Deep Learning-Based Optimal Sizing of a Grid-Tied Microgrid Under Real-Time Pricing

IET Smart Grid

WILEY

The Institution of Engineering and Technology

ORIGINAL RESEARCH OPEN ACCESS

Deep Learning-Based Optimal Sizing of a Grid-Tied Microgrid Under Real-Time Pricing

Rahmat Khezri[1] | Peyman Razmi[2] | Amin Mahmoudi[3] | Mohammad Hassan Khooban[4]

[1]Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden | [2]Department of Electrical and Computer Engineering, University of Vaasa, Vaasa, Finland | [3]College of Science and Engineering, Flinders University, Adelaide, Australia | [4]Department of Electrical and Computer Engineering, Aarhus University, Aarhus, Denmark

Correspondence: Amin Mahmoudi (amin.mahmoudi@flinders.edu.au)

Received: 5 August 2025 | **Revised:** 28 November 2025 | **Accepted:** 19 December 2025

Keywords: battery storage plants | deep reinforcement learning | microgrid design

ABSTRACT

This paper presents a novel optimal sizing model for a grid-tied microgrid operating under real-time pricing (RTP) for electricity trading with the main grid. The model determines the optimal capacities of solar photovoltaic (PV), wind turbine (WT), battery energy storage (BES) and inverter using a deep reinforcement learning approach. A double deep Q-network (DDQN) is proposed to solve the complex sizing problem efficiently. The sizing model incorporates a rule-based energy management strategy, where the average of day-ahead electricity price forecasts is used to guide the battery's state of charge (SoC) decisions. Although the model is designed to be generic, a residential building in Australia is used as a case study to validate its practical applicability. Numerical results demonstrate that the proposed method under RTP conditions achieves a lower net present cost (NPC) of electricity compared to existing sizing models from previous studies. The effectiveness and robustness of the proposed deep learning-based approach are further confirmed through comparative analysis with other machine learning techniques and metaheuristic algorithms.

1 | Introduction

1.1 | Background and Motivation

Residential electricity tariffs are undergoing significant transformation with the advancement of smart grid technologies. Traditionally, consumers have been charged using flat tariffs, where the electricity price remains constant throughout the day and across the year. To encourage more efficient energy use, time-of-use (TOU) tariffs were later introduced. Under TOU tariffs, electricity rates vary by time of day, typically classified into off-peak, shoulder and peak periods, allowing consumers to reduce costs by shifting consumption to off-peak hours [1].

To provide even greater flexibility and align prices more closely with real-time supply and demand conditions, real-time pricing (RTP) tariffs have been developed. In an RTP scheme, electricity prices fluctuate dynamically, typically on an hourly or half-hourly basis, based on the utility's real-time generation costs and system load [2]. With the continued deregulation of electricity markets and the deployment of smart grid infrastructure, RTP is expected to gradually replace both flat and TOU tariffs. Beyond improving overall power system efficiency, RTP offers financial opportunities for microgrid operators, who can leverage the variability in electricity prices to minimise operational costs [3].

To capitalise on these opportunities, microgrids often employ battery energy storage (BES) to manage the intermittent output from renewable energy sources such as solar photovoltaic (PV) and wind turbines (WT). However, a key challenge under RTP conditions lies in the optimal sizing of BES and generation assets. Proper sizing is critical for balancing cost, reliability and performance in the face of uncertain generation and dynamic pricing.

1.2 | Literature Review

Previous studies have explored the optimal sizing of residential microgrids primarily under flat and TOU tariff structures. For example, a conventional energy management system (EMS) was proposed for a home microgrid operating under a TOU tariff in ref. [4]. Sizing of wind turbine-battery energy storage (WT-BES) systems [5] and photovoltaic-battery energy storage (PV-BES) systems [6] was carried out using rule-based EMSs under flat tariff conditions. In ref. [7], the uncertainties associated with PV generation and residential load were considered for the optimal sizing of PV and BES under a TOU tariff. Additionally, ref. [8] demonstrated the use of load shifting strategies to shift microgrid consumption from high-price to low-price periods under TOU tariffs.

However, the sizing models developed in these studies are relatively simplistic, primarily because the electricity pricing structures are either static (flat tariff) or vary at only two or three fixed intervals throughout the day (TOU tariff). This limited price variability does not adequately capture the complexities introduced by RTP environments.

Optimal microgrid sizing under RTP tariffs presents a significant challenge due to the dynamic and unpredictable nature of electricity prices. The uncertainty around when to buy or sell electricity from/to the main grid or when to charge or discharge the BES, complicates the decision-making process, especially when determining the appropriate capacities of PV, WT and BES components. In ref. [9], a stochastic optimisation approach was used to determine the capacities of WT and BES for a smart home, accounting for uncertainties in load, wind generation and market prices. However, the model employed a relatively basic rule-based operation strategy and stochastic models in general require advanced statistical and computational techniques.

In ref. [10], the optimal sizing of PV and BES was investigated under various tariff structures, including RTP. However, a uniform operational strategy was applied across all tariffs, failing to account for the distinct variability inherent in RTP. Some studies have proposed BES control strategies specifically tailored for RTP environments [11, 12], but these did not address the sizing problem. In ref. [13], peak shaving techniques were applied in response to RTP fluctuations to optimise operational performance, yet the system's component sizing was not optimised. Similarly, ref. [14] developed an optimal scheduling model for a residential load based on electricity price forecasting and load shifting under RTP. However, load shifting can often be impractical or undesirable due to potential impacts on customer satisfaction and comfort.

In ref. [15], battery sizing was performed under an RTP tariff; however, other critical components, such as solar PV, were not considered. In ref. [16], demand response programmes were incorporated to optimise microgrid sizing under RTP, whereas ref. [17] addressed the optimal sizing problem using various demand-side management strategies and evolutionary algorithms. Despite these advancements, studies [15–17] did not account for battery degradation, a factor that significantly influences both the operational strategy and optimal capacity planning. Moreover, the applicability of these models to real-world scenarios remains unclear due to limited practical validation.

Reference [18] focused on optimising the size of a grid-tied hybrid energy system using forecasted meteorological data. Hourly temperature and solar radiation were predicted using machine learning techniques, with Gaussian process regression yielding the highest accuracy. The tunicate swarm algorithm was employed for system sizing and benchmarked against particle swarm optimization (PSO) and harmony search. Results showed that incorporating forecasted data reduced the cost of energy by 0.33%, with tunicate swarm algorithm producing the most cost-effective and reliable configuration. In ref. [19], a mixed-integer linear programming framework was proposed for optimal sizing and energy management of a grid-connected microgrid. The objective was to minimise both energy costs and carbon emissions by integrating TOU pricing, demand response and renewable energy sources. The model incorporated battery storage, electric vehicles and renewable generation into a unified optimisation strategy.

In ref. [20], the authors proposed a fuzzy logic-based control strategy for a weak grid-tied PV-BES system that has been optimally sized. The objective is to achieve economical power regulation by dynamically adapting to seasonal TOU tariffs. The controller efficiently managed power flows to minimise energy costs while maintaining grid stability. Reference [21] introduced a novel analytical method for determining the optimal sizing, placement and scheduling of BES in a grid-connected microgrid. By accounting for various heating load scenarios, the method aims to enhance system efficiency through reduced operational costs and improved energy reliability. In ref. [22], the authors presented a systematic review of over 180 recent studies on the optimal sizing, techno-economic feasibility and reliability of hybrid renewable energy systems, with a specific focus on energy storage system integration. Reference [23] proposed a comprehensive framework for the optimal sizing and placement of distributed generation and BES in distribution networks. A multiobjective optimisation approach is adopted to balance energy loss minimisation, voltage profile enhancement and reliability improvement, with several planning scenarios evaluated. In ref. [24], an optimisation strategy was presented for sizing a hybrid energy system intended to supply auxiliary services in substations. The system includes PV panels, batteries and diesel generators. A hybrid algorithm combining genetic algorithm and variable neighbourhood search was employed to minimise cost while ensuring reliability.

These studies highlight that microgrid sizing problems are inherently nonlinear, particularly due to battery degradation and its impact on system lifetime and cost. Consequently, both the development of a robust sizing model and the choice of optimisation method are critical. Although classical optimisation techniques have been applied, they often struggle with the nonlinear and complex nature of the problem. Metaheuristic methods, such as PSO, used in refs. [4, 5], have been widely adopted, but they often suffer from low exploration efficiency in large search spaces and high computational requirements. Recently, machine learning (ML) algorithms have emerged as a promising alternative. ML methods are model-free and capable of handling problem complexity by learning from training data

[25]. Their flexibility allows for efficient handling of diverse scenarios, and they have been successfully applied in recent power system operation and planning studies.

To achieve a practical model of microgrid sizing, several factors, such as real data, salvation cost, grid constraints and components' degradations should be applied. However, previous studies overlooked these factors in most cases. Table 1 lists a summary of the deficiencies in the existing works on microgrid's sizing based on the electricity tariff, energy management, applied algorithm and the practicality of the study.

1.3 | Contributions

This paper makes several key contributions that extend beyond the existing literature on electricity pricing-based models and PV/BES sizing. From a practical standpoint, it addresses a current and industry-relevant challenge through a real-world Australian case study. The key contributions of this study, distinguishing it from existing pricing-based and PV/BES sizing literature, are as follows:

- A comprehensive optimisation framework is developed to jointly optimise the capacities of PV, WT, BES and inverters. The approach integrates a rule-based operational strategy based on day-ahead price forecasts, enabling explicit consideration of RTP volatility, an aspect most often overlooked in models designed.
- A deep Q-Network reinforcement learning is adapted to the sizing problem, extending the use of RL beyond operational control. The method learns cost-optimal capacity decisions through interaction with the environment and incorporates practical factors including battery degradation, grid constraints and salvage value.
- The proposed methodology is demonstrated through a real-world Australian case study. Real load, meteorological and pricing data are applied to demonstrate the proposed model's performance. Comparative results with machine learning and metaheuristic techniques confirm the economic advantages and robustness of the proposed approach.

2 | Microgrid Model

Figure 1 shows the proposed system configuration and sizing framework for the grid-tied microgrid in this study. The PV, WT and BES will be connected to the microgrid through a central IVT. This section describes the model of components and the rule-based operation for the grid-tied microgrid.

2.1 | Model of Components

The generated power by PV (P_{PV}) is given by ref. [8]:

$$P_{PV}(t) = \eta_{PV} P_{PV}^{rt} \times \frac{G_c(t)}{G_{stc}}\left[1 - \gamma(T_c(t) - T_{stc})\right] \quad (1)$$

$$T_c(t) = T_b(t) + G_c(t) \cdot \frac{\mathrm{NOCT} - 20}{0.8} \quad (2)$$

where G_c is the radiation on the PV collector in kW/m^2, η_{PV} is the PV efficiency (accounting for cable and converter losses) and T_c and T_b are the cell and ambient temperatures in °C, and NOCT is the normal operating cell temperature (assumed as 45°C). G_{stc} and T_{stc} are the irradiance and temperature at

TABLE 1 | Summary of deficiencies of the existing works.

Ref.	Electricity tariff	Optimal sizing	Energy management	Optimization algorithm	Practicality		
					CD	SC	GC
[4]	TOU	✓	×	EA	✓	×	×
[5]	Flat	✓	✓	EA	×	×	✓
[6]	Flat	✓	×	Classic	×	×	×
[7]	TOU	✓	×	Classic	×	×	×
[8]	TOU	✓	✓	Classic	×	×	✓
[9]	RTP	✓	✓	PSO	×	×	✓
[10]	RTP	✓	×	Classic	×	×	×
[11]	RTP	×	×	Classic	×	×	×
[12]	RTP	×	×	Classic	×	×	×
[13]	RTP	×	×	Classic	×	×	✓
[14]	RTP	×	✓	Classic	×	×	×
[15]	RTP	✓	×	Classic	×	×	×
[16]	RTP	✓	×	EA	×	×	×
[17]	RTP	✓	×	EA	×	×	×
This paper	RTP	✓	✓	DML	✓	✓	✓

Abbreviations: CD: components' degradations; DML: deep machine learning; EA: evolutionary algorithm; GC: grid constraint and SC: salvage cost.

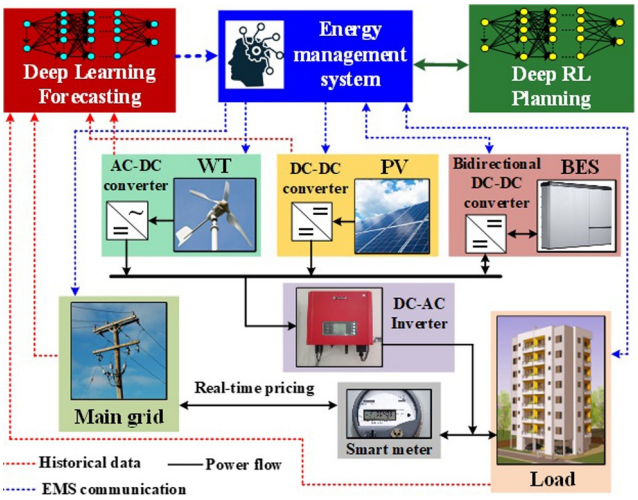

FIGURE 1 | Proposed system configuration and sizing framework of the grid-tied microgrid.

standard test conditions, considered to be 1 kW/m^2 and 25°C. γ is the temperature derating coefficient (0.4%/W/°C).

The output power (P_{WT}) of the wind turbine is given by a piecewise function of wind velocity (v) [6]:

$$P_{WT}(t)=\begin{cases}0, & v<v_c \text{ or } v>v_f\\ P_{WT}^{rt}\left(\dfrac{v-v_c}{v_t-v_c}\right)^3, & v_c\le v<v_t\\ P_{WT}^{rt}, & v_t\le v\le v_f\end{cases} \tag{3}$$

where v_f, v_c and v_t are the cut-out, cut-in and rated wind speeds, respectively. The WT's rated power is indicated by P_{WT}^{rt}. The power of PV and WT degrades over time [26]. In this study, constant annual degradations are considered for PV and WT to obtain a more realistic model of those components. Based on the output powers, annual degradations and numbers of PV and WT in the optimisation problem, the power of renewable energy (RE) is measured as follows:

$$P_{RE}(t)=\mathcal{N}_{WT}(1-d_{WT})^l P_{WT}(t)+\mathcal{N}_{PV}(1-d_{PV})^l P_{PV}(t) \tag{4}$$

The BES model contains the state of charge (SOC) formulation based on the available input/output power limit of the battery. The current battery SOCt depends on the SOC value in the previous sample ($t-1$) and the power extracted or injected in the current sample t.

$$SOC_{ES}(t)=SOC_{ES}(t-\Delta t)+\frac{\left(P_{ES}^{cg}(t)\,\eta_{ES}^{cg}-\frac{P_{ES}^{dc}(t)}{\eta_{ES}^{dc}}\right)\Delta t}{E_{ES}^{\max}} \tag{5}$$

The charging and discharging powers of BES are limited to available input and output powers of the battery. The following equations indicate the limits of input and output powers for the battery in each time interval.

$$P_{ES}^{in}(t)=\min\left(P_{ES}^{\max},\ \frac{E_{ES}^{\max}}{\Delta t}\left(SOC_{ES}^{\max}-SOC_{ES}(t)\right)\right) \tag{6}$$

$$P_{ES}^{ou}(t)=\min\left(P_{ES}^{\max},\ \frac{E_{ES}^{\max}}{\Delta t}\left(SOC_{ES}(t)-SOC_{ES}^{\min}\right)\right) \tag{7}$$

The maximum energy and power of the BES are formulated based on the number of the batteries in the optimisation problem as follows:

$$P_{ES}^{\max}=\mathcal{N}_{ES}\,P_{ES},\quad E_{ES}^{\max}=\mathcal{N}_{ES}\,E_{ES} \tag{8}$$

The capacity of the IVT, through the optimisation, should be greater than the power that passes through it.

$$\eta_{IV}\cdot\left(P_{RE}(t)+P_{ES}^{dc}(t)\right)\le\mathcal{N}_{IV}\,P_{IV} \tag{9}$$

2.2 | Microgrid's Operation Strategy

Operation strategy under a dynamic pricing is important to achieve efficient and economic power flow between the components. In this study, a rule-based operation is developed based on the next hour and day-ahead forecasts of data (Figure 2). Rule-based operations offer valuable advantages such as straightforward and understandable model, faster execution of decisions, easy implementation and user-friendliness [5]. Figure 2 shows that the operation strategy starts with obtaining the forecast data. In this study, the long short-term memory (LSTM) method is used for prediction of next hour wind, load, electricity price and radiation as well as the day-ahead price. The architecture of LSTM is an artificial recurrent neural network applied in the field of deep learning [27]. The main reason for using LSTM is its feedback connections unlike standard feedforward neural networks. This means LSTM can process not only single data points but also entire sequences of data. Afterwards, the average of day-ahead price of RTP is calculated. Then, the next hour forecast of RTP price at each time interval is compared with the calculated average price. The main idea behind the new rule-based EMS is to avoid the BES to discharge when the purchasing electricity price is lower than the average of forecasted day-ahead purchasing RTP (RTP_P^{avg}). On the other hand, the battery should not be charged if the selling electricity price is higher than the average of forecasted day-ahead selling RTP (RTP_S^{avg}).

In the EMS, if the generated power by RE is greater than the load, and the selling electricity price (RTP_S) is lower than RTP_S^{avg}, then the available input energy of BES (P_{ES}^{in}) is evaluated. If the excess power of RE is less than P_{ES}^{in}, then all the power is used for battery charging as follows:

$$P_{ES}^{cg}(t)=P_{RE}(t)-P_M(t) \tag{10}$$

On the other hand, if the excess power of RE is higher than P_{ES}^{in}, then it would be sold to the main grid (P_S).

$$P_S(t)=P_{RE}(t)-P_M(t)-P_{ES}^{cg}(t) \tag{11}$$

For the times that the selling electricity price is higher than RTP_S^{avg}, the excess power of RE would be first sold to the grid due to high price of electricity. In this condition, the sold power to the grid is calculated by:

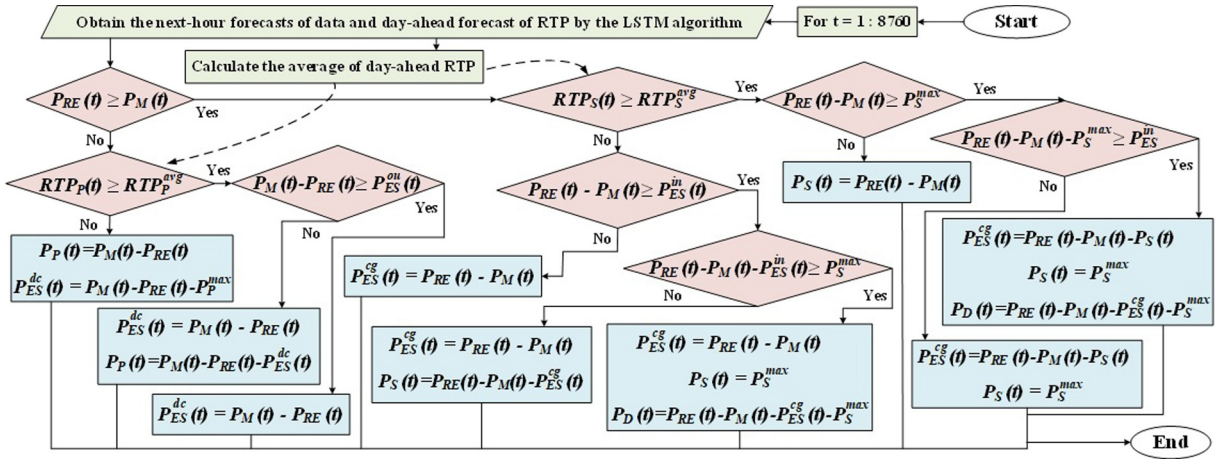

FIGURE 2 | The proposed rule-based operation strategy for the grid-tied microgrid.

$$P_S(t) = P_{RE}(t) - P_M(t) \tag{12}$$

If the excess power exceeds the maximum allowable selling power to the grid $\left(P_S^{\max}\right)$, then the maximum power would be first sold to the grid and the remaining power is used to charge The BES is as follows:

$$P_{ES}^{cg}(t) = P_{RE}(t) - P_M(t) - P_S^{\max} \tag{13}$$

If the extra power is beyond the maximum selling power limit and the available input power limit of BES, the extra power should be curtailed. The curtailed power (P_D) is calculated as follows:

$$P_D(t) = P_{RE}(t) - P_M(t) - P_S(t) - P_{ES}^{cg}(t) \tag{14}$$

It is to be noted that the curtailed power is not a physical load, and it is assumed to be controlled by the control system of inverter via a feedback loop [6]. If P_{RE} is lower than microgrid's load, and the purchasing electricity price (RTP_P) is lower than RTP_P^{avg}, then the deficit power is first supplied by purchasing from the grid (P_P) and then by discharging the BES $\left(P_{ES}^{dc}\right)$ as follows:

$$P_P(t) = P_M(t) - P_{RE}(t) \tag{15}$$

$$P_{ES}^{dc}(t) = P_M(t) - P_{RE}(t) - P_P^{\max} \tag{16}$$

However, if the purchasing electricity price (RTP_P) is greater than RTP_P^{avg}, then the deficit power is first supplied by discharging power of BES by considering the available output power limit of battery $\left(P_{ES}^{ou}\right)$. If P_{ES}^{ou} is higher than the power deficit, then no power is purchased from the grid and the discharging power of BES is sufficient.

$$P_{ES}^{dc}(t) = P_M(t) - P_{RE}(t) \tag{17}$$

If P_{ES}^{ou} is lower than the power deficit, then the deficit power after discharging the BES is purchased from the grid by:

$$P_P(t) = P_M(t) - P_{RE}(t) - P_{ES}^{dc}(t) \tag{18}$$

The developed operation strategy is then used in the sizing stage to achieve the optimal capacity of components.

3 | Optimal Sizing Model

Optimal sizing is a compromise problem to achieve the lowest operation cost in long-term while considering the investment cost of components. In this paper, the sizing model optimises the primary system capacities, PV, WT, BES and the central inverter, whereas the associated DC/DC and AC/DC converters are implicitly linked to these capacities and remain within realistic operating ranges. Therefore, the optimisation does not assume oversized or impractical converter designs; rather, converter ratings follow the selected component capacities in accordance with typical engineering practice.

In the proposed framework, the microgrid first operates based on the rule-based real-time dispatch strategy, which manages charging, discharging and power exchange decisions under RTP conditions. Building on this operational foundation, the DRL-based methodology is then employed to determine the optimal capacities of the microgrid components.

This section indicates the problem formulation and the methodology to solve the problem.

3.1 | Optimisation Formulation

The objective function is defined as the total net present cost (NPC) of the grid-tied microgrid. The total NPC (NPC_{tot}) contains the NPC of components (NPC_{com}) and the NPC of operation which is the electricity trading between the microgrid and the main grid (NPC_{tra}).

$$f = \min_{N_i} \text{NPC}_{tot} \tag{19}$$

$$\text{NPC}_{tot} = \text{NPC}_{com} + \text{NPC}_{tra} \tag{20}$$

The NPC_{tra} is calculated based on the real capital recovery factor (RCRF) and the annual electricity trading cost between the microgrid and the main grid (C_{tra}):

$$\text{NPC}_{tra} = \frac{C_{tra}}{\text{RCRF}} \tag{21}$$

The C_{tra} is the summation of three values which are the total annual cost for purchased electricity from the grid (AC_P), the total annual cost of sold electricity to the grid (AC_S) and the annual supply of charge for the grid connection (AC_C):

$$C_{tra} = \underbrace{\sum_{t=1}^{T} RTP_P(t)P_P(t)\Delta t}_{AC_P} - \underbrace{\sum_{t=1}^{T} RTP_S(t)P_S(t)\Delta t}_{AC_S} + \underbrace{\tilde{D}C_S}_{AC_C} \quad (22)$$

The RCRF is calculated based on the real interest rate (u) and the project lifetime (l) as follows:

$$\text{RCRF} = \frac{u(1+u)^l}{(1+u)^l - 1} \quad (23)$$

The real interest rate is obtained based on the interest rate (g) and an escalation rate (e):

$$u = \frac{g-e}{1+e} \quad (24)$$

The NPC_{com} is calculated based on the net present discounted values of capital cost $\left(PD_i^{cap}\right)$, operation and maintenance cost $\left(PD_i^{opm}\right)$, replacement cost $\left(PD_i^{rep}\right)$ and salvage cost $\left(PD_i^{sal}\right)$.

$$\text{NPC}_{com} = \sum_{i=1}^{4} N_i\left(PD_i^{cap} + PD_i^{opm} + PD_i^{rep} - PD_i^{sal}\right) \quad (25)$$

The PD_i^{cap} is the initial investment cost for the components.

$$PD_i^{cap} = C_i^{cap} \quad (26)$$

The PD_i^{opm} is calculated based on the annual cost of operation and maintenance $\left(C_i^{opm}\right)$ for each component and the capital recovery factor (CRF).

$$PD_i^{opm} = C_i^{opm} \cdot \frac{1}{\text{CRF}} \quad (27)$$

The CRF is obtained based on the interest rate and the project lifetime as follows:

$$\text{CRF} = \frac{g(1+g)^l}{(1+g)^l - 1} \quad (28)$$

The PD_i^{rep} is formulated based on the replacement cost $\left(C_i^{rep}\right)$ for each component and its replacement year.

$$PD_i^{rep} = C_i^{rep} \sum_{n=1}^{nR_i<l} \frac{1}{(1+g)^{nR_i}} \quad (29)$$

The PD_i^{sal} is calculated based on the remaining lifetime of each component at the end of the project lifetime (M_i) over the component's lifetime (L_i) as follows:

$$PD_i^{sal} = C_i^{cap} \cdot \frac{M_i}{L_i} \cdot \frac{1}{(1+g)^l} \quad (30)$$

It is notable that the lifetime of BES should be computed based on its degradation. The degradation of BES depends on the depth of discharge (DOD) in each cycle of battery and hence the system operation. When the annual system operation is terminated, the number of charge/discharge cycles is extracted and their SOCs are obtained and then the DOD is calculated (DOD = 1 − SOC). The *rainflow* cycle counting algorithm is used to extract the cycles and their DODs [16]. Based on an experimental model, the degradation of BES for each full cycle can be formulated as follows [28]:

$$d_{ES}(c) = \frac{b_1}{b_2 e^{-b_3 DOD(c)} + b_4} \quad (31)$$

The parameters b_1 to b_4 are constant and they are taken from [16]. If the counted cycle by *rainflow* algorithm is half, then the degradation is half of the full cycles. When the degradations for all full and half cycles are obtained, the total degradation of the BES is calculated for the annual operation:

$$d_{ES}^{tot} = \sum d_{ES}(c) \quad (32)$$

If the total annual degradation exceeds 20%, the BES requires replacement. Hence, the lifetime of BES based on its calendar lifetime (20-year) and d_{ES}^{tot} can be obtained as follows:

$$L_{ES} = \min\left(L_{ES}^{ca}, \left|\frac{20\%}{d_{ES}^{tot}}\right|\right) \quad (33)$$

Since the levelised cost of electricity (LCOE) is another useful economic index to compare the electricity generation system of microgrids, it has also been used in this study. The LCOE is calculated based on the net present cost (NPC) of components and electricity trading as well as the CRF, RCRF and the total electricity demand of the microgrid (E_M):

$$\text{LCOE} = \frac{\text{NPC}_{com} \cdot \text{CRF} + \text{NPC}_{tra} \cdot \text{RCRF}}{E_M} \quad (34)$$

The design constraints to run the optimal sizing model of the grid-tied microgrid are as follows:

$$P_{RE}^{ac}(t) + P_p(t) + P_{ES}^{dc}(t) = P_M(t) + P_{ES}^{cg}(t) + P_S(t) + P_D(t) \quad (35)$$

$$\mathcal{N}_i^{\min} \le \mathcal{N}_i \le \mathcal{N}_i^{\max} \quad (36)$$

$$SOC_{ES}^{\min} \le SOC_{ES}(t) \le SOC_{ES}^{\max} \quad (37)$$

$$0 \le P_{ES}^{cg}(t) \le P_{ES}^{in}(t) \quad (38)$$

$$0 \le P_{ES}^{dc}(t) \le P_{ES}^{ou}(t) \quad (39)$$

$$SOC_{ES}(T) \ge SOC_{ES}(t=0) \quad (40)$$

$$0 \le P_p(t) \le P_p^{\max} \quad (41)$$

$$0 \le P_s(t) \le P_s^{\max} \quad (42)$$

Equation (35) presents the power constraint of the microgrid in which the power balance should be maintained between the generation and consumption sides. The number of the components is limited to a maximum value as shown by Equation (36). Equation (37) shows that the SOC of the battery is maintained between minimum and maximum values to avoid heavy

degradations of BES. The charging and discharging powers of BES are limited to its available input and output powers as shown by Equations (38) and (39), respectively. The SOC of BES at the end of annual operation should be equal or higher than the SOC at the beginning Equation (40). The purchased and sold powers from/to the main grid are limited to their maximum values as presented in Equations (41) and (42), respectively.

3.2 | Deep Reinforcement Learning for Sizing

The described model by Equations (1–42) indicates that the objective function is not calculated in a closed form. In fact, the rainflow cycle counting algorithm needs the future data of DOD level, the BES degradation; hence, its lifetime cannot be calculated at the beginning. This challenge hinders connecting the objective function and decision variables in a closed form. Thus, the classic approaches cannot solve the problem. To overcome this difficulty, a reinforcement learning (RL) method known as double deep Q-network (DDQN) is used in this study.

3.3 | Problem Modelling for RL

The main goal of RL is to find capacity of devices so that the $\mathrm{NPC_{tot}}$ is minimised. This objective is pursued via taking optimal action by an agent, which is artificial intelligence. Here, the sizing problem of the microgrid is converted into a suitable form of RL. Towards this goal, the RL model is formulated as an ε-greedy which includes following steps:

State Space S: The main goal is to minimise the cost function which represents the state of problem in each step. In this paper, there is only one state which is $\mathrm{NPC_{tot}}$, that is,

$$S_i = \{\mathrm{NPC_{tot}}\} \tag{43}$$

Then, by choosing the optimal actions, the state is moved from a current state to an optimal state or from a current cost to a lower cost by adopting the optimal policy.

Action Space A: Action space is a collection of applied RL-agent actions to transit from state s_t at step t to state s_{t+1} at step $t + 1$. In this study, the action vector (A_t) is defined by the number of components and can be given by:

$$A_t = \{\mathcal{N}_{PV}, \mathcal{N}_{WT}, \mathcal{N}_{ES}, \mathcal{N}_{IV}\} \tag{44}$$

where $\mathcal{N}_{PV}$, $\mathcal{N}_{WT}$, $\mathcal{N}_{ES}$ and $\mathcal{N}_{IV}$ are the number of PV, WT, BES and IVT, respectively.

Reward R: A reward signal R_t is released to evaluate the performed action in a certain state. To maximise the accumulated reward, the RL agent inclines to choose the optimal actions under policy π as follows:

$$\max_{\pi} \sum_{t'}^{\infty} \gamma^{t'} R_{t+t'} \tag{45}$$

where t' is an index variable representing the future time steps. From the current time step t, and $0 \le \gamma < 1$ is a discount factor. The term $\gamma^{t'}$ ensures that rewards received further in the future are given less weight compared to immediate rewards. This reflects the principle of time preference in reinforcement learning, where immediate rewards are usually more valuable than distant future rewards. $R_{t+t'}$ represents the reward received at the future time step $t + t'$.

Policy: The agent's behaviour in a specific environment is described by policy π, which makes a mapping from s_t to a_t.

Action-Value Function: The expected long-term reward is described by this function. The main aim is to measure the quality of executing an action under state s_t, which is as follows:

$$Q_{\pi}(a_t) = \mathbb{E}[R_t | A_t] \tag{46}$$

where $\mathbb{E}$ represents the expected value of accumulated reward and A_t is the performed action at t and can be expressed by the following relation:

$$A_t = \arg\ \min Q_{\pi}(a_t) \tag{47}$$

To maintain a precise balance between exploitation and exploration methods in the training process, the actions are selected by a ε-greedy mechanism, that is, the agent selects a random action by the probability ε and the maximum Q-value will be selected by the probability $1 - \varepsilon$. To solve the formulated problem in the form of ε-greedy, a Q-Network method is applied in this paper.

3.4 | Deep Q-Network

Q-learning is one of the RL methods, which is a model-free method and can deal with successive decision-making tasks. The main aim of the learning in Q-learning is to achieve an optimal policy π^* that can maximise the long-term profit or reward. The previous experiences, such as rewards, states and actions as well as the next states are stored in replay memory which is then used in the training process. Indeed, the data are sampled randomly from the buffer (replay memory) and fed to the train network. To avoid overfitting issues, this operation is performed in small batch sizes. One of the best classical and outstanding deep RL algorithms is Deep Q-Network (DQN). The Q-Network is applied to learn the Q-value function $Q(s_t, a_t)$ of the expected future reward. A novel target Q-Network is used in DQN which distinguishes it from the basic Q-learning algorithm by:

$$Q_{\mathrm{target}} = R_{t+1} + \gamma \max_{\alpha} (Q(S_t, a_t; \theta)) \tag{48}$$

wher, Q_{target} represents the value of target Q obtained by the equation of Bellman and θ shows the Q-Network parameters. Two different Q-Networks exist in the DQN structure: target and main Q-Networks. At every step, the target Q-Network values are updated recurrently. These updated values are the copy of the values of the main network. It is notable that if only one Q-Network is used in the model, it leads to suboptimal or delayed convergence especially when frequency of the incoming

data are high. In this case, the data for training are greatly correlated which may cause unstable target function. The DQN performs all actions which results in Q-value's overestimation, since with the iterations' number, the errors will be accumulated [29]. In this study, the overestimation problem of Q-value is overcome by using a Double DQN in which another neural network is utilised to reduce the error's impact.

3.5 | Double Deep Q-Network

DDQN algorithm effectively addresses the overestimation issue inherent in standard DQN by decoupling the action selection and evaluation processes. In the DDQN, two neural networks are utilised with similar structure as in DQN, in which the overall network moves towards stability since target value is updated frequently [29]. In ref. [30], the impact of overestimation of DDQN compared to DQN is further analysed and confirmed that DDQN consistently outperforms standard DQN in various benchmark tasks. The current action is chosen by the main Q-Network and at the same time uses the target state-action value from the target Q-Network. Therefore, at each epoch, in the present state for all possible scenarios of actions, all the action-value pairs are obtained from the main Q-Network. This value pair is updated at each epoch. Then, an operator of *argmax* is applied over the values of state-action of the possible actions. In this process, the state-action value to maximise the Q-value for that particular action is described by:

$$Q_{\text{target}} = R_{t+1} + \gamma Q\left(S_t, \operatorname*{Arg\,max}_{\alpha} Q(S_t, a_t; \theta);\ \dot{\theta}\right) \tag{49}$$

In each time step, the value corresponding to the chosen state-action pair is obtained from the target Q-Network so that the main Q-Network to be updated. Thus, this reduces the bias towards overestimation, resulting in more accurate value estimate and more stable learning. It is notable that the DDQN includes nonsequential network architecture. In such network, the layers are separated into two streams, in such a way that both subnetworks have output layers. Subnetwork one is related to the Q-value function which is to measure the value of the given state. The other subnetwork measures the benefit value of taking a special action in the current state.

$$Q(s_t, a_t; \theta, \alpha, \beta) = V(s_t; \theta, \beta) + \left(A(s_t, a_t; \theta, \alpha) - \max_{a \in |\mathcal{A}|} A(s_t, a; \theta, \alpha)\right) \tag{50}$$

where A represents the value of advantage. α and β are the vector parameters of the advantage subnetwork and state-value functions, respectively. θ is a common parameter for both networks. By combining the outputs of the first subnetwork, the Q-values are achieved, which are the base values of the state with the benefit values of the second subnetwork actions.

For special state-action pair, the Q-value is equal to summation of the value of that state which is measured from the state-value (V) and the benefit of taking that special action in that special state. Therefore, Equation (48) can be written as follows:

$$Q(S_t, a_t; \theta, \alpha, \beta) = V(S_t; \theta, \beta) + A(S_t, a_t; \theta, \alpha) \tag{51}$$

From the above equation, the Q-value can be obtained if we have the values for A and S. However, if the Q-value is known, the values of A and S cannot be obtained. In Equation (48), the last part is modified as below to enhance the stability of the algorithm.

$$Q(S_t, a_t; \theta, \alpha, \beta) = V(S_t; \theta, \beta) + A(S_t, a_t; \theta, \alpha) - \frac{1}{|\mathcal{A}|}\sum_{\widehat{a}} A(S_t, \widehat{a}; \theta, \alpha) \tag{52}$$

The Q-Network consists of three layers: input, hidden and output layers. The hyperparameters are adjusted as 1 to get the optimal weights. In time-series problems, adjusting the hyperparameters is so important to achieve optimal Q-value and long-term reward. At the end, the Q-network will be trained through minimisation of the loss function.

$$L(\theta) = \mathbb{E}\left(Q_{\text{target}} - Q(S_t, a_t; \theta)\right)^2 \tag{53}$$

Figure 3 demonstrates the application of the DDQN-based sizing of the building microgrid.

4 | Case Study and Results

This section describes the case study and the obtained numerical results for the proposed model.

4.1 | Case Study

The developed model and deep learning strategies are general and applicable to any standard case studies. In this paper, to achieve practical results, a building in Australia is selected as the grid-tied residential microgrid. The LSTM is used to forecast the data of load consumption, wind speed, solar radiation and market price for the sizing model. The mean absolute error (MAE) of the next hour forecasted data has been obtained as 1.5%, 4.2%, 2.5% and 3.5% for load, wind, radiation and price, respectively. It is notable that the MAE of the day-ahead electricity price is obtained as 9.8%. The day-ahead forecast data are generated for all 365 days of a year on an hourly basis, which means 8760 h of data. Table 2 lists the numerical data associated with the project, including grid PV, WT, BES and IVT. The project lifetime is selected as 10 years, and the microgrid building is prohibited to export more than 20 kW to the grid at any moment. The lifetime of PV, WT and IVT are 25, 20 and 10 years, respectively. The annual degradations of WT and PV are considered as 1.6% and 0.95%, respectively [26]. The BES lifetime will be decided based on the annual degradation of battery after the microgrid's operation. The replacement cost of BES,

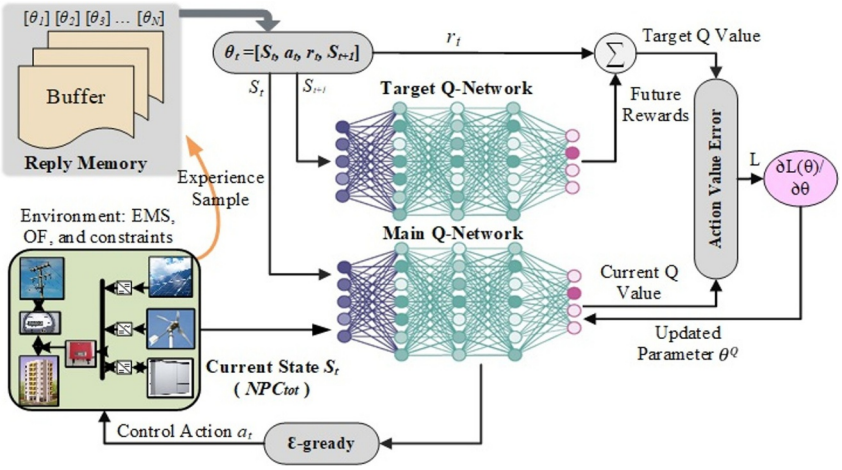

FIGURE 3 | Illustration of the DDQN-based sizing for residential microgrid.

TABLE 2 | Input data for the optimal sizing problem.

	Technical parameters	Economic parameters
Project and grid	Grid export constraint = 15 kW	Interest rate = 8%
	Grid import constraint = 20 kW	Escalation rate = 2%
PV	Unit capacity = 1 kW	Capital cost = $1200/kW
	PV lifetime = 25 years	Maintenance cost = $25/kW/yr
WT	Unit capacity = 1 kW	Capital cost = $2500/kW
	Lifetime = 20 years	Maintenance cost = $50/kW/yr
IVT	Unit capacity = 1 kW	Capital cost = $1000/kW
	Lifetime = 10 years	Inverter's efficiency = 95%
BES	Unit capacity = 0.4 kW/1 kWh	Capital cost = $500/kWh
	Minimum state-of-charge = 10%	Replacement cost = $350/kWh
	Maximum state-of-charge = 95%	Maintenance cost = $10/yr

if needed, is $350/kWh. The BES's charge/discharge efficiency is assumed as 93%, and its initial SOC is considered as 60%.

4.2 | Computational Setup and RL Training Configuration

The computational settings and reinforcement learning configuration used in the DDQN-based planning framework are provided in this subsection. The DDQN algorithm was trained using 1500 episodes with an ϵ-greedy exploration strategy, where ