresholds of 800 and 4000 Wh, respectively. In the action space for the AC and ESS, the EC increments ΔE^{AC} and ΔE^{ESS} were 40 and 150 Wh, respectively. In the reward functions, the discomfort cost penalties for the ESS and WM were set at $\bar{\delta} = 50$, $\underline{\delta} = 50$, $\bar{\tau} = 50$, and $\underline{\tau} = 50$. The ϵ-greedy policy was configured with ϵ to balance exploitation and exploration. The Bellman equation was modified with a learning rate $\theta = 0.1$ and a discount factor $\gamma = 0.9$. MATLAB R2019a was utilized for the algorithm's implementation.

4.2 | Efficiency for the suggested RL-driven HEMS

Here, the proposed CCG integrated with a DL-based HEMS algorithm is modelled, examining the ECs for CAs and the charging/discharging strategy of ESS. The ECs, as influenced by WM and AC agents, is illustrated in Figure 5a. Referring to Figure 5a, and targeting specific operation intervals (6 AM, 10 PM) over two sequential cycles (L = 2), the optimal schedule for WM is identified as (7 AM, 8 AM).

This scheduling approach ensures operation during periods of lowest TOU rates, thereby reducing energy expenses without affecting user preferences. Figure 5c,d displays the ESS's charging/discharging and SOE patterns. Similar to Figure 5a, Figure 5c indicates that ESS charging predominantly occurs during periods of lower TOU rates and discharging happens when TOU rates are higher, facilitating cost reductions for end-users. Furthermore, Figure 5d demonstrates a correlation between SOE levels and cost variations, with SOE decreasing as costs rise and vice versa.

4.3 | Effects of various variables in reward function on the suggested scheme

This section delves into examining how varying penalty values k and preferred start and end times $[\omega_s^{\mathrm{pref}}, \omega_f^{\mathrm{pref}}]$ influence the efficiency of WMs and ACs through their respective reward functions. The impact of different k values (10, 50, 100) on the indoor temperature T_t^{in} at time t, given an external temperature T_t^{out}, is depicted in Figure 6a,b. Figure 6a demonstrates that with $k = 10$, $k = 50$, and $k = 100$. For $k = 10$, there is a notable deviation in indoor temperature from the preferred range (23°C to 25°C). However, it reveals that increasing k leads to a reduced deviation from this desired temperature range. This

