International Journal of Innovative Research and Scientific Studies, 8(3) 2025, pages: 46-58

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: **www.ijirss.com**

Optimal sizing of the PV-BESS energy system for off-grid electric vehicle charging station using deep reinforcement learning techniques

Kedsara Palachai[1], Mongkonchai Thongpiam[2], Boonruang Marungsri[3], Terapong Boonraksa[4], Promphak Boonraksa[5*]

[1]*Department of Electrical Engineering, Faculty of Engineering and Architecture, Rajamangala University of Technology Suvarnabhumi, Nonthaburi, Thailand.*
[2]*Provincial Electricity Authority's (PEA), Thailand.*
[3]*School of Electrical Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima, Thailand.*
[4]*School of Electrical Engineering, Rajamangala University of Technology Rattanakosin, Nakhon Pathom, Thailand.*
[5]*Department of Mechatronics Engineering, Faculty of Engineering and Architecture, Rajamangala University of Technology Suvarnabhumi, Nonthaburi, Thailand.*

Corresponding author: Promphak Boonraksa (*Email: promphak.b@rmutsb.ac.th*)

Abstract

This research applies Deep Reinforcement Learning (DRL) techniques to determine the optimal sizing of PV systems and BESS for off-grid EV charging stations by comparing the performance of the PPO, A2C, and DQN algorithms. The study found that the PPO technique yielded the best results, reducing total costs and maximizing energy efficiency, thereby reducing electricity costs by 48.79%. In terms of economics, the project is investment-worthy with an NPV of 9.7 million baht, an IRR of 20.89%, a BCR of 1.804, and a payback period of 5 years. Environmentally, the system can reduce CO_2 emissions by up to 1,383 tons over the 20-year project lifespan. The developed off-grid system helps reduce dependence on fossil fuels, enhances energy security, and promotes sustainable energy use. Therefore, the PPO technique is the most suitable approach for sizing energy production systems and evaluating the viability of off-grid EV charging stations.

Keywords: Deep reinforcement learning, Hybrid energy system, Off-grid EV charging station.

DOI: 10.53894/ijirss.v8i3.6430
Funding: This study received no specific financial support.
History: Received: 6 March 2025 / **Revised:** 9 April 2025 / **Accepted:** 11 April 2025 / **Published:** 24 April 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception, methodology, validation, review and editing, formal analysis, and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Acknowledgment: The authors would like to thank the Rajamangala University of Technology Suvarnabhumi, Suranaree University of Technology, and Rajamangala University of Technology Rattanakosin, Thailand, for their technical and financial support. Thank you to the PEA, Thailand, for providing information on the EV charging stations.
Publisher: Innovative Research Publishing

1. Introduction

The transportation sector is a major contributor to air pollution and greenhouse gas (GHG) emissions, primarily due to the widespread use of internal combustion engine (ICE) vehicles that rely on petroleum fuels. These vehicles are among the largest sources of harmful emissions, releasing significant amounts of air pollutants and GHGs into the atmosphere. As a result, the transportation sector is the second-largest global contributor to GHG emissions, trailing only electricity generation. It accounts for 23% of global carbon dioxide (CO_2) emissions from energy use and 14% of total GHG emissions. Within the sector, road transportation, including passenger vehicles and freight trucks, is the dominant source, responsible for 73% of transportation-related CO_2 emissions. Addressing these emissions is crucial for mitigating climate change and improving air quality [1, 2].

As a result, many countries, including Thailand, have been actively promoting the use of electric vehicles (EVs) as a sustainable alternative to traditional internal combustion engine vehicles. EVs offer a way to reduce reliance on fossil fuels and decrease CO_2 emissions, contributing to cleaner air and a lower carbon footprint [3]. This has driven a significant rise in the adoption of electric vehicles, as illustrated in Figure 1; however, the growing number of EVs in the transportation system poses challenges for the power grid, particularly in managing energy demand during peak charging periods. Without proper planning, the increased load from EV charging could adversely affect the electrical system's performance and lead to grid overloads. This highlights the need for innovative energy management solutions to ensure a stable and efficient power supply. Therefore, the development of an energy system that can efficiently support electric vehicle charging at standalone charging stations is essential.

The adoption of the EV is being actively encouraged in many countries, including Thailand, as a sustainable alternative to reduce reliance on fossil fuels and lower CO_2 emissions [2]. This shift has led to a noticeable increase in the use of electric vehicles. However, the growing number of EVs in the transportation system poses challenges for the electricity grid, especially during peak charging times. The surge in energy demand from EV charging can strain the grid, potentially causing overloading and negatively impacting the overall performance of the electrical grid system [4]. Developing efficient energy systems for off-grid EV charging stations is essential to tackle this issue. These systems must ensure stability, reliability, and the ability to meet the growing demand for clean transportation. An off-grid EV charging station operates independently of the main power grid, relying on renewable energy sources such as photovoltaic (PV) panels and energy storage systems (ESS), including batteries, to supply power [5]. However, designing an optimal hybrid energy system (HES) for such stations presents a significant challenge. It requires careful consideration of multiple factors, including the energy demand of the charging station, solar panel efficiency, geographical and climatic conditions, as well as cost and investment feasibility. Addressing these factors through detailed research is crucial to creating sustainable and effective off-grid charging solutions [6-8].

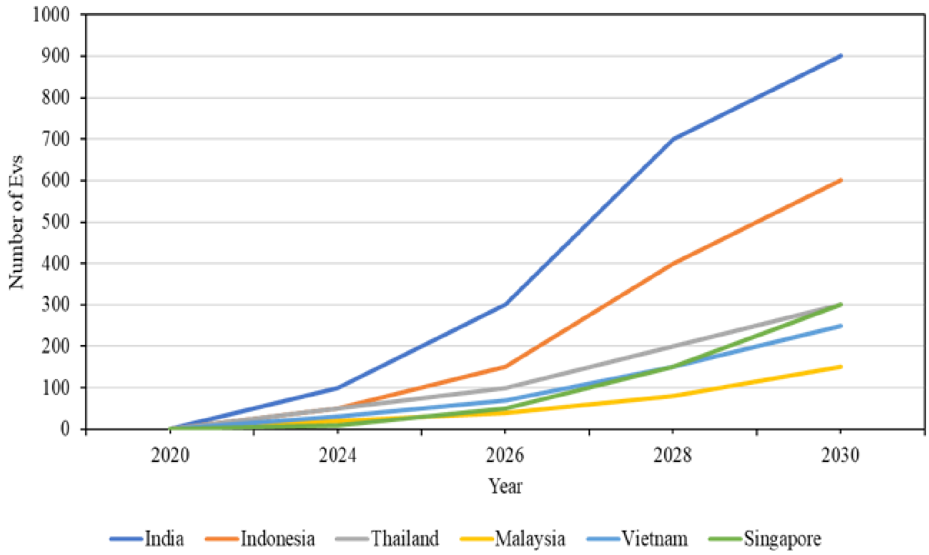

Figure 1.
The Growth of Electric Vehicles [7].

