

The Development of an Intelligent Home Energy Management System Integrated with a Vehicle-to-Home Unit using a Reinforcement Learning Approach

Ohoud Almughram^{1,*}, Sami Ben Slama² and Bassam Zafar²

omalmughram@stu.kau.edu.sa, sabdullah1@kau.edu.sa, bzafar@kau.edu.sa

Department of Information Systems, College of Computer Science, King Khalid University, Abha, Saudi Arabia;
Analysis and Processing of Electrical and Energy Systems Unit, Faculty of Sciences of Tunis El Manar, Tunis 2092,
Tunisia;

Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia;

Abstract:

Vehicle-to-Home (V2H) and Home Centralized Photovoltaic (HCPV) systems can address various energy storage issues and enhance demand response programs. Renewable energy, such as solar energy and wind turbines, address the energy gap. However, no energy management system is currently available to regulate the uncertainty of renewable energy sources, electric vehicles, and appliance consumption within a smart microgrid. Therefore, this study investigated the impact of solar photovoltaic (PV) panels, electric vehicles, and Micro-Grid (MG) storage on maximum solar radiation hours. Several Deep Learning (DL) algorithms were applied to account for the uncertainty. Moreover, a Reinforcement Learning HCPV (RL-HCPV) algorithm was created for efficient real-time energy scheduling decisions. The proposed algorithm managed the energy demand between PV solar energy generation and vehicle energy storage. RL-HCPV was modeled according to several constraints to meet household electricity demands in sunny and cloudy weather. Simulations demonstrated how the proposed RL-HCPV system could efficiently handle the demand response and how V2H can help to smooth the appliance load profile and reduce power consumption costs with sustainable power generation. The results demonstrated the advantages of utilizing RL and V2H as potential storage technology for smart buildings.

Keywords:

Deep Learning, Home Energy Management System, Q-learning, Reinforcement Learning, 18 Solar Photovoltaic, Vehicle-to-Home

1. Introduction

Conventional energy sources, including petroleum, natural gas, and coal, contribute to environmental pollution. The effect of these fossil fuels on the environment makes it essential to use Renewable Energy Sources (RnEs) such as solar and wind. However, the fluctuating nature of RnEs introduces new challenges to the energy exchange and load balancing of Smart Grids (SGs). One of the most effective ways to achieve power systems stability is to use the Electric Vehicle (EV) battery as distributed energy storage.

Previous studies have shown that 40% of energy is consumed in buildings, including homes, offices, and

hospitals [1]. Nowadays, smart buildings are equipped with RnEs as a power supply and controlled via an Energy Management System (EMS). Therefore, SBs are ideal for energy support to evaluate the impact of the EMS. Furthermore, though EV increases the SG loading, it can function as potential storage to facilitate the demand response (DR) [2]. Therefore, incorporating EVs into SBs is a technique for optimizing cost reduction [3, 4] and enhancing the usage of energy sources. Thus, the novel Vehicle-to-Home (V2H) technology is appealing due to its ability to reduce electricity prices in the home. In addition, when the SG is overloaded, V2H technology transfers the peak load and supplies constant emergency lighting [5].

According to research conducted by Pearre et al., the development of EV technology by focusing on discharging modes, namely, Vehicle to Grid (V2G), Vehicle to Building (V2B), and V2H, has been evaluated [6]. In the discharging modes, EVs discharge their batteries and distribute the excess energy to the power grid, buildings, or residences. The authors argued that the V2G mode is less efficient than V2B and V2H due to the significant transmission power losses. On the other hand, the V2H discharging mode can utilize the excess energy from the EV battery more efficiently. For example, in V2H, residences consume the excess energy from the EV battery through appliances. Therefore, the Home Energy Management System (HEMS) manages smart appliances and EVs. In addition, a smart meter is connected to the HEMS to support bi-directional information exchange between the customers and the utility service provider.

Moreover, the HEMS can rely on a prediction algorithm to determine the charging and discharging periods. The HEMS charges the EV battery for cost reduction when the price is low. In contrast, the HEMS schedules the discharge of the EV battery to power home appliances and avoids purchasing energy from the utility service provider when the price is high. Furthermore, this strategy serves as an emergency energy source during a power outage. Other studies have indicated that EVs and V2H technologies were expected to be used as domestic

electricity storage systems [7]. Thus, it is crucial to promote PV power self-consumption. In addition, the authors identified V2H as a technology that uses an EV battery as a controlled voltage source to supply home loads during power outages. As a result, EVs provide an uninterrupted power supply, increasing grid reliability. The study compared V2H and V2G, demonstrating that V2G has a complex infrastructure, which increases losses due to location distance. Therefore, EV application in V2H is more promising in terms of its controlling strategy. V2H technology has gone through several optimization procedures [8]. Conventional optimization methods, including model-based linear and dynamic programming, have been used in different applications for integrating V2H into HEMSs, such as scheduling EV charging and discharging [9], economic load reduction [10], and load shifting [11]. However, these optimization methods cannot adapt to the environment's stochastic nature, including unpredictable driving patterns, load profiles, and pricing. The global search technique, which uses genetic algorithms [12] and swarm intelligence [13] to resolve power management issues, is another model-based optimization strategy. However, the application of these methods in real-time systems is limited due to extensive computations, which makes them slow. Thus, data-driven approaches, including machine learning (ML), are being introduced to offer more advanced services [14–16]. Furthermore, reinforcement learning (RL) algorithms provide superior alternative energy decisions, control, and management solutions due to their outstanding ability to make decisions and resolve issues without prior knowledge of the environment. Additionally, before being activated online for any driving patterns, load profiles, or dynamic electricity pricing, RL models may be tested offline for a generation profile and general load [8, 17]. Compared to existing optimization methods, this study proposes a Home Centralized Photovoltaic (HCPV) technique that uses Deep Learning (DL) and a model-free RL algorithm, the Fuzzy Q-learning algorithm. The following are the main contributions of this paper:

- The proposed RL-HCPV model optimizes the energy usage of a smart home equipped with PV solar panels, stationary batteries, EVs, smart appliances, and V2H units.
- Various DL algorithms predict PV power generation, appliance consumption, and prices one hour in advance. Then, an RL algorithm is employed to make efficient power decisions based on the predicted data. Finally, the RL-HCPV utilizes Fuzzy Q-learning to make optimal decisions during on-peak and off-peak power supply to reduce costs and provide sustainable power for the appliances.

The simulation results demonstrate that the proposed DL and RL models can minimize the consumer's power costs

while maintaining user comfort. Furthermore, the RL-HCPV algorithm's performance with and without the V2H unit is analyzed, and it is confirmed that integrating the V2H unit results in considerable energy savings under the different penalty constraints of the reward function.

The remainder of this paper is organized as follows: (1) Section 2 provides a literature review, (2) Section 3 defines the system modeling of PV generation, appliance energy consumption, V2H units, EV availability, and prices, (3) Section 4 presents the proposed RL-HCPV algorithm formulation using various DL algorithms and Fuzzy Q-learning methods, (4) The simulation results for the proposed RL-HCPV model are provided in Section 5, and (6) The conclusions are provided in Section 6.

2. Related Work

According to an independent statistical analysis by the US Energy Information Administration (EIA) in 2021, the transportation sector consumes 28% of the energy produced annually, most of which comes from fossil fuels. On the other hand, residential energy consumption reaches 27% [18]. The analysis reveals an existing CO₂ concentration of 286 ppm, predicting an annual increase of 2 ppm. According to another study, if drastic measures are not taken, some regions could experience increased CO₂ emissions of up to 700 ppm by 2100 [19]. Therefore, it is crucial to radically change fossil fuel usage to decrease CO₂ emissions in the atmosphere. A major challenge to achieving this goal is decarbonizing transportation. Therefore, electrifying transport to overcome this challenge is the most promising option. Consequently, RES, such as PV solar panels and wind, are becoming the most important power sources for electrifying EVs and residences. Furthermore, V2H technology supports using clean energy for EVs and smart homes.

V2H technology refers to transferring electricity from vehicles to the home. Several pieces of literature have explained the origins of this technology. Recently, many countries have implemented rules to reduce gas and CO₂ emissions to address global warming and environmental issues and raise global energy consumption awareness. Moreover, the steady rise in oil prices necessitates the development of alternative fuel sources to reduce oil dependence [20]. As a result, electrical energy could be a viable alternative to traditional fuel sources.

Furthermore, electricity is considered an extraordinary future technology due to its infrastructure readiness, safety, and reliability. However, the storage of electricity has advantages and disadvantages. Additionally, the continuous rise in fuel prices and the emerging environmental awareness motivate manufacturers to

develop transportation solutions that overcome such issues.

Previously, EVs were not an ideal alternative due to their weak batteries. As a result, EVs and batteries have been developed as suitable alternatives for conventional vehicles. However, the driving distance of EVs is low, and battery prices are high compared to conventional vehicles. Therefore, EVs can benefit society in many ways, explicitly and implicitly. Many electric transportation projects and EV technology trade are in operation worldwide with rapid development [21]. For this reason, it is essential to adapt and acquire more innovations and conceptual advancements [22].

Studies have indicated that each EV can store up to 5–40 kWh of energy [23], which is not only used for traveling. It can also be used for V2H, V2B, and V2G applications. The study reviewed the application's characteristics, drawbacks, and pilot projects. The review revealed that V2H and V2B are easier to implement than V2G and do not need external entities to be fully operational. Also, V2H and V2B provide more apparent advantages for users. The characteristics of V2H applications are: operate at the home level, are normally suitable for one EV, reduce electricity bills, provide backup power, and are easy to implement. The main drawback of V2H is that it is only suitable for single-family homes in residential blocks. Some examples of V2H pilot projects are Leaf to Home [24], Toyota Smart Home [25], and Honda Smart Home [26]. So far, Leaf to Home is the only commercially available pilot project among the three.

A study utilized a battery-swapping strategy to integrate an EV into a building consisting of a reserve battery and solar energy [27]. The strategy supported flexible EV exiting for the owner even when the EV reaches the building with an empty battery. The simulation indicated that the building could operate continually under the grid-tied and off-grid states. In another study, the authors aimed to decrease the domestic loads' dependence on the grid, increase the power utilization generated from rooftop PV solar panels, enhance the power supply reliability compared to the residential loads during peak periods, and decrease the energy cost of a household [28]. A V2H model based on the Fuzzy Inference System (FIS) was developed to manage home energy effectively and sufficiently utilize the stored energy of EVs for residential loads. Researchers suggested a HEMS scheduling algorithm that uses V2H battery storage and solar energy to meet load demand and execute accurate system operations [29]. The study explored the impact of V2H and PV on net power usage and home energy consumption. Four energy flow models were investigated, including Green EV (GEV) charging without Green PV (GPV) (Model 1), V2H charging and discharging without GPV (Model 2), GEV charging with

GPV (Model 3), and V2H charging and discharging with GPV (Model 4). The simulation demonstrated the proposed approach's ability and proficiency to prevent unanticipated changes.

In another study, the researchers simulated eight system control modes to evaluate a proposed V2H concept [30]. The main operating modes consisted of an in-house battery charging mode where the PV-generated power is used simultaneously for in-house battery charging and load power. The in-house battery discharge mode uses the stored power of the in-house battery to satisfy the load demand. The Nissan Leaf charging mode involves charging the Nissan Leaf using the excess power remaining after the in-house battery is charged and the load demand is satisfied. The Nissan Leaf discharge mode includes using the Nissan Leaf to provide the power to satisfy the demand when the load demand is high, and the generated PV power and State-Of-Charge (SOC) of the in-house batteries are low. The in-house battery protection mode prevents the degenerative discharge condition of the in-house battery, and the system isolation mode prevents system failure. The excess power handling mode manages the excess generated power, and finally, the nighttime charging mode charging the Nissan Leaf should be prepared for the next day.