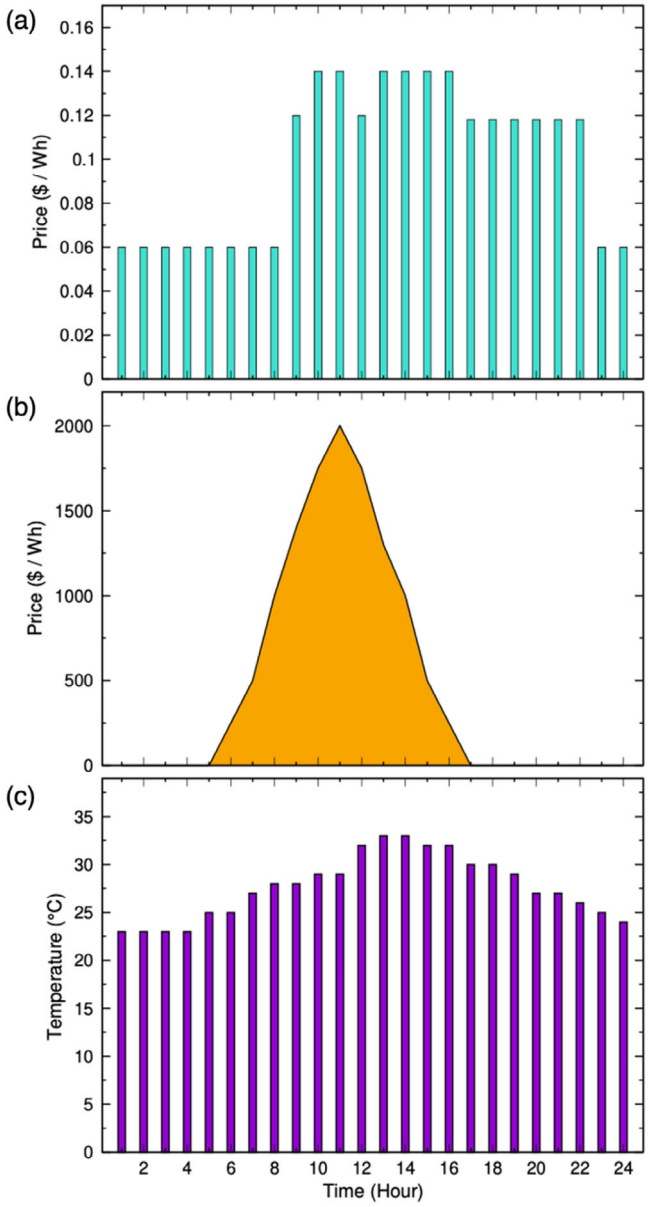

FIGURE 4 The profile of energy cost and weather. (a) Time-of-use (TOU) cost; (b) photovoltaic (PV) production; (c) outdoor temperature.

is attributed to the AC's strategy of prioritizing comfort, which may lead to higher energy costs. A comparison of Figure 6a with Figure 6b highlights the trade-off between energy conservation and occupant comfort. As per Figure 6b, the energy consumption of the AC (AC-EC) escalates with an increase in k to meet user comfort needs.

The influence of different preferred operating times on the energy consumption of the washing machine (WM-EC) is illustrated in Figure 7. This figure consistently shows results across three distinct operating windows: (6:00, 22:00), (12:00, 22:00), and (17:00, 22:00), maintaining a constant end time of 10 PM (ω_f^{pref}) but varying start times (ω_f^{pref} = 6:00, 24:00, and 17:00). Figure 7 suggests that the WM's optimal operational schedule is selected within the preferred timeframe when Time-of-Use (TOU) costs are minimal.

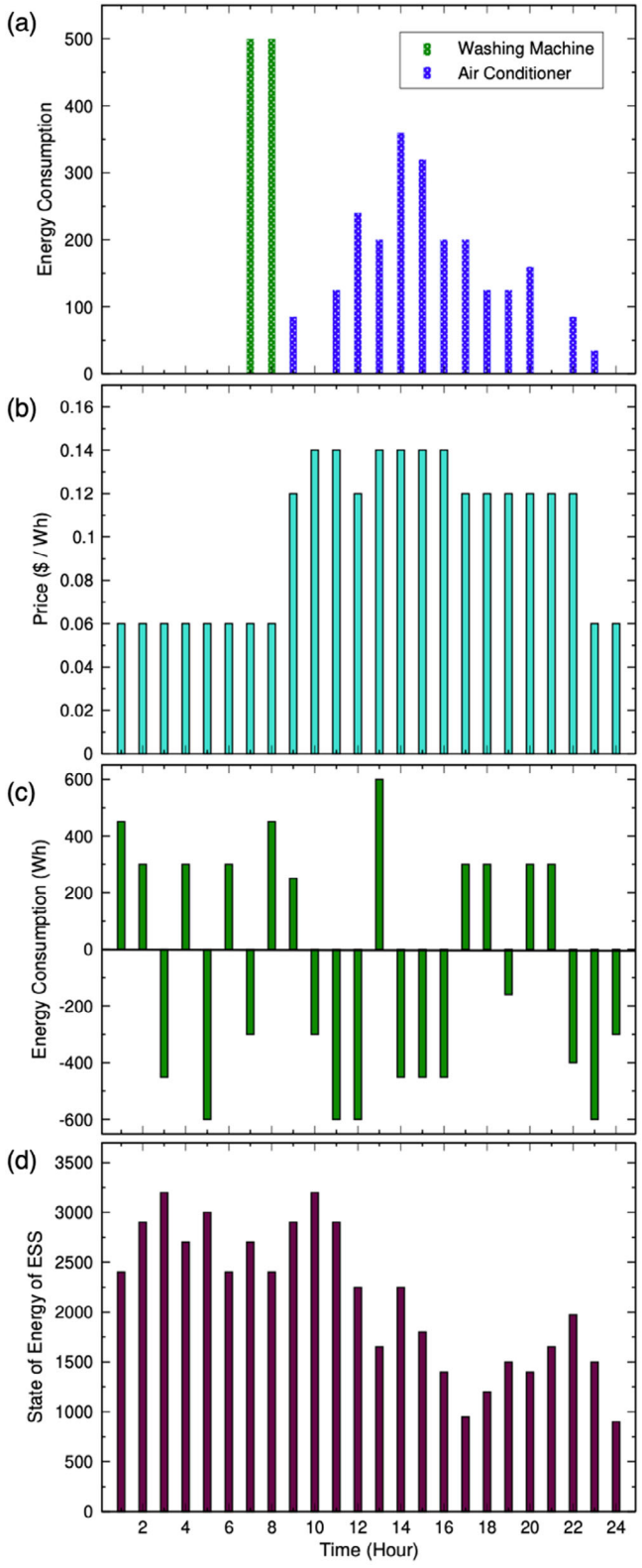

FIGURE 5 Reinforcement learning (RL)-driven day-ahead operation plan of device using time-of-use (TOU) cost tariff. (a) Energy consumption (EC) of air conditioner (AC) and washing machine (WM); (b) TOU cost; (c) charge and discharge of energy storage system (ESS); (d) state of energy (SOE) for ESS.

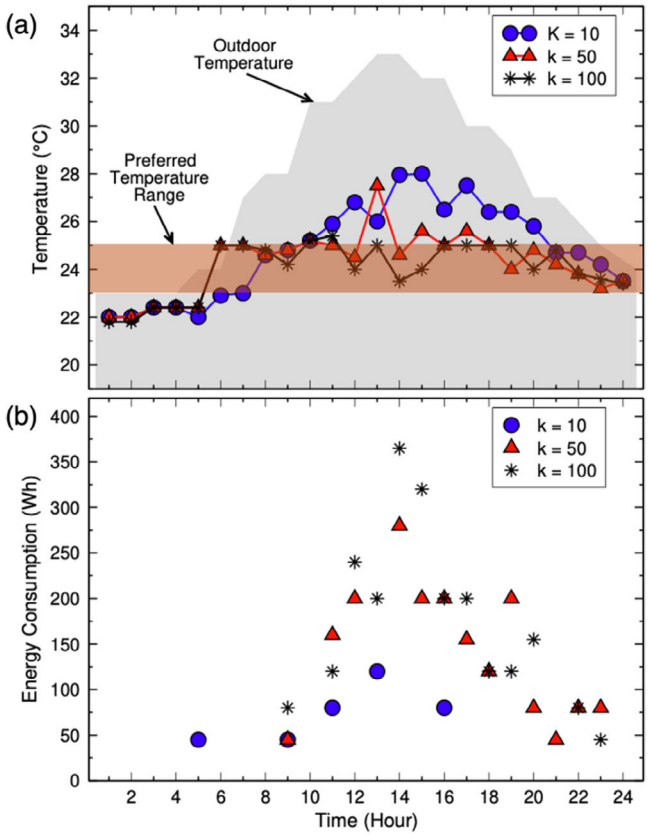

FIGURE 6 Impact of different penalty values (k) on air conditioning (AC) strategy: Analyzing indoor temperature (T_t^{in}) and energy consumption ($E_{a,t}$) (a) Indoor temperature with $k = 10$, $k = 50$, and $k = 100$; (b) energy consumption with $k = 10$, $k = 50$, and $k = 100$.

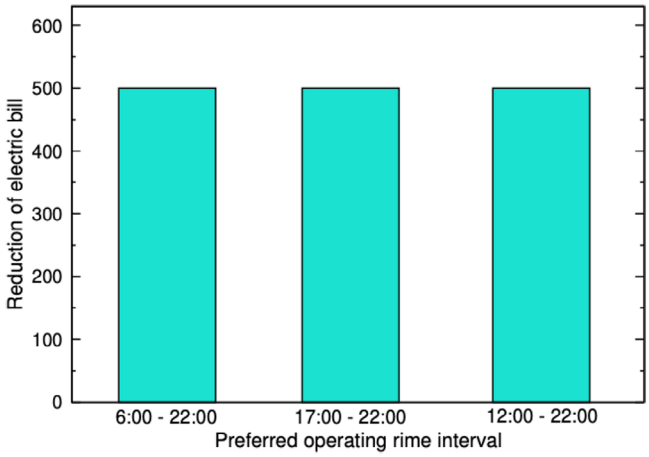

FIGURE 7 Impact of various desired operation periods $[\omega_s^{pref}, \omega_f^{pref}]$ for washing machine (WM) planning on energy consumption (EC, $E_{a,t}$). (a) (6:00, 22:00); (b) (24:00, 22:00); (c) (17:00, 22:00).