Artificial intelligence technology plays a significant role in analyzing and determining the optimal size for renewable energy installations. In particular, deep learning, a branch of AI, has been widely applied in various research projects. Several studies have explored using DRL techniques to develop systems that adapt to environmental conditions and user behavior and enhance energy management efficiency [9-13]. For example, one study examined energy management at EV charging stations using approximate dynamic programming techniques to optimize energy use from renewable sources and reduce the management cost burden at electric vehicle charging stations. On another front, research studies the energy management of electric vehicle charging stations using ADP techniques to optimize energy use from renewable sources and reduce the burden of energy management costs at electric vehicle charging stations [14, 15]. On another front, it employs DRL techniques to balance charging and energy discharge in lithium-ion (Li-ion) batteries using DQL for model training, which enhances the efficiency of charge balancing and reduces balancing time, resulting in improved battery system performance [16].

Additionally, DRL techniques develop policies for EV charging stations, considering user flexibility and Time-of-Use (TOU) electricity rates, which can help reduce users' electricity costs by more than 20% and increase charging flexibility [17]. Moreover, DRL techniques, specifically Proximal Policy Optimization (PPO), have been employed to develop an adaptive EV charging system that responds to environmental factors and user behavior, thereby reducing charging costs and enhancing user convenience [18]. DRL techniques were used to develop a smart grid system that can improve energy management and Demand Response in the energy system, utilizing PPO to find suitable policies for energy management and Demand Response, which helps enhance the stability and efficiency of the Smart Grid [19]. Deep learning was applied to predict solar energy generation and EV charging demands while optimizing battery storage and charging plans. Simple RNNs effectively forecast PV performance, while bidirectional LSTMs are well-suited for EV load predictions [20]. The integration of solar and wind power into EV charging stations using DOA and SBNN optimization improves energy efficiency and reduces harmonic distortion [21]. A framework optimizes PV and EV sizing using SCSB, improving load regulation. The proposed model integrates EV charging with PV and storage to minimize costs, showing significant savings through optimized investments [22]. There are domestic EV charging challenges while promoting clean solar energy consumption. Using DRL, it shifts EV loads to peak solar generation times, leveraging real-time pricing and historical data for optimal flexibility [23]. This research applies DRL algorithms in complex processes, featuring algorithms such as SARSA, Q-Learning, DQN, PPO, and SAC, and carefully considers reward functions and appropriate states or actions for each situation. These algorithms can be applied to solve various problems in renewable energy. So, the contribution of this paper as follows:

- This research leverages DRL techniques-PPO, A2C, and DQN- to optimize the sizing of photovoltaic (PV) and battery energy storage systems (BESS) for off-grid EV charging stations. All three DRL techniques were compared for their optimization results. The findings indicate that the PPO algorithm delivers the most cost-effective and energy-efficient solution.
- This study analyzed economic feasibility and sustainability impact, demonstrating the economic viability of off-grid EV charging stations powered by hybrid energy systems.
- By utilizing the most suitable hybrid energy system from an operator's perspective, this research provides a data-driven decision-making framework for investments in off-grid EV charging stations.

2. Off-Grid Electric Vehicle Charging Station

Off-grid electric vehicle charging stations are self-sufficient power systems that utilize Electric Vehicle Supply Equipment (EVSE) to charge EVs without relying on the power grid [24]. Figure 2 illustrates the layout of the off-grid solar charging station. It consists of a PV system that receives energy from sunlight and converts it into DC electricity. In order to have energy to use at night or during low sunlight and to allow the electric vehicle charging station to work continuously, there must be a battery energy storage system.

Off-grid EV charging stations offer several advantages, such as not relying on the electrical grid, enabling independent operation, reducing the risks of grid disruptions, and increasing service reliability. Additionally, these stations use clean energy from solar panels and energy storage systems, which helps reduce greenhouse gas emissions and promotes the use of environmentally friendly energy. Furthermore, they help reduce long-term costs since there is no dependence on grid electricity. However, off-grid charging stations also have some drawbacks, including high installation costs due to the use of solar panels and high-performance energy storage systems needed for reliability. There are also issues with energy fluctuation since the efficiency of solar panels depends on weather conditions [25].

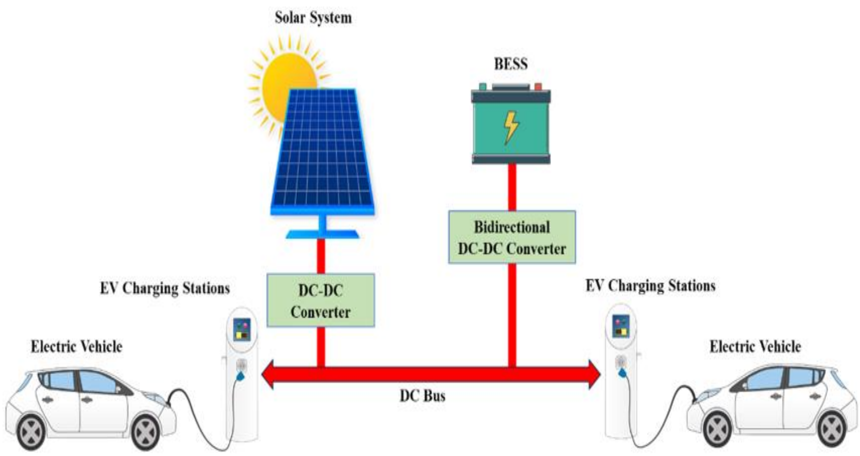

Figure 2.
The Growth of Electric Vehicles [7].

3. Methodology

3.1. Vehicle Charging Station: A Case Study

This research study investigates the charging load of the electric vehicle charging station of the Provincial Electricity Authority (PEA) in Sakon Nakhon Province, Thailand, with geographical coordinates 17.1595°N and 104.14321°E. The

Global Solar Atlas (GSA) tool was used to identify related factors such as light intensity and temperature for simulation data. The daily load curve of the EV charging station in a case study is shown in Figure 3.

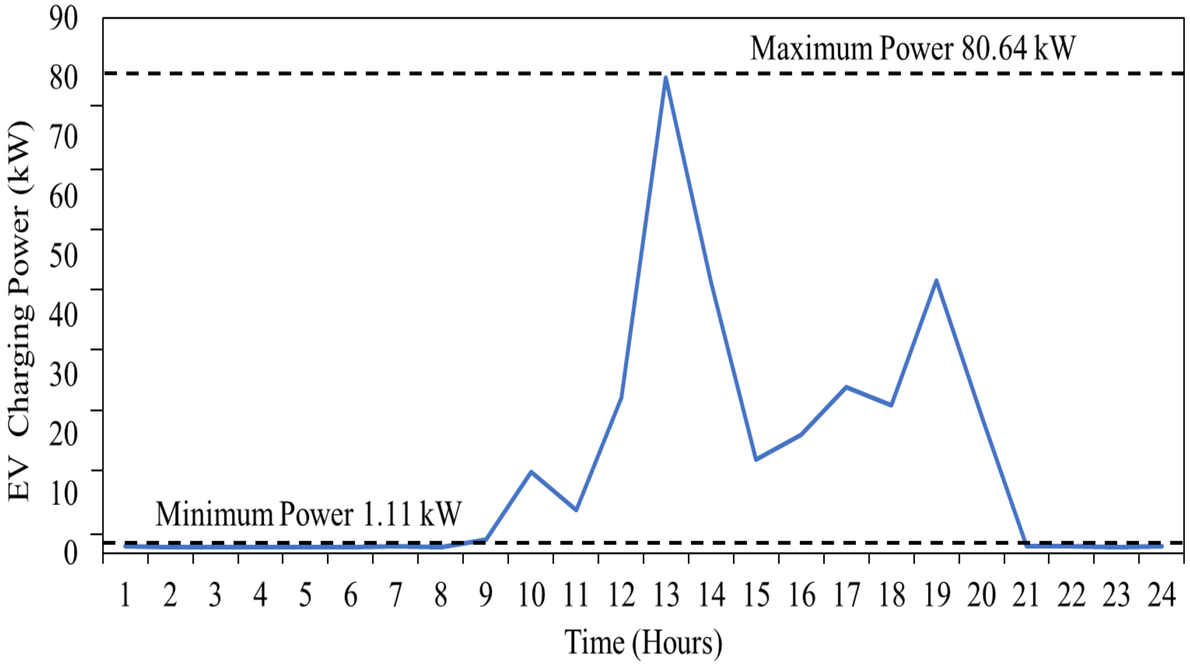

Figure 3.
Load curve of the EV charging station.