Researchers simulated an optimized strategy for the operations and facilities of a residential energy system [31]. The authors set up several scenarios and investigated the optimized strategy with different PV panel installation locations, automotive usage patterns, and energy demand characteristics. For houses with extended EV absence times, the results demonstrate that combining a V2H system with a stationary battery (SB) is more cost-effective. Thus, when designing a household energy system, it is advisable to evaluate numerous battery alternatives to determine whether a V2H system should be placed alone or in conjunction with a SB.

The above studies slightly deviate from uncertainties. Thus, efficient uncertainty prediction models are required to improve decision-making for cost reduction and DR optimization. For instance, the RnEs can predict the amount of energy produced during the day and the household energy demand when each production source and electricity consumption profile are known. For energy prediction, household energy consumption and production values are used to predict how much energy is consumed and produced by appliances. Researchers argued that EVs as energy resources can create ancillary service markets in SGs by providing ancillary services [32]. Therefore, the grid-integrated-vehicle (GIV) technology is appropriate for investing in EVs for ancillary service markets.

Furthermore, it is vital to consider individual consumption behavior when deciding optimal operational EV points for increasing user profit and resource

effectiveness. However, a mechanism that considers individual consumption behavior has not been created. Thus, a proposed V2H aggregator uses the in-vehicle batteries of EVs to enable individuals to participate in a regulated market. The results indicated that the proposed V2H aggregator could predict the power required for the grid. Studies have been conducted in which the researchers present a partially implemented software to control the energy exchange in V2H technology [33]. They implemented various basic algorithms to satisfy the load demand. The software can receive requests as XML documents from several control sources, which include long and short-term plans from smart homes or EV viewpoints. In addition, the model includes an Energy Flow Control Unit (EFCU) as a system for bidirectional EV connection within the smart home. Conventionally, EFCUs are operated manually to start and stop the charging and discharging mechanism. In contrast, the optimized model enables automatic control of the charging and discharging mechanism based on the predicted long and short-term plans using any communication technology, e.g., WiFi or Bluetooth.

A study compared two house microgrids (MG) that consisted of PV panels and wind, energy storage battery, and EV [34]. One MG incorporates the notion of a V2H unit installed in a home linked to the main grid because the EV is not available daily. The economic benefit of V2H was examined, and its cost-significance may persuade residences to move from thermal to EVs. Furthermore, the quantity of CO₂ released by EVs is half that produced by combustion vehicles, reducing pollution dramatically. As a result, CO₂ emissions will be significantly reduced during a 20-year timeframe. In addition, the study demonstrated the success of V2H in terms of electrical service continuity, energy self-sufficiency, and pricing. Finally, the study concluded by demonstrating a new research approach: Applying artificial intelligence (AI) and algorithms to predict meteorological data for optimal energy generation and EV availability and use.

Furthermore, several parameters should be considered when modeling MGs, such as pricing, generation of renewables, and energy consumption for better-managing demand and load shedding. ML algorithms are prevalent wherever AI exists [35] and are classified as supervised or unsupervised [36]. Supervised

ML algorithms are trained from a labeled training dataset. For instance, V2H applications can be charged and discharged case labels. On the other hand, unsupervised ML algorithms lack labeled output variables, whereas the training dataset is comprised of only input variables, such as producing time-series data for load forecasting.

Researchers proposed a supervised ML model to predict EV availability in V2H services to understand the vehicle availability's impact on capacity [15]. First, the model predicts the EV's start and end location (classification problem). Then, it predicts the total distance of the trip (a regression problem). Finally, the framework used for the previous three ML tasks is a Light Gradient Boosted Machine (LightGBM). This model was applied to five distinct vehicle usage profiles. First, the usage profiles predict EV availability by classifying the number of trips per week. Finally, the model is used on each user profile to earn the lowest electricity bill for each class. The findings of this evaluation indicated that this model predicts the EV's traveled distance and location with an accuracy of 85%.

Studies have been performed in which the authors proposed an unsupervised ML model for the economic analysis of the profits of dynamic pricing schemes in the different operation modes of an EV charging and discharging approach in HEMS [7]. The V2H operation mode revealed its relative financial advantage without affecting customer comfort. Furthermore, the results indicated that using V2H for minimizing electricity consumption from the grid in peak pricing periods is more beneficial than V2G or G2V.

Several subsets of ML algorithms, such as DL, utilize artificial neural networks (ANNs) by applying a layered concept hierarchy. Moreover, RL based on rewards and penalties has gained significant attention in dynamically changing environments, and there is a rising interest in data-driven techniques for modeling HEMSs. As a result, researchers are seeking elements that may provide insights and predictive analytics. Table 1 summarizes algorithms used in the most recent literature (2019–2022) to predict several parameters in HEMSs that integrate V2H technology.

Table 1: ML application in HEMS with V2H unit

Reference	Algorithm	Prediction	ML Type	Real-Time
[8]	Q-learning	Charge and discharge EV battery	RL	Charging and discharging schedules in real-time.
[9]	Mixed integral linear programming (MILP)	Optimize the energy generation of a smart home with minimal costs while considering the user's comfort	AI	Real-time pricing

[15]	One regression model and two classification problems	Predict the EV availability to provide (V2H) services	DL	-
[37]	A stochastic programming approach based on the Wasserstein distance metric and K-medoids	Manage the residential PV solar power uncertainty	AI	-
[38]	Regression	Predicted consumed energy in Madrid is calculated	DL	-
[39]	Binary Particle Swarm Optimization (BPSO) [CLASSIFICATION]	Minimizes customer dissatisfaction, peak-to-average ratio (PAR), and energy cost	AI	Real-time appliances
[29]	Energy balance control algorithm	Characterize the source details utilizing energy demand	AI	-
[40]	Twin delayed deep deterministic policy gradient (TD3) learning approach.	DR management strategy	Deep RL	Real-time DRM
[41]	Controller algorithm	Prediction for the trip	AI	-
[42]	ANN	Day-ahead solar PV generation and household load	RL	-
[43]	Double-stage MILP [Neural Network (NN)]	Optimal energy trading in real-time and day-ahead energy signals	RL	Real-time energy markets
[44]	Long Short-term Memory (LSTM) [Type of recurrent Neural Network (RNN)]	Short-term prediction of PV power production	RL	-
[45]	A stochastic model, a predictive control strategy, consists of (ANN, RNN, radial basis function NN (RBF-NN), and CNN)	Ensures home appliance thermal comfort conditions, plug-in EV battery charging requirements, and home power demand are met.	RL	Operates in real-time to minimize the influence of the gap between the actual and predicted HEMS performance
[46]	MILP	Day-ahead load scheduling provides optimal DR, cost-reduction, and PV self-consumption	RL	-
[47]	A stochastic adaptive dynamic programming (ADP)	Reduce the energy cost	AI	Including the occupant's and appliances' status for execution in real-time
[48]	A model predictive HEMS	Optimizes the charge and discharge profile of an EV and the operation of an HPWH to minimize excess electricity considering the disconnection of the in-vehicle storage battery	AI	-
[49]	Q-learning	Reduces the electricity bill within the consumer comfort level and the appliance operation characteristics	DL and RL	-

Despite extensive research on residential energy management using the RL method, real-time smart energy scheduling algorithms considering the operation of multiple smart home appliances, PV solar generation, SB, and V2H units have not been studied. The proposed approach differs from existing model-based optimization methods for HEMSs in the following ways: (1) a model-free Fuzzy Q-learning method is used to schedule energy for a smart home appliance, PV power generation, SB, and V2H unit based on Time-of-Use (ToU) and Real-Time Pricing (RTP) tariffs; and (2) prediction of appliance consumption, PV generation, prices, EV availability, and SOC using appropriate DL algorithms aids the proposed Fuzzy Q-learning algorithm. Combined with the DL models, the suggested RL-HCPV algorithm lowers customer expenditure while maintaining the desired appliance operation. The simulations depict a single home with PV solar panels, SB, smart appliances, and a V2H device under ToU and RTP tariffs.

3. System Modeling

In this study, we analyze the instance in which the HEMS, which schedules and regulates the shiftable and non-shiftable household appliances under the ToU and RTP tariffs, performs automated energy management for a single house. Figure 1 depicts the conceptual system model for the proposed RL-HCPV and the data classification for the customer, weather station, and utility company.

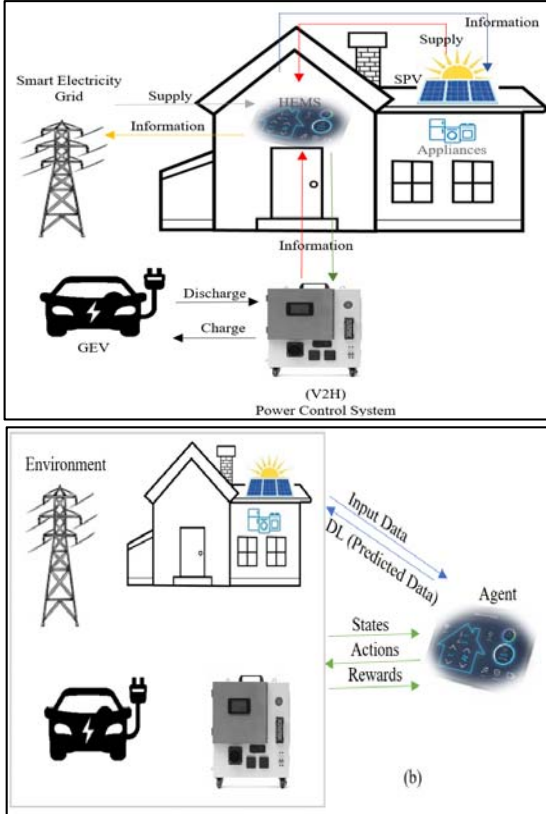


Figure 1: (a) A schematic of the RL-HCPV concept (b) DL and RL in RL-HCPV

3.1 Objective Function

This model aims to build an ideal smart MG that reduces investment and energy expenses while ensuring V2H technology integration. This is accomplished by increasing home autonomy while decreasing grid energy consumption. The investment cost is comprised of the purchase and installation of all equipment, including PV panels, SB, V2H units, electronics, cables, and operation and maintenance costs. On the other hand, the energy cost mainly consists of the grid operator's energy bills. The objective function is to ensure the optimal operation of the smart MGs with minimized operating costs. Furthermore, battery installation capacity should be adjusted by incorporating the deterioration cost. Equation (1) can be used to optimize the cost function [34, 50]:

$$C_{total} = C_{opr} + C_{pr} \quad (1)$$

Where C_{opr} corresponds to the operation cost, which is given by:

$$C_{opr} = C_{PV} \cdot PR_{PV}^{ins} + C_{SB} \cdot CAP_{SB}^{max} \quad (2)$$

C_{PV} is the PV panels cost in (\$/W), and C_{SB} is the SB cost in (\$/J); PR_{PV}^{ins} is the PV power installed and E_{SB}^{max} is the battery capacity.

C_{pr} denotes the cost of the electricity supplied by the utility, and it can be expressed as follows:

$$C_{pr} = C_{MG} \cdot \sum_{i=1}^{24} (\sum_{t=1}^{hour} (PR_{MG}^i(t))) \quad (3)$$

C_{MG} is the utility grid cost, which fluctuates during the day: off-peak periods are cheaper than on-peak periods; $PR_{MG}^i(t)$ is the power provided by the utility, which is reduced in this study.

3.2 Power Balance Modeling

In residential homes, examples of both shiftable and non-shiftable appliances include lights, kitchen appliances (ovens, stoves, kettles), washing machines, dishwashers, air conditioners, and electric vehicles (EVs). Kettles, stoves, refrigerators, and air conditioners are examples of non-shiftable equipment that have a steady power output for a predetermined amount of time. However, appliances like washing machines and electric vehicles can be used at other times. For instance, shiftable equipment (washing machines) can be used to conserve energy during non-peak times.

Energy costs can be reduced by rescheduling shiftable equipment. However, this may impede user comfort. Furthermore, heavy electrical demand occurs between 5:00 and 9:00 PM as grid power prices steadily rise to their maximum. Therefore, a more flexible strategy is necessary to satisfy the comfort and financial requirements of residential customers. SB energy storage systems are an excellent choice for home power demand for peak shifting or shedding reduction. They can be classified as time and power-shiftable loads or power supply resources. In this study, SB and V2H units are utilized to balance demand and supply, reducing both expense and discomfort. The limits for PV, SB, and EV battery charging and discharging power and utility costs are provided by Equations (4–7).

$$0 < Pr_{pv}(t) < Pr_{pv,max} \quad (4)$$

$$SOC_{SB,max} < SOC_{(t)} < SOC_{SB,min} \quad (5)$$

$$SOC_{EV,max} < SOC_{(t)} < SOC_{EV,min} \quad (6)$$

$$C_{MG,min}(t) < c_1(t) < C_{MG,max}(t) \quad (7)$$

Where, $Pr_{pv}(t)$ is the power of PV at time t and $Pr_{pv,max}$ is the maximum generated power by PV panels. $SOC_{SB,max}$ and $SOC_{SB,min}$ are the maximum and minimum SOC limitations for the SB to avoid battery degradation. Likewise, $SOC_{EV,max}$ and $SOC_{EV,min}$ are the maximum and minimum SOC limitations for the EV battery. $C_{MG,max}(t)$ and $C_{MG,min}(t)$ are the maximum and minimum utility costs. Equation (8) presents the power balance over the operation.

$$Pr_{comp}(t) = Pr_{MG}(t) + Pr_{pv}(t) + Pr_{EV}(t) \quad (8)$$

$Pr_{comp}(t)$ represents the appliances power consumption at time t as the total power supplied by Pr_{MG} , Pr_{pv} , Pr_{EV} , the grid, PV panels, and EV, respectively.

3.3 Modeling of Appliance Consumption

The household appliances in this research are classified based on their energy consumption patterns as follows (See table 2):

- Non-shiftable appliances
- Shiftable interruptible appliances
- Fixed appliances

Non-shiftable appliances are designed to be easily utilized and cannot be stopped once the appliance is turned on. However, before operating, it might be moved from one-time slot to another. Shiftable interruptible appliances may be turned off at any moment. They help save power costs when turned off during peak hours [50]. Fixed appliances are base appliances that are neither shiftable nor interruptible and require a continuous power supply. Non-shiftable appliances' power consumption may be divided into two modes, signifying high and low consumption. Consequently, the price of the non-shiftable appliance power consumption at time t ($C_{NSH}(t)$) is calculated as follows (eq. 9):

$$C_{NSH}(t) = CT(t) \times Pr_{comp}(t) + \beta_{NSH} \times (Pr_{comp}^{mode}(t) - Pr_{NSH,max})^2 + \hat{C}(t) \quad (9)$$

Here $CT(t)$ is the electricity tariff at time t , and β is the dissatisfaction price, $Pr_{comp}^{mode}(t)$ is the power consumption of non-shiftable appliances during the mode at time t , $Pr_{NSH,max}$ is the power rating of the non-shiftable appliance. Moreover, $\hat{C}(t)$ indicates the predicted consumption cost at a specific time. Various requirements must be met before scheduling may begin. The following is how the cost of shiftable interruptible appliances is calculated (eq. 10):

$$C_{SH}(t) = CT(t) \times Pr_{comp}(t) + \beta_{SH} \times (S_{SH,updated} - S_{SH,scheduled}) + \hat{C}(t) \quad (10)$$

$S_{SH,scheduled}$ is the scheduled start time of the appliance and $S_{SH,updated}$ is the updated operating time. Thus, β_{SH} can be used to specify how quickly an appliance should complete its job. Moreover, $Pr_{comp}(t)$ is the power consumption of a shiftable appliance at time t . Finally, as a limitation, the appliance must complete its operation within the following twelve hours of being scheduled.

Table 2: Home appliances' parameters

Households	Category	Starting Time	Ending Time	Total Running Duration	Power rating (kW)
Washing-Machine	Non-shiftable	9 pm	10pm	2h wash	0.38
Vacuum-Cleaner	shiftable	1 pm	3pm	0.50h	1.25
Air Conditioner	shiftable	1 pm	10 pm	9 h	1.0
Rice-Cooker	Non-shiftable	10 am	12 pm	0.75 h	0.80
Humidifier	shiftable	6 am	12 am	4 h	0.05
Light	fixed	6 am	612 am	Full time-24h	1.20
Oven	fixed	2pm	10pm	2h	2.0
Hair-dryer	shiftable	7 am	8.30 am	30 min	1.80
Dish Washer	Non-shiftable	8 am	11pm	3h	0.6
Electrical Water	shiftable	6 am	8 pm	0.50h	1.5
Robot	shiftable	11am	2 pm	2h	2.0

3.4 Modeling of PV Generation

Recently, solar PV panels have been widely used in many types of buildings and are likely to be one of the key renewable energy supplies in smart MGs. However, the generated electricity performance greatly depends on seasonal and climatic conditions. The unpredictability of the climate influences generated power and negatively impacts the grid's stability, dependability, and operation. As a result, an accurate PV generation forecast is a critical requirement for ensuring the smart MG's stability and dependability [51]. The following formula (eq.11) is used

to determine a building's solar PV power generation at each time interval [52]:

$$Pr_{pv,t} = a_{pv} \times f_{pv} \times r(t) \quad (11)$$

a_{pv} , f_{pv} and r_t indicate the area of the solar panels, PV efficiency, and PV irradiance, respectively. There are three different possibilities depending on the value of the net power (eq.12) [34]:

- $\overline{Pr}_{PV} = 0$: Renewables completely meet the load, and no electricity from SB, EV, or utility is required.
- $\overline{Pr}_{PV} > 0$: SB, utilities, and EVs are necessary to compensate for the shortage of renewables. The

SOC_{SB}, the utility prices, and the availability of EVs determine the utilization of each of these resources.

- $\overline{Pr}_{PV} < 0$: Overproduction is stored in an SB-free capacity. Excess power is transferred to the grid if the battery is fully charged or the demand exceeds the maximum power limit. However, the EV battery is a backup power source if the EV is available or the grid is insufficient.

$$\overline{Pr}_{pv} = Pr_{comp} - (Pr_{pv} + Pr_{sb} + Pr_{ev}) \quad (12)$$

The SB is the first to intervene when there is a renewable energy shortage ($\overline{Pr}_{PV} > 0$). It can supply the power and maintain the balance between supply and demand when its fully charged and the demand is below its maximum discharge power limits. However, the SB supplies its maximum power while the grid supplies the remaining power if the discharge power exceeds the maximum SB discharge power limits. Although the grid is a major energy supply, its power might be restricted due to reliability and security concerns. Therefore, the grid provides the load if the demand satisfies the grid power constraints. Otherwise, the EV serves as a backup to provide an uninterrupted power supply. Finally, load shedding can be used if the EV maximum power limits are exceeded or the EV battery is prohibited due to the minimum SOC_{EV}.

In contrast, the SB is used to store the surplus generated power during the instance where there is overproduction ($\overline{Pr}_{PV} < 0$). Moreover, the utility grid interferes and maintains the power supply and demand balance if the SB is completely charged or the demand exceeds the SB maximum charge power limits. This restriction on grid power aims to increase smart MG stability and autonomy. Thus, EV stores a portion of the generated power when the grid's maximum power has been reached.

3.5 Modeling of EV SOC and Availability

The proposed management technique also considers the EV battery's SOC which is a significant metric because it indicates the amount of energy stored. Equation 13 is a typical assessment of the EV battery's SOC based on power charging and discharging.

$$SOC_{EV} = SOC_{int} + \frac{\Delta Pr_{g2v} + \Delta Pr_{v2h} + \Delta Pr_{v2g}}{CAP_{EV}^{full}} \quad (13)$$

SOC_{int} is the starting SOC and CAP_{EV}^{full} is the EV battery's power capacity. Given the EV battery's lifespan, several limits have been placed on the power produced by the EV through the V2H unit and the battery SOC. In addition, ΔPr_{g2v} represents when the EV battery is charged from the grid, ΔPr_{v2h} is the energy supplied to the home, and ΔPr_{v2g} is the energy returned to the grid.

According to equation 14, EV power is limited between the lowest operational power ($Pr_{EV,min}$) given to

the EV and the largest power range ($Pr_{EV,max}$) to be provided to the home or grid (V2G or V2H).

$$Pr_{EV,max} \geq Pr_{dis}, Pr_{ch} \geq Pr_{EV,min} \quad (14)$$

Equation 15 similarly indicates the inhibition of EV battery deep discharge and full charging by setting a minimum and maximum SOC. i.e., (SOC_{EV,min}) and (SOC_{EV,max}).

$$SOC_{EV,min} < SOC_{EV} < SOC_{EV,max} \quad (15)$$

The ability to ensure if the EV is parked at home and available for V2H connections is defined as follows by equation 16:

$$Loc_{EV} = \begin{cases} 1 & EV \text{ available} \\ 0 & EV \text{ not available} \end{cases} \quad (16)$$

3.6 Modeling of SB

The energy balance equation for SB is given as follows (eq.17):

$$L_{SB}(t) = (1 - L_{SB})L_{SB}(t-1) + T(Pr_{SB,ch}(t) f_{SB,ch} - \frac{Pr_{SB,dch}(t)}{f_{SB,dch}}) \quad (17)$$

$L_{SB}(t)$ represents the energy storage level at time t and L_{SB} is energy loss. $Pr_{SB,ch}(t)$, $f_{SB,ch}$ represent the SB's charging power and efficiency; $Pr_{SB,dch}(t)$ and $f_{SB,dch}$ represent the discharging power and efficiency, respectively. The SB's stored energy is constrained by its minimum and maximum values, SOC_{min} and SOC_{max}, as follows:

$$SOC_{SB,min} Pr_{SB} \leq Pr_{SB}(t) \leq SOC_{SB,max} Pr_{SB} \quad (18)$$

To guarantee that the battery is full, the energy storage level of the SB at the end of the day, T, is assumed to be identical to its original value, as follows:

$$Pr_{SB,0} = Pr_{SB,T} = Pr_{SB,int} \quad (19)$$

The following constraints ensure that $Pr_{SB,ch}$ and $Pr_{SB,dch}$ are less than the SB's maximum charging and discharging power ratings:

$$\begin{aligned} 0 &\leq Pr_{SB,ch}(t) \leq m_{SB,ch}(t) Pr_{SB} \\ 0 &\leq Pr_{SB,dch}(t) \leq m_{SB,dch}(t) Pr_{SB} \end{aligned} \quad (20)$$

$m_{SB,ch}(t)$ and $m_{SB,dch}(t)$ are binary variables representing the charging and discharging operation modes of the SB. It is worth noting that equation 20 requires the charging power to be zero if the SB is not in charging mode (i.e., $m_{SB,ch}(t) = 0$) and $m_{SB,dch}(t)$ for the discharge mode.

The SB can only function in the charging or discharging modes at a time, as defined by the following restriction (eq. 21):

$$m_{SB,ch}(t) + m_{SB,dch}(t) \leq 1 \quad (21)$$

3.7 Modeling of ToU and RTP Pricing

Different power tariffs (e.g., ToU, CPP, and RTP.) have been introduced under the DR program to minimize electricity use during on-peak hours. The DR method

incentivizes load transfer from on-peak to off-peak hours, where costs are lower during off-peak hours. Prices increase when the demand for power rises during on-peak hours, and the extra demand is satisfied by expensive utility tariffs. Several factors, including the generation cost, determine the electricity price per unit. The key factors influencing electricity prices are distribution, transmission, and generation costs. In addition, prices are greatly impacted by the utility and consumer demand. The utility is a service provider that regulates power rates on a seasonal (ToU), weekly, daily, or hourly basis (RTP).