4.4 | Effect of ANN on AC agent efficiency

This section aims to evaluate the performance of the proposed RL-based approach in forecasting indoor temperatures using the ANN method outlined in Section 3.2. Figure 8a,b illustrates

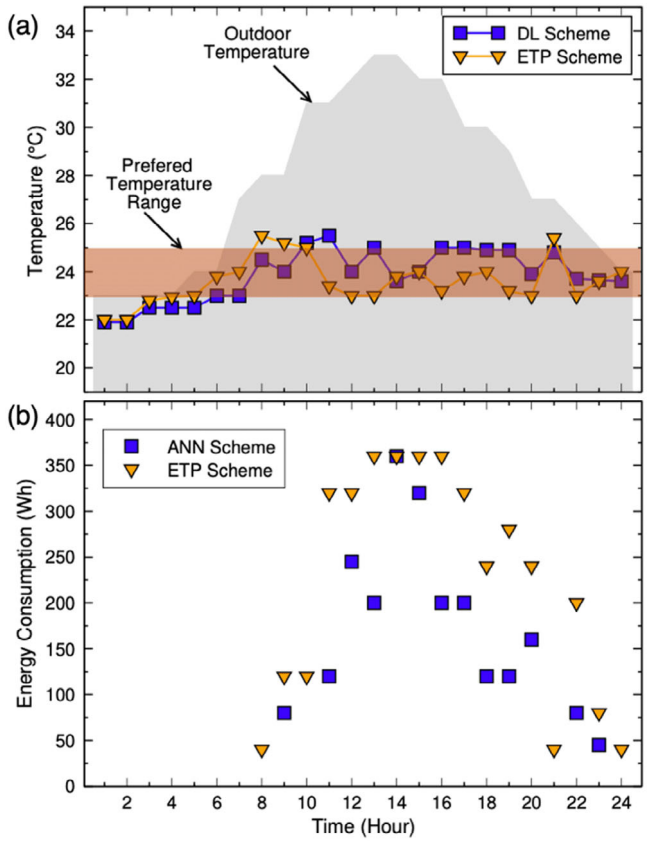

FIGURE 8 Utilizing reinforcement learning for advanced energy control systems in air conditioning: Forecasting indoor temperatures to optimize (a) ambient temperature and (b) power usage. ANN, artificial neural network; DL, deep learning; ETP, empirical transfer pricing.

a comparison of air conditioning energy consumption (AC-EC) and indoor temperatures between the ETP and ANN models, employing the QL technique. Upon analysis, the ANN model demonstrates a reduced power requirement in comparison to the ETP model, as depicted in Figure 8a.

4.5 | Comparing MILP and RL-driven HEMS in terms of efficiency

This section delineates a comparative analysis of the proposed DL-enhanced CCG approach against a traditional MILP-based HEMS strategy. The operational dynamics of AC-EC under MILP and DL-augmented CCG frameworks are depicted in Figure 9a,b. Additionally, Figure 9c,d juxtaposes the ESS charging and discharging strategies under both methodologies. Notably, the ESS's energy management in Figure 9b under the CCG with DL paradigm demonstrates more effective energy utilization, translating into a cost reduction of $94, with the ElC for MILP and CCG with DL being −$194 and −$288, respectively. As per Figure 9e,f, the DL-based CCG approach significantly lowers AC-EC compared to the MILP approach, with ElC reductions of $228, against $463 and $234 for MILP and DL-based CCG, respectively.