From Figure 3, the minimum power appears between approximately 8 o'clock, with the lowest load value at 1.1 kW, and the maximum power occurs around 1 PM, with the highest electric power value at 80.64 kW. This period has the highest usage of electric vehicle charging services. The electricity charges can be calculated in Table 1.

Table 1.
Electricity Charge Calculation for 20 Years [26].

Item	On Peak	Off Peak
Peak Power Demand (kW)	80.64	11.29
Energy Consumption (kWh)	10,078.50	338.70
Monthly Electricity Cost (THB)	73,292.85	
Annual Electricity Cost (THB)	879,514.20	
20-Year Electricity Cost (THB)	17,590,284	

From Table 1, the calculation of electricity costs for EV charging stations shows that the On-Peak power demand is 80.64 kW and the Off-Peak is 11.29 kW, with the total energy consumption during On-Peak at 10,078.50 kWh and during Off-Peak at 338.70 kWh. For the automatic electricity tariff adjustment (Ft), the Ft rate is 0.3672 Baht/unit, and the value-added tax (VAT) is 7%. The monthly electricity bill, including the base electricity cost, Ft, and VAT, is 73,292.85 Baht. The annual electricity cost, including VAT, will be 879,514.20 Baht, which results in a total electricity cost over 20 years of 17,590,284 Baht.

3.2. Hybrid Energy System

A HES is a system that integrates multiple energy sources to improve the stability and efficiency of electricity generation. Typically, HESs incorporate renewable energy sources, such as PV solar and wind energy, combined with ESS or diesel generators. This study focuses on utilizing solar PV and ESS. The present work focuses on determining the optimal hybrid energy configuration for an off-grid electric vehicle (EV) charging station, as depicted in Figure 4. This involves finding the most suitable combination of energy sources to efficiently meet the station's power demands while ensuring sustainability and reliability.

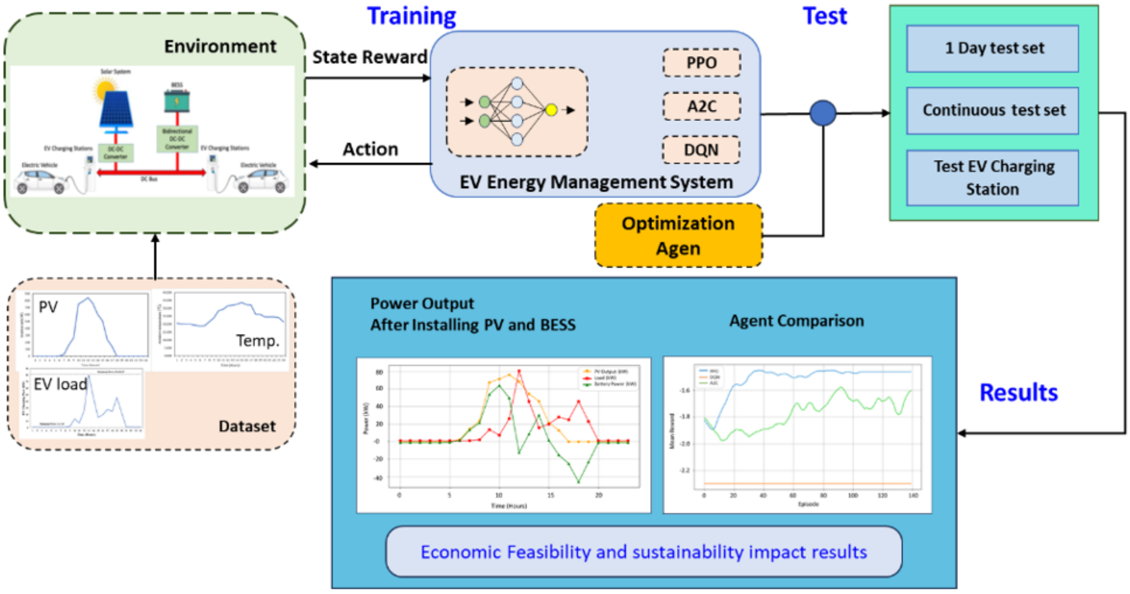

Figure 4.
Schematic of the Proposed Approach.

In designing and calculating the energy produced by a solar energy system, it is necessary to use equations that assist in accurately calculating the amount of energy solar panels can produce. The energy production calculation is determined by Equation 1.

$$P_{PV} = Y_{PV} \times f_{PV} \times I_r \times S_C T[1 + \alpha p(T_C - T_{C,s})] \quad (1)$$

where αp is referred to as the temperature coefficient of the solar cell panel, which indicates how the panel's efficiency decreases as the temperature increases; generally, this value ranges from -0.004 to -0.05 (K^{-1}), depending on the type of solar cell panel used. $T_{C,s}$ is the temperature at which the solar panel operates most efficiently, typically around 25°C. Y_{PV} represents the energy output of a solar panel under specific illumination conditions, such as 200 W/m² or 300 W/m², depending on the panel type. Depending on the system type and panel installation design, the loss-related factor f_{PV} varies between 0.85 and 0.9. Irradiance I_r is determined by the location of the solar power system, with an average of approximately 1,000 W/m² under good sunlight conditions [27].

The mathematical model in the energy management system of an off-grid EV charging station with PV solar panels and a BESS accounts for energy production and consumption from multiple sources during the PV generation period. Effective energy management must allocate energy appropriately between the EVs, BESS, and charging station loads. This can be achieved using the energy balance equation, which governs the charging station's operation by managing the energy from PV solar panels stored in the BESS and regulating the energy within the BESS, as shown in Equation 2.

$$P_{BESS} = P_{PV} - P_{EV} \quad (2)$$

Where P_{EV} is the electric vehicle's charging power, P_{PV} is the PV power produced, and P_{BESS} is the power of the BESS. Since BESS must be operate within safe limits, Equation 3 shows the limits for charging and discharging power of BESS.

$$P_{BESS\,\min} \leq P_{BESS} \leq P_{BESS\,\max} \quad (3)$$

Where $P_{BESSmin}$ is the minimum power discharge limit from BESS, and $P_{BESSmax}$ is the maximum power charging limit from BESS. The state of charge (SOC) of the BESS battery should not be lower than 20% and should not exceed 90% to reduce battery degradation, as stated in Equation 4.

$$SOC_{BESS\,\min} \leq SOC_{BESS} \leq SOC_{BESS\,\max} \quad (4)$$

Where $SOC_{BESSmin}$ is BESS's minimum safe charging status, and $SOC_{BESSmax}$ is BESS's maximum safe charging status. Equation 5 is used to limit the battery's charging status in EV to prevent battery deterioration.

$$SOC_{EV\,\min} \leq SOC_{EV(t)} \leq SOC_{EV\,\max} \quad (5)$$

Where SOC_{EVmin} is the minimum charging state of the battery at 10% in EV, and SOC_{EVmax} is the maximum charging state of the battery in EV at 90%. In the simulation of calculating costs and expenses related to installing and maintaining the solar power system PV and batteries, including selling electricity to the load, various details are provided in Table 2.