RTP and ToU are popular pricing tariffs across the world. However, the RTP fluctuates in response to the demand and wholesale price. This variation may cause the customer some discomfort. The ToU tariff is divided into three blocks: on-peak, shoulder-peak, and off-peak hours, with pricing remaining constant throughout the season during these periods. Utilities determine pricing rates for each block of ToU after monitoring the consumers' past energy consumption behavior.

The adjusted ToU tariff is based on extra generation costs and user load demand during particular hours. Changes are implemented per hour unit for on-peak and shoulder-peak hours. The formal expression of the adjusted ToU price signal (C_{ToU}) is given as (eq. 22) [54]:

$$C_{ToU}(t) = \begin{cases} ToU & \text{if } t_{off} \\ C^+ & \text{else} \end{cases} \quad (22)$$

t_{off} indicates the off-peak hours and C^+ denotes extra cost. Utilities set the pricing tariffs based on demand, with electricity costs being higher during on-peak hours and cheaper during off-peak hours. Off-peak hours are when households and companies consume the least power. The RTP price is an hour-ahead, day-ahead, and week-ahead price. The day-ahead and week-ahead pricing approaches are based on prior power usage, while RTP hourly price is the daily prediction of demand and cost. It is simple to program electrical appliances for day-ahead and week-ahead pricing systems. However, hourly price scheduling is rather challenging because the price fluctuates between on-peak and off-peak hours. Load forecasting can be employed to address this issue [55]. However, shiftable appliances are adjusted from on-peak to off-peak prices. Vacuum cleaners, humidifiers, and hairdryers are regarded as shiftable appliances. The total hourly power consumed by all shiftable appliances (SH) in a day is determined as follows:

$$C_{RTP,SH}(t) = \sum_{SH}^t (P_{sh} \cdot m_{sh}(t)) \quad (23)$$

$T = 24$ and P_{sh} signifies the power rating for each shiftable appliance. $m_{sh}(t)$ denotes the hourly status of each

shiftable appliance (ON or OFF). Each appliance's state for the next period is indicated as follows:

$$\hat{m}_{sh}(t) = \begin{cases} 1 & \text{if } \hat{m}_{sh} \text{ is ON} \\ 0 & \text{if } \hat{m}_{sh} \text{ is OFF} \end{cases} \quad (24)$$

The second appliance type is non-shiftable, meaning they are not interrupted or shifted. These types of appliances may not help with load scheduling or energy management. Non-shiftable appliances must follow a preset execution pattern. This study considers dishwashers and rice cookers as non-shiftable appliances. Equation 25 calculates the hourly total power consumption (kWh) in a day for non-shiftable appliances ($C_{RTP,NSH}(t)$).

$$C_{RTP,NSH}(t) = \sum_{NSH}^t (P_{nsh} \cdot m_{nsh}(t)) \quad (25)$$

Where P_{nsh} is the rating of power and m_{nsh} indicates the status of each non-shiftable appliance (ON or OFF).

$$\hat{m}_{nsh}(t) = \begin{cases} 1 & \text{if } \hat{m}_{nsh} \text{ is ON} \\ 0 & \text{if } \hat{m}_{nsh} \text{ is OFF} \end{cases} \quad (26)$$

3.8 Modeling of V2H unit

Equation 27 expresses the total power saved in the EV battery [8].

$$\overline{Pr}_{EV} = Pr_{EV,int} + \Delta Pr_{g2v} - \Delta Pr_{v2h} - \Delta Pr_{v2g} - Pr_{trip} \quad (27)$$

\overline{Pr}_{EV} indicates the overall amount of energy in the EV battery, $Pr_{EV,int}$ is the initial power stored in the EV battery, ΔPr_{g2v} is the power transferred from the grid to the EV battery, ΔPr_{v2h} indicates the power supplied from the EV to the home, ΔPr_{v2g} is the discharged power from EV to the grid, and Pr_{trip} indicates the total EV power consumption during the trip.

$$\Delta Pr_{v2h} = m_{V2H} \int_{in}^{out} P_{ach}(t) \cdot dt \quad (28)$$

P_{ch} and P_{dis} are the EV charging and discharging power, respectively. Additionally, m_{V2H} represents the EV connection mode (0 or 1).

4. Proposed RL-HCPV

Reinforcement Learning (RL) is an ML computational technique in which an agent attempts to increase overall reward by executing actions and engaging with an anonymous environment. The suggested RL-based data-driven technique in this research is divided into two sections, as seen in Figure 2, (i) the appropriate DL algorithms are trained for predicting the future PV generation, appliance consumption, EV SOC, EV availability, and electricity prices (ToU and RTP) (ii) a Fuzzy Q-learning algorithm is created for generating efficient one-hour-ahead DR decisions. The next

subsections go into the specifics of this data-driven solution strategy.

The one-hour-ahead prediction of PV generation, appliance consumption, EV SOC, EV availability, and electricity prices allows effective scheduling for charging/discharging the energy storage units, including SB and EV batteries, using a V2H unit. To ensure optimal and intelligent management of the DR of this SG, an RL-HCPV algorithm is proposed in Table 3 and Figure 3. The predicted data are used as input parameters in the Fuzzy

Q-Learning model to create the ideal schedule for energy storage units charging and discharging while taking into account all energy and transport demand restrictions. This section explains the process of predicting input parameters of the model. Several DL algorithms, including artificial neural networks (ANN), Long Short-Term Memory (LSTM), and multi-target regression, are used to predict the uncertainty parameters; PV generation, appliance consumption, EV SOC, EV availability, and electricity prices from historical time-series datasets.

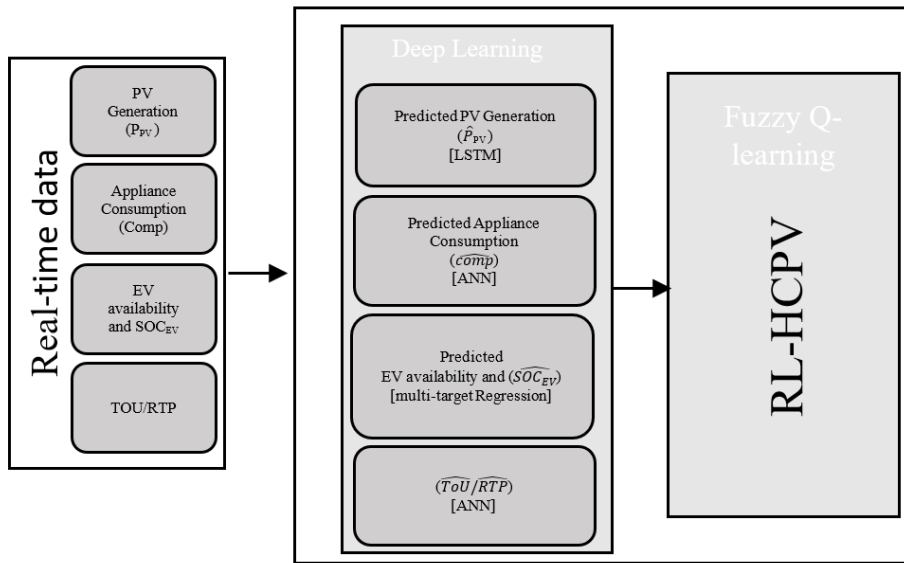


Figure 2. Schematic of the RL-HCPV system

Table 3: Proposed RL-HCPV algorithm

Algorithm 1 Proposed RL-HCPV Algorithm
Initialize using time, \hat{P}_{pv} , \hat{P}_{comp} , \hat{SOCEV} , $\hat{TOU/RTP}$, \hat{C}_{RTP} , \hat{C}_{TOUs} For time slot $t=1:T$ do For HEMS do Execute LSTM for PV Generation End Receive extracted information about future PV Generation Execute ANN for Appliance Consumption' End Receive extracted information about future total appliance consumption Execute Multi-target regression for EV availability and SOC End Receive extracted information about future EV patterns Execute ANN for Pricing End Receive extracted information about future ToU and RTP For each t do Sort descending by T Execute Algorithm 2 End End
Algorithm 2 Fuzzy Q-Learning (Decision Making)
Initialize Environment ActionSet = [], S = [] Get current state S, ActionSet Initialize Fuzzy Set S, u=[] Initialize FIS StateSpace Set Membership function Initialize FQL model S, γ , α , ee-rate, ActionSet, FIS Set Q=0, V=0, Error=0 Set Q-table=matrix ([index]) For each episode σ Select Action α Set Inferred Action α_i If ($\alpha_i \in u$) Reward l

```

Else
Reward 0
Receive immediate reward  $r_t$ 
Calculate TruthValue, V, Q, Quality
Return  $Q(S_t, \alpha_t)$ 
Update Q-Value
Return  $Q(S_{t+1}, \alpha_{t+1})$ 
Keep State History
End
    
```

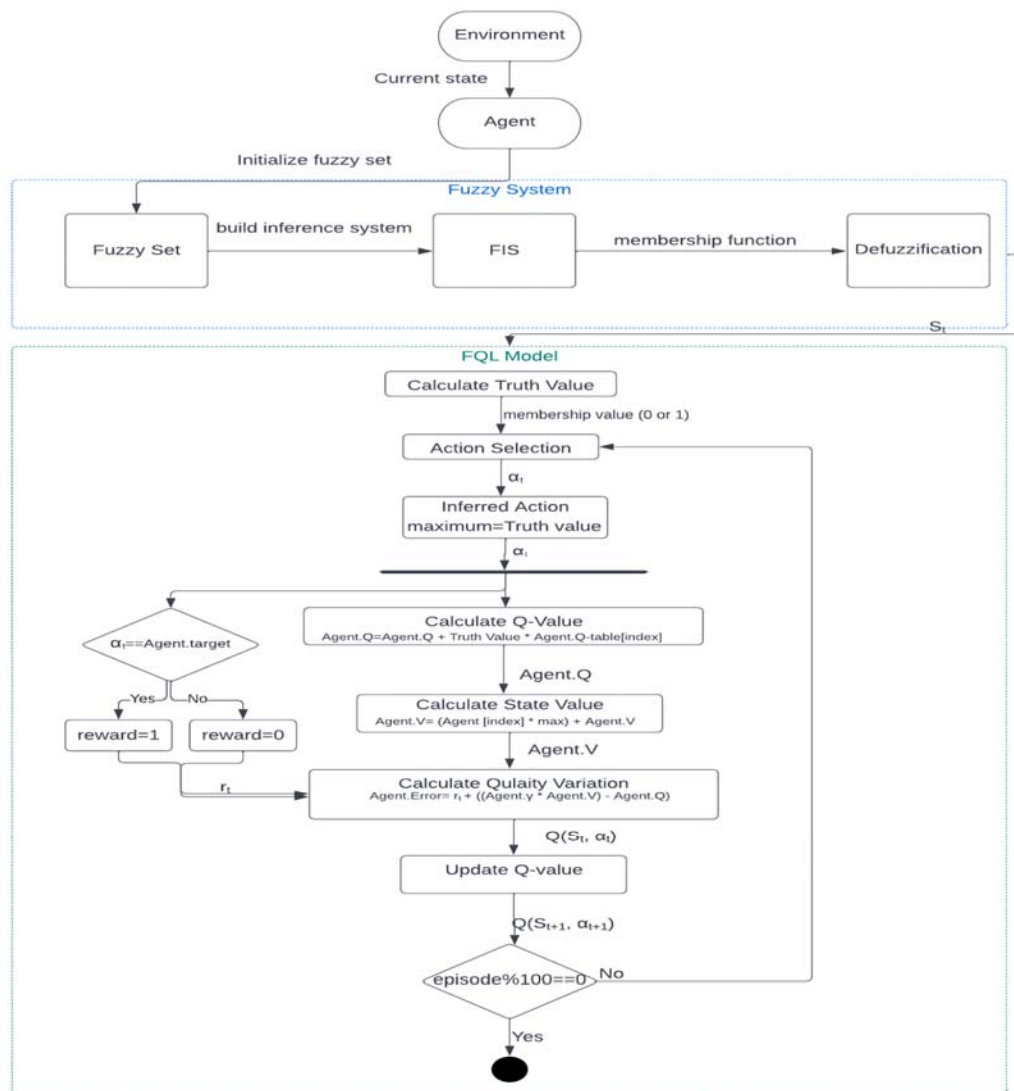


Figure 3. FQL algorithm for RL-HCPV system

4.1 Prediction of Uncertainty Parameters of HEMS

A. DL algorithms for uncertainty prediction

Applying DLs to predicting PV power generation helps to overcome the limitations of statistical approaches in handling non-linear data caused by weather variation. DL is an effective prediction approach with many applications, such as appliance consumption, weather conditions, and

wind speed prediction [27, 42]. The ANN algorithm has become a prominent tool for predicting renewable power generation [58], electricity price [57], and load [56]. ANN is one of the easiest ways to set up building models, even with only minimal physics knowledge. ANNs are constructed by training, validating, and testing datasets. The input, hidden, and output layers are the three primary components of ANNs. The input layer receives the input

data, and the hidden layer, which consists of numerous layers, examines the input data. After acquiring analyzed information from the hidden layer, the output layer provides the output.

B. Uncertainty management strategy for RL-HCPV

In this study’s proposed model, to deal with the uncertainties of PV generations, appliance consumption,

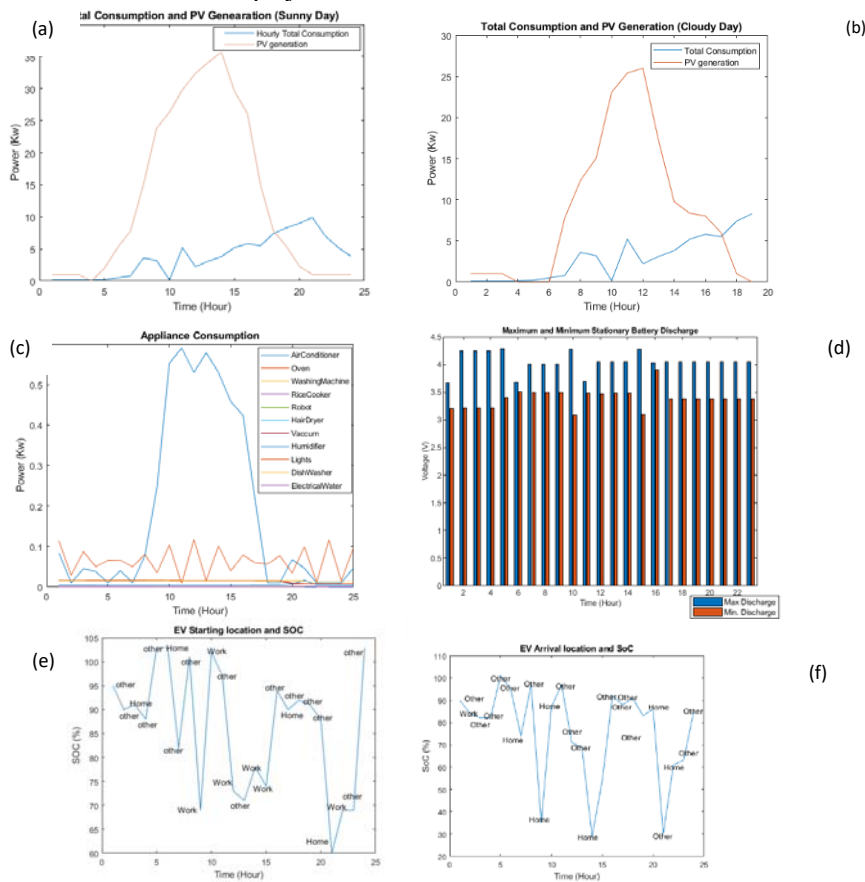
EV SOC, availability, and electricity prices, several DL algorithms were applied to predict the future trends of these uncertainty parameters. Several DL algorithms (ANN, LSTM, SVR, Multi-target regression, DT) have been employed to decide the most accurate algorithm for each uncertainty parameter. Therefore, the appropriate DL algorithm is set for each uncertainty parameter (see Table 4). The input for each DL algorithm is shown in Figure 4.

Table 4: DL algorithms for each uncertainty parameter

Uncertainty Parameters	DL algorithm	Input	Output
PV Generation	LSTM	PV solar generation (training data set)	Predicted PV solar generation
appliance consumption	ANN	appliance consumption (training data set)	Predicted appliance consumption
EV SOC and availability	Multi-target regression	Historical data (SOC-End and End-location)	Predicted SOC and End-Location
ToU tariffs	ANN	ToU Price (training data set)	Predicted ToU Price
RTP tariffs	ANN	Startup Price, Intermediate Price and no load Price (training data set for each column separately)	Predicted Startup Price, Intermediate Price and no load Price

Every hour, the inputs of the historical data of PV generations, appliance consumption, EV SOC, location, and electricity price data are used for the DL algorithms. The outputs are predicted for 24 hours. The projected

information will then be placed into the decision-making process, Fuzzy Q-learning for efficient power decision-making.



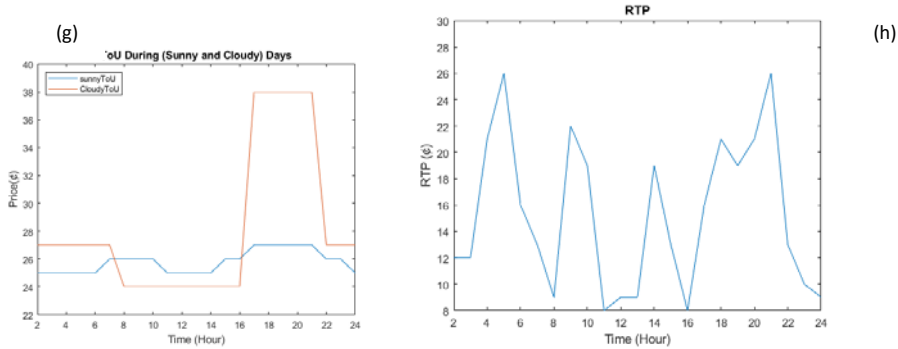


Figure 4: Input experimental profiles: (a–b) Total appliance consumption and PV generation (c) Appliance consumption (d) Stationary battery SOC (e–f) EV availability and SOC (g) ToU (h) RTP

In this model, each uncertainty parameter was manipulated separately to ensure the most accurate output predictions. The DL models were implemented for each uncertainty parameter using Python. First, an ANN model for appliance consumption was applied using a training dataset in the input layer. Three hidden layers were included, with 32 neurons for each layer and one output layer for the predicted power consumption values. The model displayed the accuracy of mean absolute error (MAE) at 0.21 and mean square error (MSE) at 0.22. In addition, LSTM modeling was applied for PV power generation. The PV generation training dataset was used in the input layer. Three hidden layers were included, with 96 neurons for each layer and one output layer for the predicted PV solar generation values. The model results displayed the accuracy of MAE at 0.37 and MSE at 0.27. Multi-target regression is advantageous for EV SOC and availability because it predicts several continuous target variables concurrently based on the same input data set. Two target columns were considered; EV arrival location and SOC at every time interval (one hour). The model displayed the accuracy of MAE at 0.31 and MSE at 0.15. Finally, ANN modeling was applied for ToU and RTP tariffs. The observations, including the trained data set prices, were regarded as the inputs for the neural networks at every time interval (one-hour). The input layer had 15 neurons for the prices training dataset. Three hidden layers were included, 32, 18, and 9, respectively, to obtain the output layer for the predicted prices. The model displayed the accuracy of MAE at 0.37 and MSE at 0.32.

4.2 RL for Decision-making in HEMS

A. Fuzzy Q-Learning Overview

The Fuzzy Q-learning algorithm is a fuzzy variation of the Q-learning method. The input states and their related fuzzy sets must be identified before building a FIS to combine with the Q-Learning method [59]. As a new ML method, the Fuzzy Q-learning algorithm [60] is

frequently employed in decision-making to maximize cumulative rewards. This algorithm's primary mechanism is to establish a Q-table in which the Q-value, $Q(s_t, a_t)$ in each iteration is modified using equation 29 until the convergence condition is met.

$$Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha[r(s_t, a_t) + \gamma \cdot \max_{a'} Q(s_{t+1}, a_{t+1})] \quad (29)$$

The discount rate, γ defines the relationship between future and present rewards. It accepts a value between 0 and 1, where “0” indicates that the agent depends solely on the present reward, whereas “1” implies that the agent will compete for future rewards. While ϵ ($0 < \epsilon < 1$) represents the learning rate and indicates how much the new reward influences the previous value of $Q(s_t, a_t)$. For example, $\epsilon = 0$ indicates that the information currently collected is negligible and the received reward will not affect the Q-value. Whereas $\epsilon = 1$ means that only the most recent information is considered.

B. Fuzzy Q-Learning model for decision making in RL-HCPV

Under ToU and RTP dynamic power pricing and varied energy consumption patterns, RL produced a real-time smart energy scheduling option for a smart home. The RL process is defined by a state space and a numerical reward given to the agent after performing a specific action to assess the new state. The Fuzzy Q-learning in our proposed model consists of state-space and reward function as follows:

- *State-space implementation using fuzzy logic:*

The state space here was represented by 16 states, as displayed in Table 5. PV power generation and appliance consumption were classified as low or high levels to simplify the model and decrease calculation time. The SB and EV battery SOC was defined as efficient or not efficient, while the EV availability was defined as

available or unavailable. In addition, the weather was either cloudy or sunny. Finally, the prices were categorized as cheap and expensive.

• *Action space:*

The suggested model aims to supply the smart home with power from PV, V2H, and SB during on-peak hours. During a power outage, the power supply is scheduled during off-peak hours when electricity rates are cheaper. As a result, the activities can be classified as V2H charging, SB charging, ToU tariff, and RTP tariff. PV-generated energy serves as the standard power source. Based on the situation, the agent chooses one action from action space A (see Table 5) that given by :

$$A=[V2H \text{ discharging, PV discharging, SB discharging, ToU tariff, RTP tariff}] \quad (30)$$

Table 5: Fuzzy Q-learning states and actions for RL-HCPV

States	Actions				
	V2H discharging	PV discharging	SB discharging	ToU tariff	RTP tariff
PV generation-High		•			
PV generation-Low	•				
SB SOC efficient			•		
SB SOC not efficient	•				
Appliance consumption-high	•				
Appliance consumption-low		•			
EV SOC efficient	•				
EV SOC not efficient			•		
EV available	•				
EV not available			•		
ToU Cheap					
ToU expensive	•				
RTP Cheap					
RTP expensive	•				
Weather (Cloudy)			•		
Weather (Sunny)		•			

• *Reward function implementation using fuzzy logic:*

Once the agent performs a random action and observes the new state, it earns a numerical reward r_t (s_t, α_t). The value of this reward reflects the adequacy of the agent's action for a present state. Thus, fuzzy logic is utilized to evaluate the action to respond to a specific state. The output of the FIS system's state space is s_t , which indicates the present state. Then, s_t will be used as the input

variable of the reward function's FIS. Therefore, evaluations of the random actions resulting from FIS make up fuzzy Q-learning. The fuzzy sets are defined as very good (VG), good (G), and bad (B) actions for each action taken (output). According to the membership function (MF) discourse, all possible acts are either rated as 0 or 1. First, the state-space FIS system will determine the present state. Next, the agent will choose a random action from the action space depending on the current space. The FIS system of the reward function also evaluates all potential actions with a value of 0 or 1. Then the agent will receive a numerical value that represents the action taken.