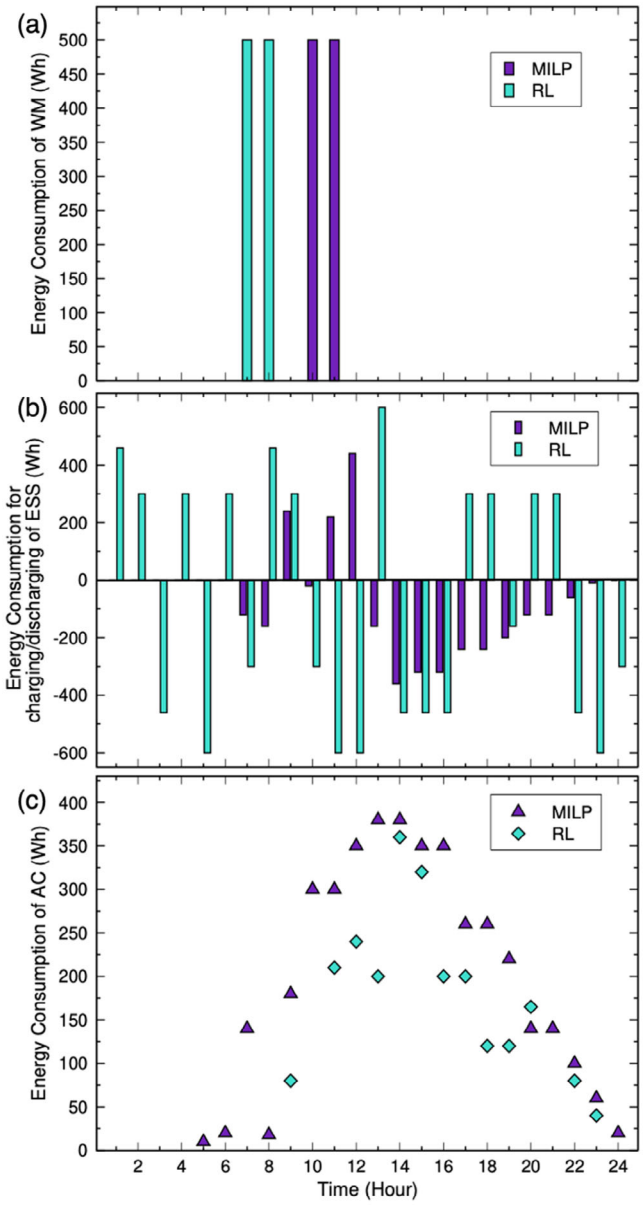

FIGURE 9 Illustrating the contrast in energy consumption and power fluctuations during charging/discharging: A study of mixed integer linear programming (MILP) versus reinforcement learning (RL) methods. (a) Power management in washing machines; (b) energy storage systems; (c) air conditioning power usage.

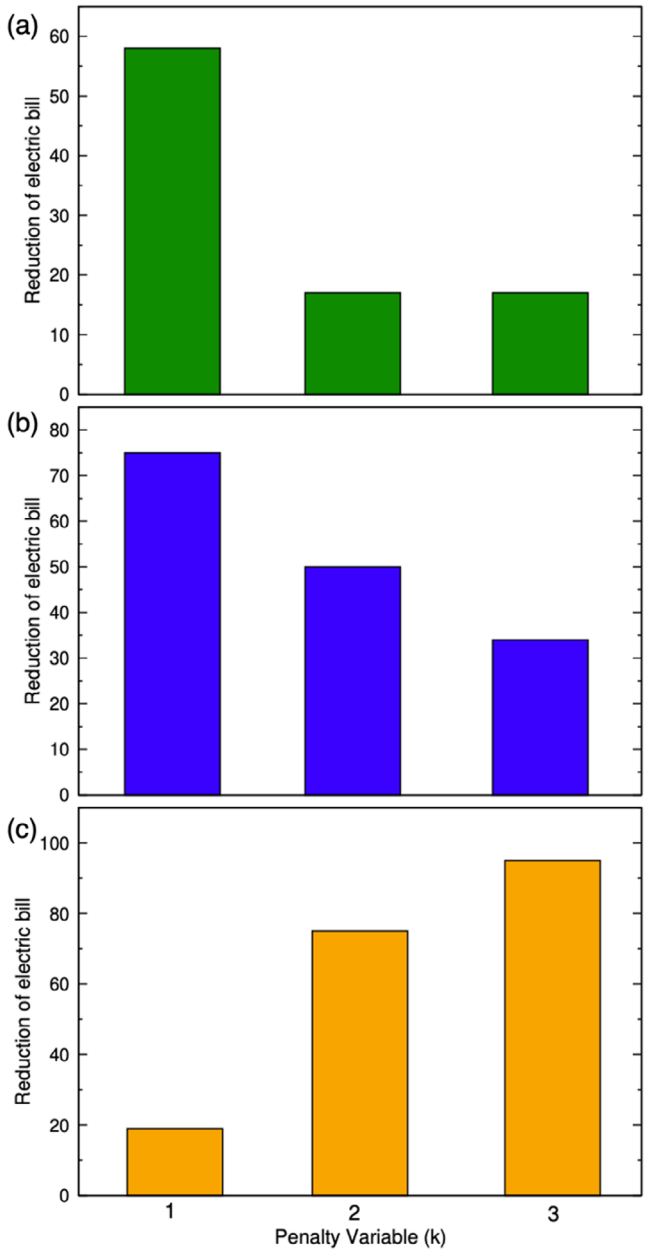

FIGURE 10 Reduction in energy expenditures across diverse operational scenarios: (a) Variation in washing machine (WM) during assorted targeted operational durations; (b) fluctuations in energy storage system (ESS) with varying storage capacities; (c) changes in air conditioner (AC) subject to distinct penalty rates.