Table 2.
Simulation parameters of the investment cost on off-grid EV charging station systems [28-30].

Parameters	Values
PV Installation Cost	30,707.44 THB/kW
PV Maintenance Cost	170 Baht/kW/year
Battery Installation Cost	40,800 THB/kW
Battery Energy Cost	20,400 THB/kWh
Battery Maintenance Cost	672.18 THB/kW/year

3.3. Reduction of CO_2 Emissions

Reducing CO_2 emissions is a key issue in climate change [30]. Generating electricity from solar energy is an effective way to reduce CO_2 emissions, as the energy produced from this renewable energy source can replace the use of energy from fossil fuels, a major source of CO_2 emissions [31]. The Emission Reduction (ER, in kg CO_2) is determined by multiplying the annual electricity activity by the Emission Factor, as shown in Equation 6.

$$ER = A \times EF \tag{6}$$

The annual electricity activity (*A*) (in kWh) calculated from the previous steps can be obtained from Equation 7. The *A* was calculated using the total capacity (*C*) (in kW) of the system and the total operational hours in a year (*H*) (in hours). The Emission Factor (*EF*) (in kg CO_2/kWh) of the electricity grid, which represents the amount of CO_2 emitted per unit of electricity generated, is based on data from the Electricity Generating Authority of Thailand (EGAT). The Emission Factor for Thailand is approximately 0.5124 kg CO_2/kWh [30].

$$A = C \times H \tag{7}$$

3.4. Charging Station Benefit

In calculating the profit from a charging station for an off-grid system with a long-term investment of 20 years, the following equations can be used to calculate net profit and capital control.

3.4.1. Net Present Value

Net Present Value (NPV) is the difference between the present value of money saved from energy costs in monetary terms that are expected to be received each year over the life of the project and the present value of money that must be spent on the project being considered, calculated at a predetermined discount rate or cost of capital [32]. The *NPV* calculation formula can be obtained from Equation 8.

$$NPV = \sum_{t=0}^{n} \frac{ES_t}{(1+i)} - I \tag{8}$$

Where ES_t is the energy savings in year t, I is the Initial investment, i is the discount rate, n is the project lifespan. If the *NPV* is greater than zero, the project has investment potential. If the *NPV* is higher compared to other projects, it indicates the project is attractive for investment. However, *NPV* alone may have limitations when comparing projects of different sizes with the same *NPV*. Therefore, other evaluation tools should be used alongside NPV for better decision-making accuracy.

3.4.2. Internal Rate of Return

Internal Rate of Return (IRR) is the discount rate that makes the present value of the expected cash flows to be paid for the investment equal to the present value of the expected cash flows to be received from energy savings over the project's lifespan [31], which can be calculated by Equation 9.

$$0 = \sum_{t=0}^{n} \frac{ESt_t}{(1+i)^t} - I \tag{9}$$

If the IRR is greater than or equal to the discount rate (interest rate) set by the investor sets, the project is worth investing in. Although NPV and IRR often lead to the same decision, differences in assumptions, such as reinvesting received cash flows or using different depreciation methods (e.g., double-declining balance vs. straight-line method), may lead to different outcomes. Therefore, it is important to consider the assumptions used in the calculations to improve the accuracy of the analysis [31].

3.4.3. Benefit-Cost Ratio

Benefit-Cost Ratio (BCR) is the ratio between the net present value of the expected cash flows over the life of the project and the initial investment. It represents a comparison between the returns in the form of income adjusted to present value throughout the project's lifespan and the initial investment incurred at present. A project with a BCR is more significant than one that indicates it is suitable for investment. The calculation formula for the internal rate of return is in Equation 10 [31].

$$BCR = \frac{\sum_{t=1}^{n} \frac{ES_t}{(1+i)^t}}{I_0} \tag{10}$$

If the BCR is greater than 1, the project is financially viable, as the present value of returns from energy savings exceeds the initial investment. A project with a BCR more significant than 1 is considered attractive and suitable for investment.

3.5. Proximal Policy Optimization

The PPO is a deep reinforcement learning algorithm that has gained significant attention due to its ability to optimize an agent's policy effectively while maintaining stability and efficiency. The PPO is classified as a model-free algorithm and combines the benefits of value-based methods with policy gradient approaches. It directly optimizes the agent's policy by adjusting. A key feature of PPO is its ability to balance exploration and exploitation. This is achieved through a surrogate objective function, which constrains policy updates to prevent them from being too large. Such a mechanism helps ensure that the learning process avoids drastic shifts in the agent's behavior that could destabilize training. Additionally, PPO employs an actor-critic architecture, which enhances its effectiveness by balancing exploration and exploitation while ensuring stable learning. Its use of the surrogate objective function improves training efficiency, making it a reliable choice for various DRL applications [32]. The pseudocode structure of the PPO techniques is depicted in Figure 5.

Algorithm 1 PPO-Clip

1:Input: initial policy parameters θ_0, initial value function parameters ϕ_0

2:for $k = 0, 1, 2, \ldots$ do

3:Collect set of trajectories $D_k = \{T_i\}$ *by running policy* $\pi k = \pi(\theta k)$ in the environment.

4:Compute rewards-to-go $\hat{R}_t$.

5:Compute advantage estimates, $\hat{A}_t$ (using any method of advantage estimation) based on the current value function $V_{\phi k}$.

6:Update the policy by maximizing the PPO-Clip objective:

$$\theta_{k+1} = \arg\max_{\theta} \frac{1}{|D_k| T} \sum_{T \varepsilon D_k} \sum_{t=0}^{T} \min\left(\frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta k}(a_t \mid s_t)} A^{T\theta_k}(s_t, a_t),\ g(\varepsilon, A^{\pi\theta_k}(s_t, a_t)) \right),$$

typically via stochastic gradient ascent with Adam.

7:Fit value function by regression on mean-squared error:

$$f_{k+1} = ang \min_{f} \frac{1}{|D_k| T} \sum_{T \varepsilon D_k} \sum_{t=0}^{T} \left(Vf(st) + \hat{R}_t \right)^2,$$

typically via some gradient descent algorithm.

8: end for

Figure 5.
Pseudocode structure of the PPO techniques [33].

3.6. Objective Function

Optimal sizing hybrid energy configuration for an off-grid EV charging station uses the DRL technique. The objective function of this work is shown in Equation 11. It aims to reduce costs and minimize energy deficit.

$$f = \min(C_{total} + E_{DP}) \tag{11}$$

The main objective is to reduce overall costs. The total system cost (C_{total}) can be calculated using Equation (12-16), and the Energy Deficit Penalty (E_{DP}) is the penalty when energy is insufficient, as derived from Equation 17.

$$C_{total} = C_{PV,install} + C_{PV,maintenance} + C_{BESS,install} + C_{BESS,maintenance} + C_{BESS,energy} \tag{12}$$

The PV system cost has details as shown in Equations 13-14.

$$C_{PV,install} = S_{PV} \times c_{PV,install} \tag{13}$$

$$C_{PV,maintenance} = S_{PV} \times c_{PV,maintenance} \times 20 \quad \text{(For 20 years)} \tag{14}$$

where $C_{PV,install}$ is the installation cost (THB), S_{PV} is the PV system size (kW), $c_{PV,install}$ is the installation price per unit of PV power (THB/kW), $C_{PV,maintenance}$ is the PV maintenance cost over 20 years (THB), and $c_{PV,maintenance}$ is the PV maintenance cost per unit per year (THB/kW/year).