C. Fuzzy Q-Learning management strategy for RL-HCPV

The Fuzzy Q-learning method was used to determine the best policy after obtaining the expected energy prices, PV solar generation, appliance use, and EV SOC. In this scenario, RL is used to determine the ideal DR, given the fluctuating cost of electricity and the variety of energy usage patterns.

The agent can employ the off-policy to learn by performing a random action in a specific state while not adhering to a current policy.

Thus, Fuzzy Q-learning supports the off-policy by making the best decision given the present state. This implies that no policies are required during the training process. Initially, all the Q-values' pairs in the Q-table should be set to zero. After that, the agent will interact with the environment and use equation 31 to update the Q-values after each action.

$$Agent.Q = Agent.Q + Truth \text{ Value} * Agent.Q\text{-table}[\text{index}] \quad (31)$$

In this paper, a random action called "exploring" was used. In this scenario, a suitable number of repetitions will be required to explore and update the values of r_t (s_t, α_t) for all the Q-value pairs at least once. The training phase of the model suggested by this study was set to 5000 iterations so that the agent could visit all pairings and gather new information. As a result, when the Q-table converges, the optimal Q-values will be obtained.

$$Agent.Error = r_t + ((Agent.\gamma * Agent.V) - Agent.Q) \quad (32)$$

Figure 3 depicts the technique of employing Fuzzy Q-learning in the RL-HCPV algorithm. Generally, Figure 5 illustrates how fuzzy logic decides on a particular state and reward. The parameters γ and α are set at 0.9 and 0.1, respectively, and the entries in the Q-value table are set to zeros. All possible actions are listed for each present state, and then an action is randomly selected. When the stated action is finished, the agent will use fuzzy logic in this step to detect the new state and the reward. Equation 31 will calculate the maximum Q-value before updating the state-

action Q-value pairs. Finally, the present state will then be the following state.

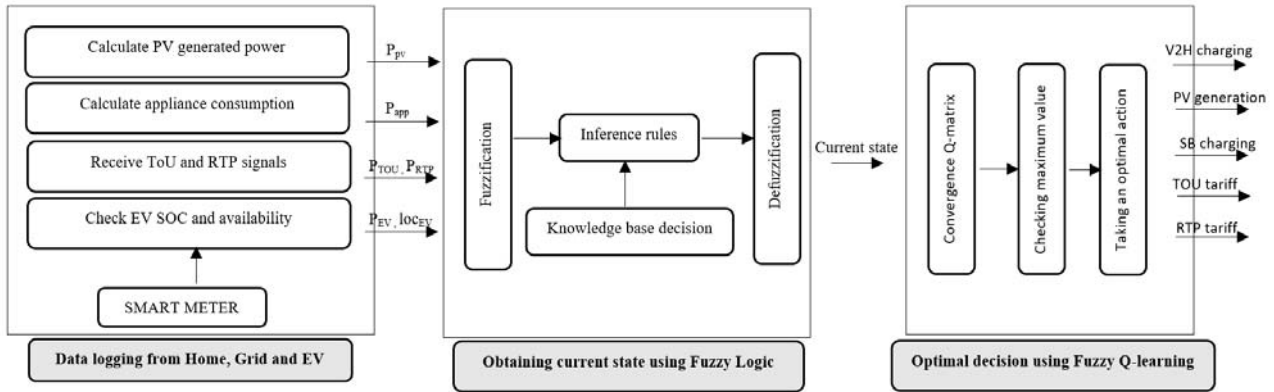


Figure 5: Flowchart describing the proposed optimal decision-making method

5. Results and Discussion

5.1 Implementing the Proposed Model for RL-HCPV

The proposed RL-HCPV method is based on the interactions between PV power generation, appliance power consumption, electricity pricing, V2H unit, EV SOC, and availability. A smart meter uses the smart house to gather power data such as PV power generation, appliance consumption, SOC_{SB}, and energy price signals from the utility, which are then sent to the HEMS. When the EV returns home, HEMS will receive a signal indicating EV availability and the SOC percentage, as shown in Figure 2. The data collected helps to determine whether the PV-generated power and SOC_{SB} are sufficient to meet the demand. If the EV can transmit electricity to the home via V2H, HEMS may be the best alternative in a power outage.

The simulation runs for 24 hours, using an hour-long time step. First, the power utility provides the price tariff signal. Then, the user is expected to depart at 7:00 AM and return at 2:00 PM. Figure 4(c) displays the overall power requirements of household appliances. There are two peak times: 2:00–5:00 PM, when the household member returns from work, and 7:00–9:00 PM, when they begin cooking, watching TV, and other activities. On sunny and cloudy days, the off-peak time begins at midnight and lasts until early morning. The suggested approach sought to understand how using DL, RL, and V2H affected power consumption efficiency while reducing expenses. The two smart MG design scenarios emphasize the necessity of

including V2H as a different energy source in the RL-HCPV model.

The simulation was tested on a Windows 11 PC with an AMD Ryzen CPU clock speed of 1.80 GHz and 16 GB RAM. The simulation software used was Real-time Simulink tool from MATLAB R2022a, while the programming software was Python.

5.1.1 Case 1: Smart MG without V2H unit

Depending on the relationship between the overall power usage, the price of electricity, and the Q-values, the standard SG, in this case, comprises PV solar panels, SB, and appliances. All the executed activities are depicted in Figure 6(a). For example, during 5:00–7:00 PM, the total energy consumption is low (3kWh–5.7kWh), the ToU energy price is 0.24 \$/kW, RTP is 0.21 \$/kW, and PV generation is low (5 kWh). As a result, the current state is $s_t = \text{low; low; expensive; expensive}$. According to the Q-value, the RTP action has the highest Q-value for this state. After 7:00 AM, the system switches to another state since demand exceeds (16 kWh), prices are high (ToU 0.28 \$/kWh and RTP 0.27\$ kW), and PV generation is high (10 kWh). The current state is $s_t = \text{high; high; expensive; expensive}$. The PV discharging action has the highest Q-value for this condition according to the Q-value.

During the evening peak period, 7:00–9:00 PM, the consumption and prices are increased, and the PV generation is low, but the SB saves the excess power during sunny days (40 kWh). Thus, the current state is $s_t = \text{low; high; expensive; expensive}$. The SB discharge mode provides energy to appliances while lowering power costs. Finally, from 9:00 PM–04:00 AM, consumption is 0.95 kWh, PV generation is 0.3 kWh, and SB power is 3.5 kW. The current states are all low. Thus, ToU (0.21\$/kW) will

be activated, which is cheaper compared to RTP (0.27\$/kW).

5.1.2 Case 2: Smart MG with V2H unit

Depending on the link between the overall power usage, the price of electricity, and the Q-values, the SG, in this instance, is made up of PV solar panels, SB, V2H units, and appliances. All the executed activities are depicted in Figure 6(b). First, the RL-HCPV checks the EV availability and SOC efficiencies. Thus, the V2H discharging mode is initiated when appliance power demand is high, and PV power output is low, typically between 7:00 and 9:00 PM when prices are also high. Thus, the current state is $s_t = \text{V2H discharge; low; high; expensive; expensive}$.

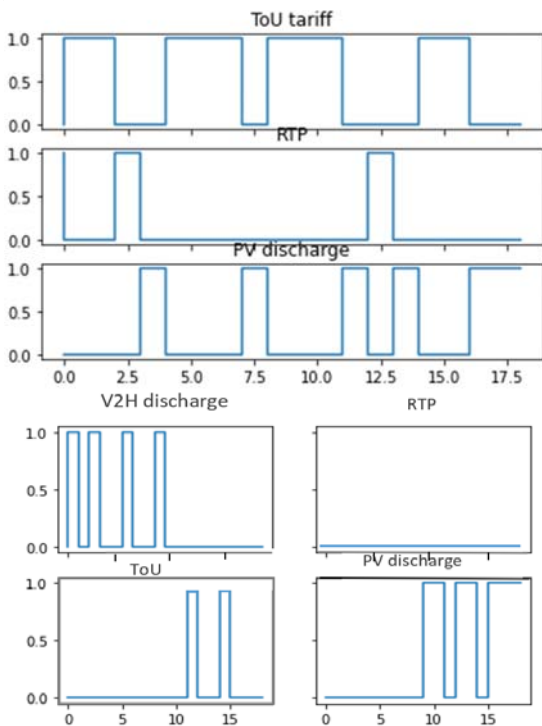
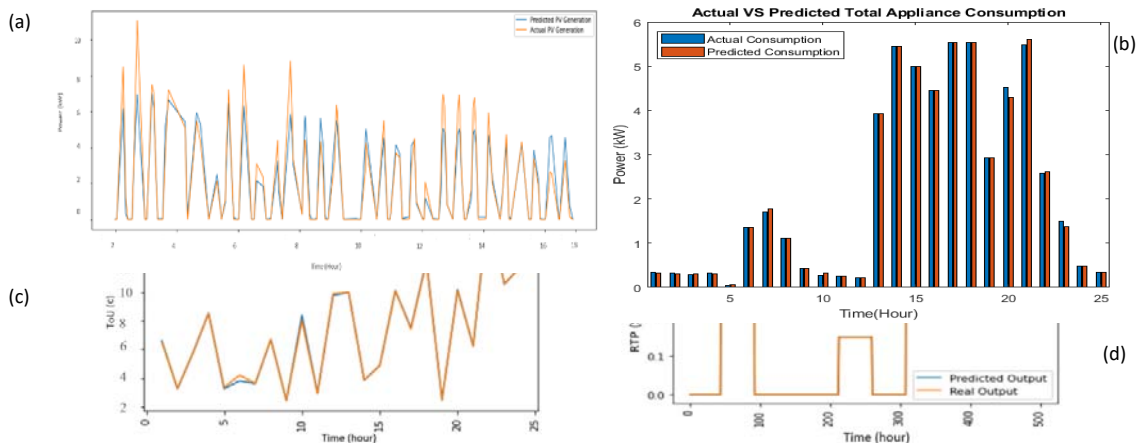


Figure 6: (a) Actions during Case 1 (b) Actions during Case 2

5.2 Performance of the Proposed RL-HCPV

5.2.1 Performance of DL algorithms

The performance of the LSTM model for PV generation is shown in Figure 7(a). In addition, the hourly prediction of appliances' consumption using ANN compared with actual consumption is shown in Figure 7(b). The multi-target regression for the hourly EV availability and SOC prediction is compared with the actual power generation, as shown in Figure 7(c and d). Finally, the performance of the ANN model for the price tariff is shown in Figure 7(e and f). The blue line in these graphs reflects the predicted data hourly, whereas the orange line represents the actual values. The actual and predicted data are similar, while slight inaccuracies can be observed in certain time intervals. In summary, the accuracy of the applied DL algorithms is listed in Table 6. As a result, the suggested DL algorithms for each uncertainty parameter generate precise and realistic predicted values, which can help in the decision-making process for DR management.



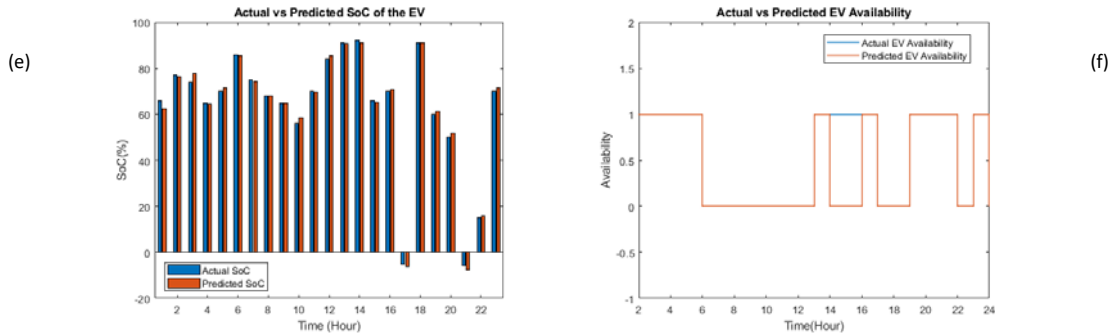


Figure 7: Actual and predicted (a) PV generation (b)Appliance consumption (c) EV SOC (d) EV availability (e) ToU pricing (f) RTP

Table 6: Accuracy of the DL algorithms

Uncertainty Parameters	DL algorithm	MAE	MSE
PV Generation	LSTMs	0.37	0.27
appliance consumption	ANN	0.21	0.22
EV SOC and availability	Multi-target Regression	0.31	0.15
ToU tariffs	ANN	0.37	0.32
RTP tariffs	ANN	0.0013	0.0066

5.2.2 Performance of Fuzzy Q-Learning

In RL, the Q-value is used to evaluate an action in a given state. The convergence can measure the learning rate of the Q-value. As a result, an RL algorithm is said to have converged when the learning curve becomes flat and no longer grows. The convergence of the Q-value for our FQL model is displayed in Figure 8. This graph shows that the agent converges with the highest Q-value. The Q-value is initially low because the agent makes bad action decisions, but it rises as the agent learns the behaviors through trial and error, eventually achieving the maximum Q-value.

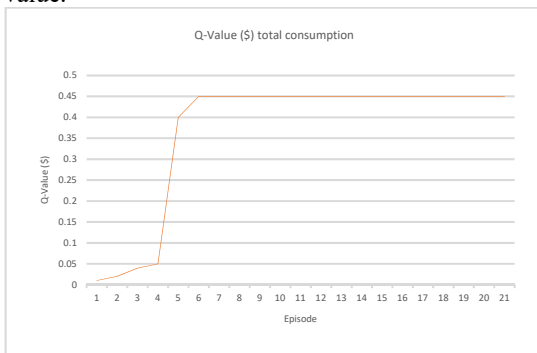


Figure 8: The convergence of Q-value

5.2.3 Performance comparison with and without V2H

While implementing the model, the main focus was minimizing costs and customer discomfort by ensuring sustainable power delivery. This is done by comparing the appliance consumption in various situations to the ToU and RTP tariffs.

The hourly appliance power consumption shown in Figure 9 illustrates the effectiveness of the proposed RL-HCPV algorithm. As displayed in this figure, when the V2H unit is not included, the dependence on grid prices grows from roughly seven hours on sunny days to 15 hours on cloudy days. In contrast, when the V2H unit is included, the dependence on grid prices on sunny and cloudy days is around three hours. The figure shows that when the V2H unit is incorporated, the suggested solution decreases reliance on grid costs by 38% on sunny days and 24% on cloudy days. Thus, integrating the V2H unit reduces reliance on the utility, indicating a significant reduction in electricity costs.

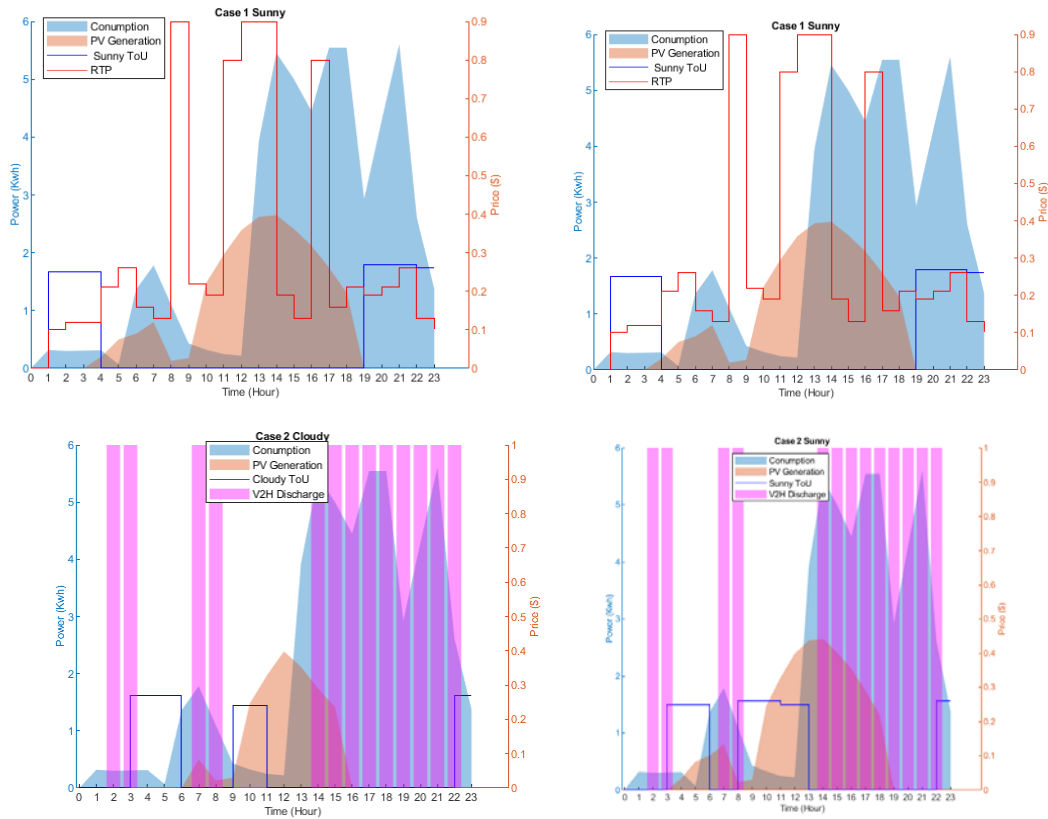


Figure 9: Dependence on ToU and RTP with or without V2H

5.2.4 Electricity bill analysis

According to the Saudi Electricity Company (SEC), the energy consumption is calculated using the following equation:

$$\frac{\text{load}}{30 \text{ days}} \times \text{actual days}$$

Table 7 compares the prices for a day of 14 KW consumption in four scenarios. The comparison shows that including the V2H unit can remove the penalties and significantly reduce the bill compared to other cases that do not include the V2H. Furthermore, when using RL, scheduling consumption when costs are the lowest significantly impacts RTP, resulting in lower costs than in other situations. Therefore, using both V2H and RL is the most effective way to lower the electricity bill.

Table 7: Comparison of electricity bills with or without V2H and RL

Elements	Electricity bill without V2H and RL	Electricity bill with V2H without RL	Electricity bill without V2H with RL	Electricity bill with V2H with RL
PV generation	10 kw	10 kw	10 kw	10 kw
Max. load	1.4 kw	1 kw	1.2 kw	0.8 kw

ToU tariff	00.18 SAR	00.18 SAR	00.18 SAR	00.18 SAR
RTP tariff	00.31-00.47 SAR	00.31-00.40 SAR	00.31-00.38 SAR	00.31-00.36 SAR
Penalty	0.05 SAR	NA	0.05 SAR	NA
Total electricity charge	34.8 SAR	17.4 SAR	25.2 SAR	14.7 SAR
	9.28 \$	4.64 \$	6.72 \$	3.92 \$

5.2.5 Analysis of computational capabilities

A mathematical HEMS model typically consists of discrete and continuous binary variables. The MILP problem related to HEMS was calculated using a number of heuristic algorithms, including Differential Evolution (DE), Binary Particle Swarm (BPS), Particle Swarm Optimization (PSO), Binary Learning Differential Evolution (BLDE), and Genetic Algorithm (GA). However, the computing demands of these approaches and algorithms are high and require tremendous processing power to meet the requirements for the best solution. As a result, the ideal convergence of these methodologies and algorithms is considered unstable.

Hence, this study utilized DL and RL to overcome these issues. Thus, Table 8 compares the obtained DL and RL results with the GA algorithm used to analyze our proposed model's productivity and effectiveness. The reduced cost statistics and average computational time of GA have been obtained for scheduled power optimizations. The result indicates that the RL generates noticeably better computational results in less time.

Table 8: Statistics comparison of simulated algorithms performance

Algorithm	Number of iterations (epochs)	Population size	Time
GA	1000 iterations	550	8ms/step
ANN	50 epochs	256	2ms/step
Fuzzy Q-learning	5000 iterations	210	3ms/step

6 Conclusion

This study proposed an RL-HCPV system for intelligently managing the DR of a smart home. The proposed model aims to make optimal decisions during on-peak and off-peak power supply to minimize costs and ensure sustainable electricity for the appliances. Several DL algorithms have been applied to the real-time data to deal with the uncertainty of PV generation, appliance power consumption, prices, EV availability, and SOC. DL algorithms have examined each uncertainty parameter to determine the most accurate model. Consequently, ANN is applied for the PV generation, ToU, and RTP tariffs. A multi-target regression is applied to predict EV availability and SOC. In addition, LSTM is applied to predict appliance energy consumption. Finally, the predicted data are transferred into the RL algorithm, the Fuzzy Q-learning model, for efficient DR decision-making. Moreover, simulations demonstrate that when the V2H unit is included, the proposed RL-HCPV algorithm can minimize the reliance on grid prices by 38% on sunny days and 24% on cloudy days.

Future analyses should consider consumer-friendly actions like selling surplus electricity to the grid to boost user revenue. In addition, Smart MGs require an autonomous sustainable power supply plan for household appliances.

Author Contributions: Ohoud Amughran planned the presented idea and wrote the manuscript with support from Sami Ben Slama and Bassam Zafar who contributed to the original draft preparation, supervised the findings and final revisions.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] J. Corral-García, F. Lemus-Prieto, J.-L. González-Sánchez, and M.-Á. Pérez-Toledano, "Analysis of Energy Consumption and Optimization Techniques for Writing Energy-Efficient Code," *Electronics (Basel)*, vol. 8, no. 10, p. 1192, 2019.
- [2] M. S. Aliero, M. Asif, I. Ghani, M. F. Pasha, and S. R. Jeong, "Systematic Review Analysis on Smart Building: Challenges and Opportunities," *Sustainability*, vol. 14, no. 5, p. 3009, 2022.
- [3] G. Hafeez et al., "Efficient Energy Management of IoT-Enabled Smart Homes Under Price-Based Demand Response Program in Smart Grid," *Sensors*, vol. 20, no. 11, p. 3155, 2020.
- [4] J. S. GK, "MANFIS based SMART home energy management system to support SMART grid.," *Peer Peer Netw Appl*, vol. 13, no. 6, 2020.
- [5] B. Sharma and J. K. Maherchandani, "A Review on Integration of Electric Vehicle in Smart Grid: Operational modes, Issues and Challenges," in 2022 International Conference on Computer Communication and Informatics (ICCCI), 2022, pp. 1–5.
- [6] N. S. Pearre and H. Ribberink, "Review of research on V2X technologies, strategies, and operations," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 61–70, 2019.
- [7] S.-H. Park, A. Hussain, and H.-M. Kim, "Impact analysis of survivability-oriented demand response on islanded operation of networked microgrids with high penetration of renewables," *Energies (Basel)*, vol. 12, no. 3, p. 452, 2019.
- [8] F. Alfaverh, M. Denai, and Y. Sun, "Electrical vehicle grid integration for demand response in distribution networks using reinforcement learning," *IET Electrical Systems in Transportation*, vol. 11, no. 4, pp. 348–361, 2021.
- [9] U. ur Rehman, K. Yaqoob, and M. A. Khan, "Optimal power management framework for smart homes using electric vehicles and energy storage," *International Journal of Electrical Power & Energy Systems*, vol. 134, p. 107358, 2022.
- [10] M. H. Elkholy, H. Metwally, M. A. Farahat, M. Nasser, T. Senjyu, and M. E. Lotfy, "Dynamic centralized control and intelligent load management system of a remote residential building with V2H technology," *J Energy Storage*, vol. 52, p. 104839, 2022.
- [11] G. S. Georgiou, P. Christodoulides, and S. A. Kalogirou, "Optimizing the energy storage schedule of a battery in a PV grid-connected nZEB using linear programming," *Energy*, vol. 208, p. 118177, 2020.
- [12] S. S. R. Nasab, A. T. Nasrabadi, S. Asadi, and S. A. H. S. Taghia, "Investigating the probability of designing net-zero energy buildings with consideration of electric vehicles and renewable energy," *Engineering, Construction and Architectural Management*, 2021.
- [13] S. Rout and G. R. Biswal, "Bidirectional EV integration in home load energy management using swarm intelligence," *International Journal of Ambient Energy*, pp. 1–8, 2022.
- [14] A. Dileep and B. M. Beena, "Global Distribution and Price Prediction of Electric Vehicles Using Machine Learning," in 2022 2nd International Conference on Intelligent Technologies (CONIT), 2022, pp. 1–6.
- [15] D. Aguilar-Dominguez, J. Ejeh, A. D. F. Dunbar, and S. F. Brown, "Machine learning approach for electric vehicle availability forecast to provide vehicle-to-home services," *Energy Reports*, vol. 7, pp. 71–80, 2021.
- [16] R. Khezri and D. Steen, "Vehicle to Everything (V2X)-A Survey on Standards and Operational Strategies," in 2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), 2022, pp. 1–6.
- [17] X. Xu, Y. Jia, Y. Xu, Z. Xu, S. Chai, and C. S. Lai, "A multi-agent reinforcement learning-based data-driven method for home energy management," *IEEE Trans Smart Grid*, vol. 11, no. 4, pp. 3201–3211, 2020.
- [18] US Energy Information Administration, "Independent Statistics and Analysis," 2021. <https://www.eia.gov/totalenergy/data/annual/>
- [19] N. S. Pearre and H. Ribberink, "Review of research on V2X technologies, strategies, and operations," *Renewable and Sustainable Energy Reviews*, vol. 105, pp. 61–70, 2019.
- [20] C. Gschwendtner, S. R. Sinsel, and A. Stephan, "Vehicle-to-X (V2X) implementation: An overview of predominate trial configurations and technical, social and regulatory challenges," *Renewable and Sustainable Energy Reviews*, vol. 145, p. 110977, 2021.
- [21] A. Ahmadian, B. Mohammadi-Ivatloo, and A. Elkamel, "A review on plug-in electric vehicles: introduction, current status, and load modeling techniques," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 3, pp. 412–425, 2020.
- [22] S. Vadi, R. Bayindir, A. M. Colak, and E. Hossain, "A review on communication standards and charging topologies of V2G and V2H operation strategies," *Energies (Basel)*, vol. 12, no. 19, p. 3748, 2019.
- [23] A. Chatzivasileiadi, E. Ampatzi, and I. P. Knight, "Electrical energy storage sizing and space requirements for sub-daily autonomy in residential buildings," *Energies (Basel)*, vol. 15, no. 3, p. 1145, 2022.

- [24] E. Solomon, I. G. Hagos, and B. Khan, "Outage management in a smart distribution grid integrated with PEV and its V2B applications," in *Active Electrical Distribution Network*, Elsevier, 2022, pp. 217–226.
- [25] S. Das et al., "Toyota smarhome: Real-world activities of daily living," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 833–842.
- [26] B. Cui, S. Lee, P. Im, M. Koenig, and M. Bhandari, "High resolution dataset from a net-zero home that demonstrates zero-carbon living and transportation capacity," *Data Brief*, vol. 45, p. 108703, 2022.
- [27] H. Mehrjerdi, "Resilience oriented vehicle-to-home operation based on battery swapping mechanism," *Energy*, vol. 218, p. 119528, 2021.
- [28] M. S. Shamami et al., "Artificial Intelligence-Based Performance Optimization of Electric Vehicle-to-Home (V2H) Energy Management System," *SAE International Journal of Sustainable Transportation, Energy, Environment, & Policy*, vol. 1, no. 13-01-02–0007, 2020.
- [29] S. A. ben Slama, "Design and implementation of home energy management system using vehicle to home (H2V) approach," *J Clean Prod*, p. 127792, 2021.
- [30] R. S. Sundararajan and M. T. Iqbal, "Dynamic Modelling of a Solar Energy System with Vehicle to Home Option for Newfoundland Conditions," *European Journal of Engineering and Technology Research*, vol. 6, no. 5, pp. 16–23, 2021.
- [31] T. Higashitani, T. Ikegami, A. Uemichi, and A. Akisawa, "Evaluation of residential power supply by photovoltaics and electric vehicles," *Renew Energy*, vol. 178, pp. 745–756, 2021.
- [32] H. Nakano et al., "Aggregation of V2H systems to participate in regulation market," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 668–680, 2020.
- [33] M. Kosinka, Z. Slanina, M. Petruzela, and V. Blazek, "V2H control system software analysis and design," in *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, 2020, pp. 972–977.
- [34] O. Ouramdane, E. Elbouchikhi, Y. Amirat, F. le Gall, and E. Sedgh Gooya, "Home Energy Management Considering Renewable Resources, Energy Storage, and an Electric Vehicle as a Backup," *Energies (Basel)*, vol. 15, no. 8, p. 2830, 2022.
- [35] O. Vermesan et al., "Automotive intelligence embedded in electric connected autonomous and shared vehicles technology for sustainable green mobility," *Frontiers in Future Transportation*, vol. 2, p. 688482, 2021.
- [36] S. Shahriar, A. Al-Ali, A. Osman, S. Dhau, and M. Nijim, "Machine Learning Approaches for EV Charging Behavior: A Review," *IEEE Access*, vol. 8, pp. 168980–168993, Dec. 2020, doi: 10.1109/ACCESS.2020.3023388.
- [37] Y. Yang and S. Wang, "Resilient residential energy management with vehicle-to-home and photovoltaic uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 132, p. 107206, 2021.
- [38] D. Borge-Diez, D. Icaza, E. Açikkalp, and H. Amaris, "Combined vehicle to building (V2B) and vehicle to home (V2H) strategy to increase electric vehicle market share," *Energy*, vol. 237, p. 121608, 2021.
- [39] A. Mohammad et al., "Integration of electric vehicles and energy storage system in home energy management system with home to grid capability," *Energies (Basel)*, vol. 14, no. 24, p. 8557, 2021.
- [40] Y. Ye, D. Qiu, H. Wang, Y. Tang, and G. Strbac, "Real-time autonomous residential demand response management based on twin delayed deep deterministic policy gradient learning," *Energies (Basel)*, vol. 14, no. 3, p. 531, 2021.
- [41] R. Kelm Pawel and I. Wasiak, "Energy Management in a Prosumer Installation Using Hybrid Systems Combining EV and Stationary Storages and Renewable Power Sources," *Applied Sciences*, vol. 11, no. 11, p. 5003, 2021.
- [42] S. Rafique, M. J. Hossain, M. S. H. Nizami, U. bin Irshad, and S. C. Mukhopadhyay, "Energy management systems for residential buildings with electric vehicles and distributed energy resources," *IEEE Access*, vol. 9, pp. 46997–47007, 2021.
- [43] A. T. Eseye, M. Lehtonen, T. Tukia, S. Uimonen, and R. J. Millar, "Optimal energy trading for renewable energy integrated building microgrids containing electric vehicles and energy storage batteries," *IEEE Access*, vol. 7, pp. 106092–106101, 2019.
- [44] S. Lee, H. Jin, L. F. Vecchietti, J. Hong, and D. Har, "Short-term predictive power management of PV-powered nanogrids," *Ieee Access*, vol. 8, pp. 147839–147857, 2020.
- [45] M. Yousefi, A. Hajizadeh, M. N. Soltani, and B. Hredzak, "Predictive home energy management system with photovoltaic array, heat pump, and plug-in electric vehicle," *IEEE Trans Industr Inform*, vol. 17, no. 1, pp. 430–440, 2020.
- [46] A. C. Duman, H. S. Erden, Ö. Gönül, and Ö. Güler, "A home energy management system with an integrated smart thermostat for demand response in smart grids," *Sustain Cities Soc*, vol. 65, p. 102639, 2021.
- [47] X. Liu, Y. Wu, and H. Wu, "PV-EV Integrated Home Energy Management Considering Residential Occupant Behaviors," *Sustainability*, vol. 13, no. 24, p. 13826, 2021.
- [48] A. T. Tran, A. Kawashima, S. Inagaki, and T. Suzuki, "Design of a Home Energy Management System Integrated with an Electric Vehicle (V2H+ HPWH EMS)," in *Design and Analysis of Distributed Energy Management Systems*, Springer, 2020, pp. 47–66.
- [49] S. Lee and D.-H. Choi, "Reinforcement learning-based energy management of smart home with rooftop solar photovoltaic system, energy storage system, and home appliances," *Sensors*, vol. 19, no. 18, p. 3937, 2019.
- [50] N. Ashenov, M. Myrzaliyeva, M. Mussakhanova, and H. K. Nunna, "Dynamic Cloud and ANN based Home Energy Management System for End-Users with Smart-Plugs and PV Generation," in *2021 IEEE Texas Power and Energy Conference (TPEC)*, 2021, pp. 1–6.
- [51] M. Kim, H. Kim, and J. H. Jung, "A Study of Developing a Prediction Equation of Electricity Energy Output via Photovoltaic Modules," *Energies (Basel)*, vol. 14, no. 5, p. 1503, 2021.
- [52] A. Vulkan, I. Kloog, M. Dorman, and E. Erell, "Modeling the potential for PV installation in residential buildings in dense urban areas," *Energy Build*, vol. 169, pp. 97–109, 2018.
- [53] H. A. Muqet et al., "IoT-based intelligent source-load-storage coordination scheme for prosumer campus microgrids," *Front Energy Res*, vol. 10, 2022.
- [54] A. Khalid, N. Javaid, A. Mateen, M. Ilahi, T. Saba, and A. Rehman, "Enhanced time-of-use electricity price rate using game theory," *Electronics (Basel)*, vol. 8, no. 1, p. 48, 2019.
- [55] S. Aslam, "An optimal home energy management scheme considering grid connected microgrids with day-ahead weather forecasting using artificial neural network," *Masters Thesis*, 2018.
- [56] M. Rafiei, T. Niknam, J. Aghaei, M. Shafie-Khah, and J. P. S. Catalão, "Probabilistic load forecasting using an improved wavelet neural network trained by generalized extreme learning machine," *IEEE Trans Smart Grid*, vol. 9, no. 6, pp. 6961–6971, 2018.
- [57] S. Chai, Z. Xu, and Y. Jia, "Conditional density forecast of electricity price based on ensemble ELM and logistic EMOS," *IEEE Trans Smart Grid*, vol. 10, no. 3, pp. 3031–3043, 2018.
- [58] W. Fu, K. Wang, C. Li, and J. Tan, "Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM," *Energy Convers Manag*, vol. 187, pp. 356–377, 2019.
- [59] H. R. Berenji, "Fuzzy Q-learning: a new approach for fuzzy dynamic programming," in *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference*, 1994, pp. 486–491.
- [60] H. R. Berenji, "Fuzzy Q-learning: a new approach for fuzzy dynamic programming," in *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference*, 1994, pp. 486–491.