In Figure 10a–c, alterations in parameters within the DL-enhanced CCG framework are shown to yield a reduction in total energy expenditure:

$$\frac{X^{\text{bill,MILP}} - X_p^{\text{bill,MILP}}}{X^{\text{bill,MILP}}} \times 100\,(\%) \tag{30}$$

Here, $X^{\text{bill,MILP}}$ represents the total energy cost in the CCG framework utilizing MILP, while $X_p^{\text{bill,MILP}}$ indicates the total energy cost under the CCG framework with DL, where p denotes a variable such as the desired operation duration for the WM, ESS capacities, and penalties related to user-preferred indoor heating conditions in the AC system.

Figure 10a illustrates that WM agents, by opting for extended operation periods, can effectively reduce energy costs. As depicted in Figure 10b, an increase in ESS capacity leads to enhanced energy charging and discharging, thereby lowering energy expenses. Furthermore, Figure 10c reveals that reduced penalties for preferred indoor thermal conditions lead to additional energy savings, as the AC agent optimizes the WM's energy consumption without compromising user comfort. Finally, Figure 11 presents an hourly comparison of overall EC between the MILP and the proposed DL-based CCG strategies.

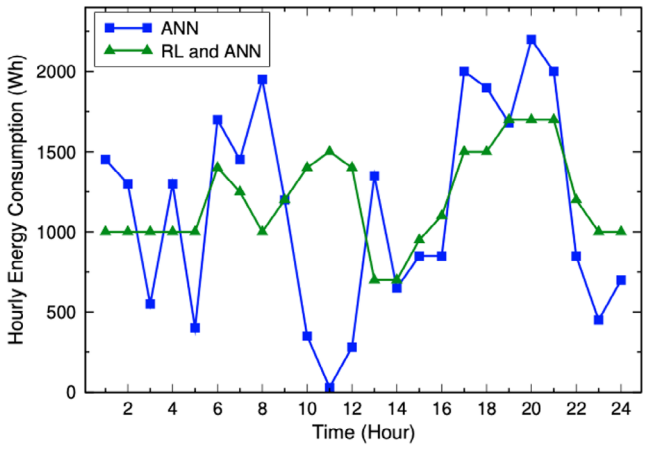

FIGURE 11 Comparing every hour energy consumptions (ECs) between mixed integer linear programming (MILP) approach and reinforcement learning (RL) schemes. ANN, artificial neural network.

5 | DISCUSSION

The simulation results and analysis presented here highlight the efficacy and versatility of the QL-based energy management approach in smart home environments. Through the optimization of appliance operation, integration of renewable energy sources, and adaptation to dynamic pricing schemes, the algorithm offers a comprehensive solution for sustainable and cost-effective energy management. Moving forward, further research and development efforts will focus on refining the algorithm, expanding its applicability, and addressing emerging challenges in smart home automation and energy sustainability. The simulation results presented in Table 1, along with Figures 4–11, offer valuable insights into the performance and efficiency of the proposed QL-based energy management approach in smart home environments. This discussion section aims to delve deeper into the implications of these results, analyzing key findings and elucidating their significance. Table 1 provides a comprehensive overview of the operational dynamics of key home appliances, including ACs, WMs, and ESSs. The data reveal that the QL algorithm effectively regulates the energy consumption of these appliances, ensuring optimal operation while minimizing costs. For instance, the scheduled operation of the WM during off-peak hours, as depicted in Figure 7, aligns with TOU tariff schemes, resulting in substantial cost savings for homeowners. Similarly, Figure 5a,b demonstrates the efficient management of AC energy consumption based on real-time pricing, highlighting the algorithm's ability to adapt appliance operation to varying electricity costs.

The influence of reward function parameters on appliance efficiency is evident from Figures 6 and 8. By varying penalty values (k) and preferred operating times, the algorithm balances energy conservation with user comfort preferences. For instance, increasing k values lead to reduced deviations in indoor temperature from the desired range, as shown in Figure 6b,c. However, this comes at the expense of higher energy consumption, as depicted in Figure 6e,f. This trade-off underscores the importance of parameter tuning in optimizing appliance performance and energy utilization.