The BESS cost equation can be calculated as follows:

$$C_{BESS,install} = (P_{BESS} \times C_{BESS,power}) + (E_{BESS} \times C_{BESS,energy}) \tag{15}$$

$$C_{BESS,replacemen} = C_{BESS,repl} \times \left(\frac{Y_{PL}}{Y_{BL}} \right) \tag{16}$$

$$C_{BESS,maintenance} = C_{BESS,maintenance} \times Y_{PL} \tag{17}$$

Where $C_{BESS,install}$ is the BESS installation cost, P_{BESS} is the power of BESS system (kW), E_{BESS} is the BESS energy capacity, $C_{BESS,power}$ is the price per power of BESS installation (THB/kW), $C_{BESS,energy}$ is the price per energy of the BESS, $C_{BESS,replacement}$ is the BESS lifetime replacement cost, $C_{BESS,maintenance}$ is the total maintenance cost, Y_{PL} is the project life, and Y_{BL} is the battery life.
Energy deficit penalty (E_{DP}) as shown in Equations 18.

$$E_{DP} = E_{deficit} \times k_{deficit} \tag{18}$$

Where $E_{deficit}$ is the amount of energy deficit during a specified period (kWh) is the penalty rate per unit of energy deficit (THB/kWh). If the SOC of the battery is below the specified level (ifE_{BESS}), an additional penalty will be imposed, as shown in Equation 19.

$$ifE_{BESS} < E_{\min,} P_{deficit} = P_{deficit} + k_{low\,SOC} \tag{19}$$

The structure of the deep learning operation for determining the optimal hybrid energy sizing for an Off-Grid EV charging station is shown in Figure 6.

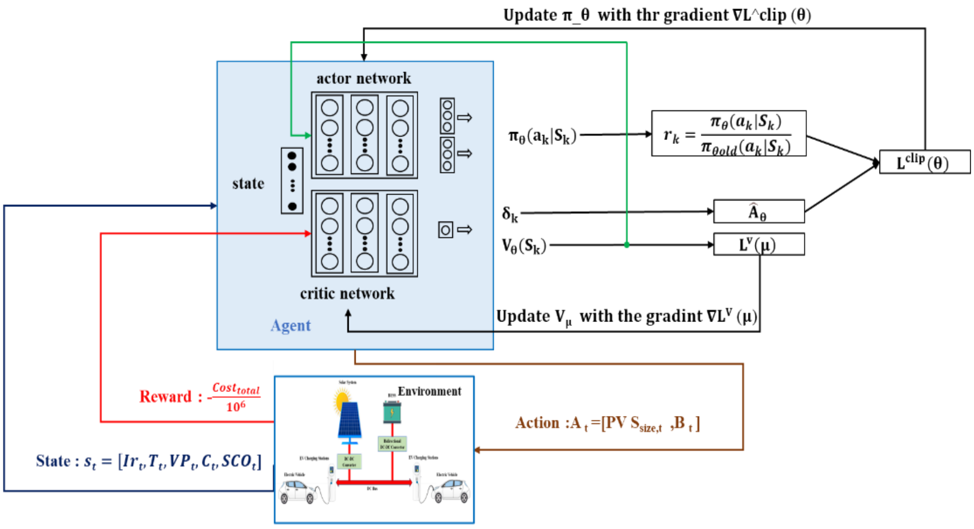

Figure 6.
The Structure of DRL operation for determining the optimal hybrid energy sizing for an Off-Grid EV Charging Station.

3.6.1. State Space

The state of the agent in the model consists of 4 variables. $S_t = [I_{rt}, PV_t, C_t, SOC_t]$ Where S_t is the state of energy at time, I_{rt} is the solar irradiance at time t (W/m²), PV_t is the temperature of the solar panel at time t (°C), C_t is the total energy consumption at time t (kW), SOC_t is the state of charge of the battery at time t (%).

3.6.2. Action Space

Actions that the agent can choose at each period include resizing the PV and managing battery power. $a_t = [PV\,S_t, B_t]$
Where a_t is the control action at time t, PVS_t is the PV system size variable at time t (range: -1 to +1), B_t is the battery power variable for charging/discharging at time t (range: -1 to +1)

3.6.3. Transition Function

The state transition function will consider the actions chosen by the agent. Results $s_{t+1} = f(s_t, a_t)$

Where s_{t+1} is the state at the next time, s_t is the state at time t, a_t is the action chosen at time t, and the function f describes the change in the state based on the action chosen by the agent, such as energy production from PV or charging/discharging energy from the battery.
The state change can be calculated from the sunlight's energy, as shown in Equation 20.

$$EP_t = PVS_t \times I_{rt} \times \eta \tag{20}$$

Where EP_t is the energy produced at time t (kW), η is the efficiency of the PV system (Dimensionless). The Battery State of Charge (SOC) update.

$$SOC_{t+1} = SOC_t + \frac{Battery\ Power_t \times \Delta t}{Battery\ Capacity} \tag{21}$$

Where Dt is the duration of each action.

3.6.4. Reward Function

This equation is a reward function used to calculate the total cost and adjust it to an appropriate value for the model's learning. The reward function is shown in Equation 22.

$$R_t = -\left(\frac{C_{total}}{10^6}\right) + E_{DP} \tag{22}$$

The total cost refers to the overall operational costs during each period, such as installation, maintenance, or system operation costs, divided by 10^6, which adjusts the reward values to appropriate units and reduces the scale of the numbers for calculation. The DRL algorithm is set with the optimal parameter values shown in Table 3 to ensure the system learns and makes decisions efficiently. The trained PPO model aims to optimize the size of PV and BESS, considering the power system's energy efficiency, cost reduction, and stability.

Table 3.
Parameters for the DRL techniques simulation.

Parameters	Set the configuration	Parameters	Set the configuration
Learning rate	1×10^{-4}	Number epoch	5
N_{step}	1,024	Entropy coefficient	0.005
Gamma	0.99	Number of eval episodes	140
Batch size	64	Total time steps	400

4. Results

The study aimed to find the optimal energy combination for an off-grid electric vehicle charging station using Stable-Baselines 3 for implementing PPO in Python. Stable-Baselines3 is a reinforcement learning library developed by Antonin Raffin and the team from the ROBUST AI team in collaboration with open-source community developers. It is an improvement over its predecessor, Stable-Baselines, developed by OpenAI Baselines, which has a systematic structure and has been optimized for high performance. Stable-Baselines3 was chosen for its compatibility with PyTorch and efficient support for deep reinforcement learning techniques, while the PPO algorithm was chosen for its ability to handle continuous action spaces well, making it suitable for the optimal sizing problem of solar PV and BESS.

The Mean Reward Progress graph in Figure 7 shows that the learning of the PPO algorithm (Blue line) increases rapidly in the early stages and remains constant because it is a policy-based RL using an actor-critic framework, which helps it learn well. A2C (Green line) is highly volatile (up and down) because it is an advantage-based approach, requiring entropy regularization to help improve, but tends to improve. The episodes increase, but they are not as stable as PPO. However, DQN (Orange line) has a constant negative mean reward throughout the training because it is a value-based RL, which is not suitable for complex problems or continuous action spaces.