Figures 9 and 10 provide a comparative analysis of the QL-based EM approach against traditional methods, such as MILP. The data illustrate the superiority of the proposed algorithm in terms of energy cost reduction and appliance efficiency. For example, Figure 9b showcases the more effective energy utilization of the ESS under the QL paradigm, leading to significant cost reductions compared to MILP-based strategies. Similarly, Figure 10 highlights the impact of parameter adjustments within the QL framework on overall energy expenditures, emphasizing the algorithm's adaptability and efficiency.

The integration of the PV system into the EM framework further enhances energy sustainability and cost-effectiveness. Figure 4a illustrates the variability of TOU tariff schemes, which the algorithm leverages to optimize energy consumption and maximize the utilization of solar energy. By forecasting PV energy generation and adapting appliance operation accordingly, the algorithm minimizes the reliance on grid power and reduces overall electricity expenses, as demonstrated in Table 1. The simulation results underscore the potential of the QL-based EM approach to promote sustainability and affordability in smart home environments. By intelligently managing appliance operation, optimizing energy utilization, and leveraging renewable energy sources, the algorithm contributes to reduced carbon emissions and enhanced energy efficiency. Moreover, the cost savings realized through the implementation of the proposed approach have significant implications for homeowners, offering tangible economic benefits and improving overall quality of life.

6 | CONCLUSIONS

Here, we introduce a novel approach for smart HEM leveraging ML, integrating DL within a combined cycle generation framework. Our approach focuses on optimizing energy expenditure for two types of manageable HAs alongside the efficient operation of ESSs. The primary objective is to reduce energy costs while ensuring user comfort and adhering to AO standards. Within our proposed QL model, devices such as WMs, ACs, and ESSs autonomously adapt their operational strategies through environmental interaction, aiming to enhance their cumulative rewards. Specifically, the WM's energy consumption is aligned with user-preferred timings. ESS units are adept at modulating charging and discharging rates to prevent overcharging or depletion. Additionally, the AC unit, utilizing DL for indoor temperature forecasting, adjusts its energy usage to achieve the user's ideal indoor climate. Comparative simulations reveal that our proposed ML-based strategy surpasses traditional MILP methods in terms of cost-efficiency. In our research, the QL algorithm stands out as the more efficient method for the proposed work of energy management in smart homes. Several key factors contribute to its effectiveness:

- **Adaptability and learning capability**: QL is a reinforcement learning algorithm known for its adaptability and ability to learn optimal strategies through trial and error. It continuously updates its action-selection policy based on past experiences, making it well-suited for dynamic

environments with changing conditions, such as smart home energy management.

- **Real-time decision-making**: The QL algorithm enables real-time decision making by dynamically adjusting appliance operation schedules and energy storage strategies based on current environmental factors and electricity pricing. This capability ensures that energy usage is optimized in response to fluctuating conditions, leading to cost savings and improved efficiency.
- **Integration with deep learning**: Our research also incorporates deep learning techniques, such as ANNs, to enhance the QL algorithm's performance. By utilizing ANNs for tasks such as indoor temperature forecasting, the algorithm gains improved predictive capabilities, enabling more accurate and proactive energy management decisions.
- **Cost reduction and sustainability**: Simulation results demonstrate that the QL-based energy management approach leads to significant reductions in energy costs compared to traditional methods, such as MILP. By optimizing appliance operation schedules, leveraging renewable energy sources like solar power, and adapting to time-of-use pricing schemes, the algorithm promotes both cost reduction and sustainability in smart home environments.

AUTHOR CONTRIBUTIONS

Sajjad Afroosheh: Conceptualization; data curation; formal analysis; investigation; methodology; project administration. **Khodakhast Esapour**: Data curation; visualization; writing—original draft; writing—review and editing. **Reza Khorram-Nia**: Validation; visualization; writing—original draft; writing—review and editing. **Mazaher Karimi**: Resources; software; supervision; validation; visualization.

ACKNOWLEDGEMENTS

The authors have nothing to report.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data would be provided as per request.