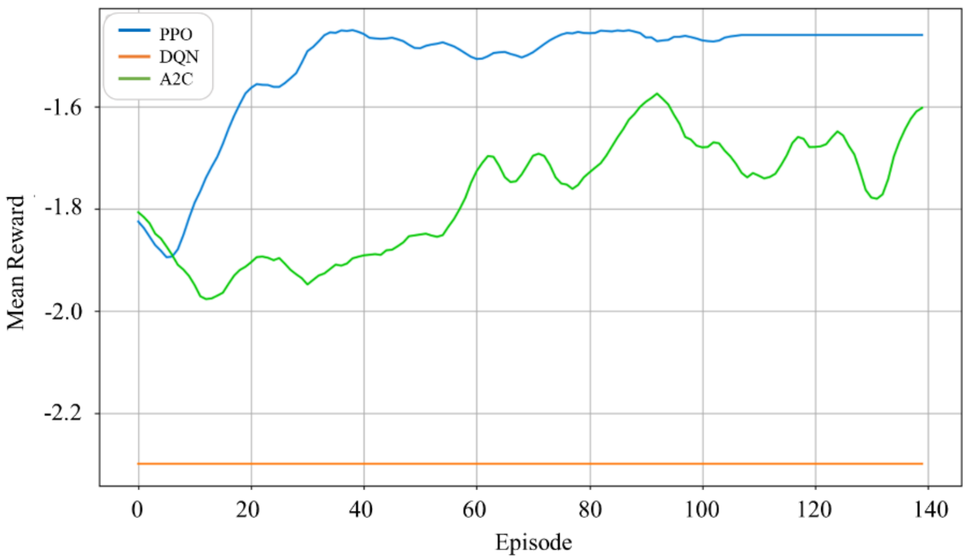

Figure 7.
Mean Reward Progress.

From the simulation comparing the learning of each model, PPO has PV and battery sizes suitable for energy production and consumption, with battery energy usage at approximately 48% of the total energy used and a final average reward of -1.43, indicating good efficiency in energy management. DQN with a larger PV size results in increased energy production, but the energy usage from the battery rises to 51% of the total energy used, with a final average reward of -2.28, indicating reduced efficiency compared to PPO. A2C has the same PV and battery size as PPO but uses 48% of the total energy from the battery, with a final average reward of -1.67, showing efficiency similar to PPO, as shown in Table 4.

Table 4.
Performance Evaluation of PPO, DQN, and A2C in PV-BESS Optimization.

Techniques	PV Size (kW)	BESS Size (kWh)	Total Production (kWh)	Consumption (kWh)	Battery Utilization	Final Mean Reward
PPO	90.00	120	411.32	345.96	0.48	-1.43
A2C	90.00	120	411.32	345.96	0.48	-1.673
DQN	108.42	120	591.64	345.96	0.51	-2.28

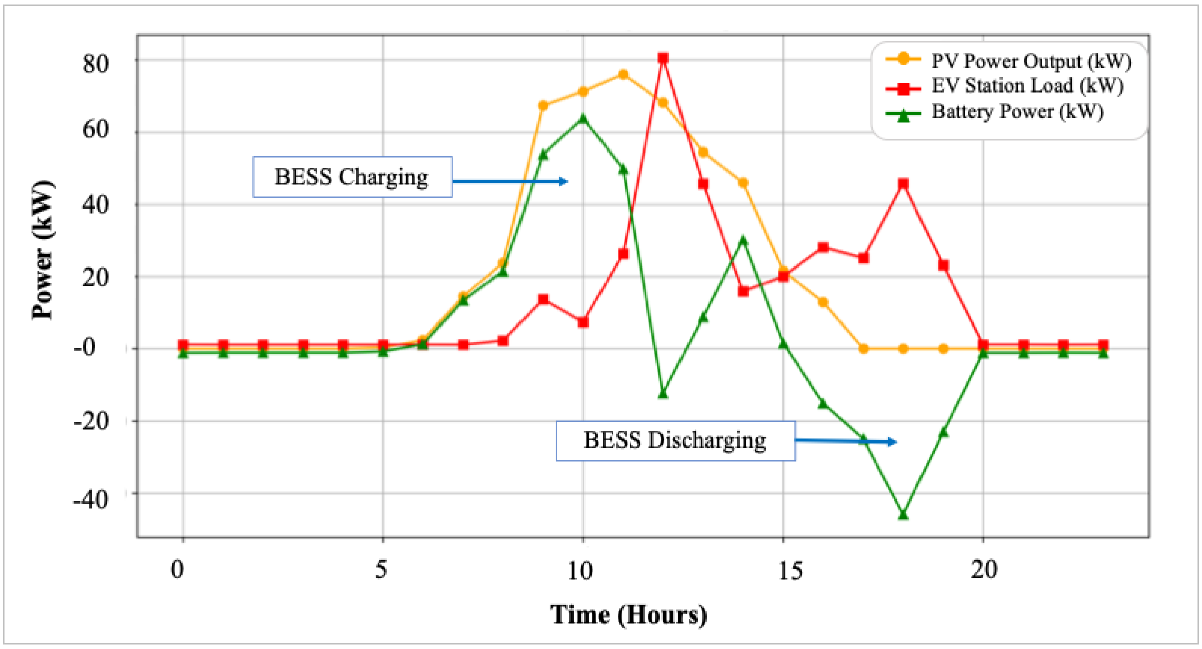

Figure 8.
The Hybrid PV-BESS Power on the EV Charging Station with the PPO Technique.

Figure 8 shows the power distribution within the system. The orange graph represents the power generated from PV, which peaks during the daytime and drops to zero in the evening. The green line indicates battery operation, where the battery charges when excess PV energy is available and discharges when PV production is insufficient. The red line represents the EV charging load, which fluctuates throughout the day, occasionally exceeding PV generation. The system has a 90 kW PV and 256.38 kWh battery, enabling efficient energy management.

Table 5 shows the daily energy and power performance of the charging station with PV-BESS optimization. The charging station uses a total energy of 345.44 kWh, all from the grid. In the case after installing PV and BESS, using deep learning simulations with the PPO and A2C models to find the appropriate sizes, the PV system can produce the energy of 411.32 kWh, and the BESS can store 256.38 kWh of energy. With the DQN model, the PV system can produce 591.64 kWh of energy, and the BESS can store 358.93 kWh of energy, which exceeds the charging station's demand, providing the station with an alternative energy source from the grid.

Table 5.
Daily Energy and Power Performance of Charging Station with PV-BESS Optimization.

Techniques (PV and BESS installation)	PV Peak power (kW)	PV Energy (kWh)	BESS Peak power (kW)	BESS Energy (kWh)
PPO Technique	90.00	411.32	120.00	256.38
A2C Technique	90.00	411.32	120.00	256.38
DQN Technique	112.00	591.64	120.00	358.93

Table 6.
Long-Term Cost Analysis for Grid-Connected and PV-BESS Energy Systems in EV Charging Stations.

Case study	Electricity bill (Baht)	Installation and maintenance costs for PV (Baht)	Installation and maintenance costs for BESS (Baht)	Total cost (Baht)
Normal case (Buy from the grid)	17,590,284	-	-	17,590,284
PV and BESS installation (PPO and A2C techniques)	-	3,069,669.60	5,938,032.00	**9,007,701.60**
PV and BESS installation (DQN technique)	-	3,439,233.28	5,938,032.00	9,758,065.28

Table 6 shows the long-term cost analysis for grid-connected and PV-BESS energy systems in EV charging stations. In the standard case, the total cost is 17,590,284 baht. After installing PV and BESS, using the simulation models (PPO and A2C), the total cost decreased to 9,007,701.60 baht, indicating a cost saving of 8,582,582.40 baht (reduced by 48.79%). In the DQN model, the total cost decreased to 9,758,065.28 baht, indicating a cost saving of 7,832,218.72 baht. The installation of PV systems and BESS significantly reduces the cost of purchasing electricity from the grid. The PPO and A2C models achieve the highest cost savings compared to the standard case.

Table 7.
Economic and Environmental Comparison: Performance Over 20 Years.

Parameters	PPO&A2C techniques	DQN technique
Net Present Value (NPV) (THB)	9,700,841	7,953,322
Internal Rate of Return (IRR) (%)	20.89	17.22
Benefit-Cost Ratio (BCR)	2.10	1.76
Payback Period (Years)	5	6
Carbon Reduction (kg CO_2)	1,383,480.0	1,721,664.0

Table 7 shows the economic and environmental performance comparison over 20 years. PPO & A2C provide superior financial performance with higher NPV, IRR, and BCR, and a quicker payback period over the 20-year project duration. However, DQN offers a higher level of carbon reduction, making it more beneficial from an environmental perspective.

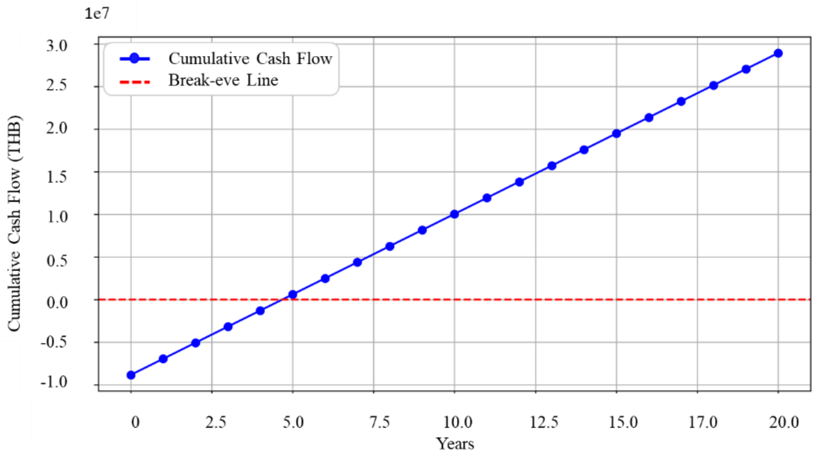

Figure 9.
Payback Period of the PPO technique.

Figure 9 illustrates the payback period for the PPO technique. The graph shows that the payback period is achieved within 5 years, indicating that the system recoups its initial investment relatively quickly. After the payback period, the energy savings continue to accumulate, further reducing the overall costs associated with grid electricity usage.

5. Discussions

This study explores the use of DRL techniques, particularly the PPO algorithm, to optimize the sizing of photovoltaic and battery energy storage systems for off-grid electric vehicle charging stations. The results demonstrate the effectiveness of DRL in addressing energy management challenges in off-grid applications, highlighting the potential of DRL as a promising solution for optimizing the RES.

The PPO algorithm demonstrated the ability to determine the optimal PV and BESS sizes, with the best configuration being a 90 kW PV system and a 120-kWh battery. This configuration allowed for efficient energy production and consumption, with the PV system generating 411.32 kWh of energy, which was sufficient to meet the charging station's needs. The battery stored excess energy during the day and discharged when PV production was insufficient at night. The DQN model, which utilized a larger PV size (108.42 kW), led to higher energy production but resulted in 51% battery utilization, which was less efficient than PPO. Both PPO and A2C models used 48% of the total energy from the battery, with PPO demonstrating the lowest final mean reward, indicating better efficiency in energy management.

Economically, the results from the PPO algorithm showed significant cost savings, reducing the electricity purchase cost by 48.79% compared to the grid-only scenario. The project's Net Present Value (NPV) was 9.7 million THB, with an Internal Rate of Return (IRR) of 20.89% and a payback period of just five years, demonstrating the economic viability of the optimized system.

From an environmental perspective, the PPO-optimized system reduced 1,383 tons of CO_2 emissions over 20 years, contributing to global sustainability efforts and underscoring the potential of renewable energy solutions to reduce greenhouse gas emissions in electric vehicle charging infrastructure.

However, there are areas for future research, such as integrating multiple renewable energy sources, including wind power, alongside PV and BESS. This could further improve the reliability and performance of the system. Exploring advanced DRL techniques like Soft Actor-Critic (SAC) may also help enhance system stability and performance under varying conditions. Future studies should also focus on scaling the model to larger systems and assessing its suitability for different geographic locations with varying solar power potential.

6. Conclusions

This paper uses deep reinforcement learning techniques to present the optimal sizing of a PV-BESS energy system for an off-grid electric vehicle charging station. The PPO algorithm has proven to be an effective method for optimizing the sizing and energy management of PV and BESS systems for off-grid EV charging stations. The optimal configuration of 90 kW PV and 120 kWh battery achieved significant energy efficiency, reducing electricity costs by 48.79% and cutting CO_2 emissions by 1,383 tons over 20 years. The project is economically viable and environmentally sustainable, with a Net Present Value (NPV) of 9.7 million THB and an Internal Rate of Return (IRR) of 20.89%. This study showcases the potential of DRL, specifically PPO, in optimizing renewable energy systems for off-grid applications. The system's success in reducing reliance on grid electricity highlights the role of advanced algorithms in driving the transition to cleaner energy solutions. Future research can enhance this approach by incorporating multiple renewable sources and further exploring more advanced DRL techniques to improve system efficiency, reliability, and scalability.

Future research will focus on enhancing renewable energy systems by integrating multiple energy sources, such as wind power, and employing more advanced DRL techniques. This will help optimize the sizing of solar panels, wind turbines, and battery systems to ensure maximum efficiency, reliability, and cost-effectiveness. The goal is to improve the overall performance and sustainability of renewable energy solutions, further contributing to the transition to cleaner and more sustainable energy sources.

References

[1] United Nations Thailand, *Causes and impacts of climate change*. Thailand: United Nations, 2025.

[2] International Energy Agency, *Global EV outlook* Paris, France: IEA, 2020.

[3] T. Boonraksa, A. Paudel, P. Dawan, and B. Marungsri, "Impact of electric bus charging on the power distribution system: A case study IEEE 33 Bus test system," presented at the Proceeding 2019 IEEE PES GTD Grand Int. Conference and Exposition Asia (GTD Asia), Bangkok, Thailand, 2019.

[4] S. Sudha Letha and M. Bollen, *Impact of electric vehicle charging on the power grid*. Sweden: Luleå University of Technology, Energy Science Sweden, 2021.

[5] I. E. Atawi, E. Hendawi, and S. A. Zaid, "Analysis and design of a standalone electric vehicle charging station supplied by photovoltaic energy," *Processes*, vol. 9, no. 7, p. 1246, 2021. https://doi.org/10.3390/pr9071246

[6] C. M. Affonso and M. Kezunovic, "Technical and economic feasibility analysis of a solar-powered electric vehicle charging station," presented at the 2019 IEEE Power & Energy Society General Meeting (PESGM), Atlanta, GA, USA, 2019.

[7] Megatech, "EV vehicle adoption and trends in ASEAN and Thailand in the next decade," Megatech," Retrieved: https://megatechthailand.com/manufacturing-trends/ev-vehicle-adoption-and-trends-in-asean-and-thailand-in-the-next-decade/, 2023.

[8] M. Saleem, S. Saha, U. Izhar, and L. Ang, "A stochastic MPC-based energy management system for integrating solar PV, battery storage, and EV charging in residential complexes," *Energy and Buildings*, vol. 325, p. 114993, 2024. https://doi.org/10.1016/j.enbuild.2024.114993

[9] A. Al Wahedi and Y. Bicer, "Techno-economic optimization of novel stand-alone renewables-based electric vehicle charging stations in Qatar," *Energy*, vol. 243, p. 122975, 2021. https://doi.org/10.1016/j.energy.2021.122975

[10] A. M. Soomro, M. K. Rathi, A. Shaikh, L. Kumar, M. E. H. Assad, and M. Ehyaei, "Techno-economic analysis of stand-alone hybrid PV-hydrogen-based plug-in electric vehicle charging station," *Energy Reports*, vol. 12, pp. 3279-3290, 2024. https://doi.org/10.1016/j.egyr.2024.09.019

[11] D. Liu, P. Zeng, S. Cui, and C. Song, "Deep reinforcement learning for charging scheduling of electric vehicles considering distribution network voltage stability," *Sensors*, vol. 23, no. 3, p. 1618, 2023. https://doi.org/10.3390/s23031618

[12] G. Liu, Y. Xue, M. S. Chinthavali, and K. Tomsovic, *Optimal sizing of PV and energy storage in an electric vehicle extreme fast charging station*. Oak Ridge, TN, USA: Oak Ridge National Laboratory, 2025.

[13] D. Yan and C. Ma, *Optimal sizing of a PV-based electric vehicle charging station under uncertainties*. Shanghai, China: UM-SJTU Joint Institute, 2025.

[14] W. Vermeer, G. R. C. Mouli, and P. Bauer, "Optimal sizing and control of a PV-EV-BES charging system including primary frequency control and component degradation," *IEEE Open Journal of the Industrial Electronics Society*, vol. 3, pp. 236-251, 2022. https://doi.org/10.1109/OJIES.2022.3161091

[15] R. A. Rachmanto, "Demand side energy management of EV charging stations by approximate dynamic programming," *E3S Web Conference*, vol. 465, p. 01004, 2023. https://doi.org/10.1051/e3sconf/202346501004
[16] Y. Tavakol-Moghaddam, M. Boroushaki, and M. Astaneh, "Reinforcement learning for battery energy management: A new balancing approach for Li-ion battery packs," *Results in Engineering*, vol. 23, p. 102532, 2024. https://doi.org/10.1016/j.rineng.2024.102532
[17] I. Ahmed, A. Pedersen, and L. Mihet-Popa, "Smart microgrid optimization using deep reinforcement learning by utilizing the energy storage systems," in *Proceeding 2024 International Conference on Smart Grid and Renewable Energy*, 2024, doi: https://doi.org/10.1109/sgre59715.2024.10428874.
[18] I. Ahmed, A. Pedersen, and L. Mihet-Popa, "Distributed energy management and demand response in smart grids: A multi-agent deep reinforcement learning framework," *arXiv preprint*, pp. 1–15, 2022. https://doi.org/10.48550/arXiv.2201.02947
[19] D. Zhao, S. Zhang, and L. Xie, "Deep reinforcement learning-based energy management in smart grids with demand response," *IEEE Transaction Smart Grid*, vol. 14, no. 7, pp. 4098–4107, 2023. https://doi.org/10.1109/TSG.2023.3147857
[20] F. Aksan, V. Suresh, and P. Janik, "Optimal capacity and charging scheduling of battery storage through forecasting of photovoltaic power production and electric vehicle charging demand with deep learning models," *Energies*, vol. 17, no. 11, p. 2718, 2024. https://doi.org/10.3390/en17112718
[21] K. N. D. V. Sai Eswar, M. Arun Noyal Doss, M. Shorfuzzaman, and A. Elrashidi, "Microgrid system for electric vehicle charging stations integrated with renewable energy sources using a hybrid DOA–SBNN approach," *Frontiers in Energy Research*, vol. 12, p. 1492243, 2025. https://doi.org/10.3389/fenrg.2024.1492243
[22] R. Fachrizal, M. Shepero, M. Åberg, and J. Munkhammar, "Optimal PV-EV sizing at solar powered workplace charging stations with smart charging schemes considering self-consumption and self-sufficiency balance," *Applied Energy*, vol. 307, p. 118139, 2022. https://doi.org/10.1016/j.apenergy.2021.118139
[23] S. Sykiotis, C. Menos-Aikateriniadis, A. Doulamis, N. Doulamis, and P. S. Georgilakis, "Solar power driven EV charging optimization with deep reinforcement learning," presented at the 2022 2nd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED), IEEE, 2022.
[24] G. Rituraj, G. R. C. Mouli, and P. Bauer, "A comprehensive review on off-grid and hybrid charging systems for electric vehicles," *IEEE Open Journal of the Industrial Electronics Society*, vol. 3, pp. 203-222, 2022. https://doi.org/10.1109/OJIES.2022.3167948
[25] M. Kumar, K. P. Panda, R. T. Naayagi, R. Thakur, and G. Panda, "Comprehensive review of electric vehicle technology and its impacts: detailed investigation of charging infrastructure, power management, and control techniques," *Applied Sciences*, vol. 13, no. 15, p. 8919, 2023. https://doi.org/10.3390/app13158919
[26] Metropolitan Electricity Authority (MEA), "Electric monthly calculate type 3," Retrieved: https://www.mea.or.th/our-services/mea-service/e-service/electric-monthly-calculate/type3, 2024.
[27] R. A. Rachmanto, "Energetic and economic viability of off-grid PV-BESS for charging electric vehicles: Case study of Yogyakarta," *E3S Web Conference*, vol. 465, p. 01004, 2023.
[28] Y. Teng and J. Wang, "New characteristics of geometric dilution of precision (GDOP) for multi-GNSS constellations," *The Journal of Navigation*, vol. 67, no. 6, pp. 1018-1028, 2014.
[29] A. Sari and R. Doğan, "“Techno-economic analysis of a stand-alone hybrid renewable energy system (solar/fuel cell/battery) and grid extension for two residential districts," *Ain Shams Engineering Journal*, vol. 15, no. 12, p. 103062, 2024. https://doi.org/10.1016/j.asej.2024.103062
[30] Metropolitan Electricity Authority (MEA), *Greenhouse gas emission factor from electricity production/consumption for carbon reduction projects and activities*. Bangkok, Thailand: Metropolitan Electricity Authority (MEA), 2020.
[31] Y. B. Muna and C.-C. Kuo, "Feasibility and techno-economic analysis of electric vehicle charging of PV/wind/diesel/battery hybrid energy system with different battery technology," *Energies*, vol. 15, no. 12, p. 4364, 2022. https://doi.org/10.3390/en15124364
[32] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017. https://arxiv.org/abs/1707.06347
[33] A. Raffin, *PPO — Stable Baselines3 2.6.0a1 documentation," Stable-Baselines3 Documentation*. USA: Stable-Baselines3, 2025.