ORCID

Khodakhast Esapour https://orcid.org/0000-0003-4746-5684
Mazaher Karimi https://orcid.org/0000-0003-2145-4936

REFERENCES

1. Su, S., et al.: Bi-level energy management and pricing for community energy retailer incorporating smart buildings based on chance-constrained programming. Int. J. Electr. Power Energy Syst. 138, 107894 (2022)
2. Qiu, D., et al.: Strategic retail pricing and demand bidding of retailers in electricity market: a data-driven chance-constrained programming. Adv. Appl. Energy 7, 100100 (2022)
3. Yang, S., et al.: Chance-constrained bi-level optimal dispatching model and benefit allocation strategy for off-grid microgrid considering bilateral uncertainty of supply and demand. J. Electr. Power Energy Syst. 146, 108719 (2023)
4. Hai, T., et al.: Optimal energy management strategy for a renewable based microgrid with electric vehicles and demand response program. Electr. Power Syst. Res. 221, 109370 (2023)
5. Zhai, J., et al.: Data-driven distributionally robust joint chance-constrained energy management for multi-energy microgrid. Appl. Energy 326, 119939 (2022)
6. Liu, H., et al.: An effective energy management Layout-Based reinforcement learning for household demand response in digital twin simulation. Sol. Energy 258, 95–105 (2023)
7. Saber, H., et al.: Transactive charging management of electric vehicles in office buildings: a distributionally robust chance-constrained approach. Sustain. Cities Soc. 87, 104171 (2022)
8. Nagpal, H., et al.: Optimal energy management in smart sustainable buildings – a chance-constrained model predictive control approach. Energy Build. 248, 111163 (2021)
9. Aghdam, F.H., et al.: A chance-constrained energy management in multi-microgrid systems considering degradation cost of energy storage elements. J. Energy Storage 29, 101416 (2020)
10. Wang, X., et al.: Chance-constrained model predictive control-based operation management of more-electric aircraft using energy storage systems under uncertainty. J. Energy Storage 55(Part C), 105629 (2022)
11. Tostado-Véliz, M., et al.: A fully robust home energy management model considering real time price and on-board vehicle batteries. J. Energy Storage 72(Part B), 108531 (2023)
12. Huy, T.H.B., et al.: Real-time energy scheduling for home energy management systems with an energy storage system and electric vehicle based on a supervised-learning-based strategy. Energy Convers. Manage. 292, 117340 (2023)
13. Zhou, B., et al.: Robust chance-constrained programming approach for the planning of fast-charging stations in electrified transportation networks. Appl. Energy 262, 114480 (2020)
14. Tostado-Véliz, M., et al.: Robust energy management in isolated microgrids with hydrogen storage and demand response. Appl. Energy 345, 121319 (2023)
15. Tong, Z., et al.: The role of smart communities integrated with renewable energy resources, smart homes and electric vehicles in providing ancillary services: a tri-stage optimization mechanism. Appl. Energy 351, 121897 (2023)
16. Umair, M., et al.: Energy management of smart homes over fog-based IoT architecture. Sustain. Comput. Inform. Syst. 39, 100898 (2023)
17. Aghdam, F.H., et al.: A stochastic optimal scheduling of multi-microgrid systems considering emissions: a chance constrained model. J. Cleaner Prod. 275, 122965 (2020)
18. Yi, Y., Verbič, G.: Fair operating envelopes under uncertainty using chance constrained optimal power flow. Electr. Power Syst. Res. 213, 108465 (2022)
19. Natarajan, A., et al.: Occupancy detection and localization strategies for demand modulated appliance control in Internet of Things enabled home energy management system. Renew. Sustain. Energy Rev. 167, 112731 (2022)
20. Wu, J., et al.: Maximum demand flexibility from the demand response of a big group of residential homes. J. Electr. Power Energy Syst. 147, 108800 (2023)
21. Tostado-Véliz, M., et al.: A novel hybrid lexicographic-IGDT methodology for robust multi-objective solution of home energy management systems. Energy 253, 124146 (2022)
22. Chu, Y., et al.: Optimal home energy management strategy: a reinforcement learning method with actor-critic using Kronecker-factored trust region. Electr. Power Syst. Res. 212, 108617 (2022)

How to cite this article: Afroosheh, S., Esapour, K., Khorram-Nia, R., Karimi, M.: Reinforcement learning layout-based optimal energy management in smart home: AI-based approach. IET Gener. Transm. Distrib. 18, 2509–2520 (2024).
https://doi.org/10.1049/gtd2.13203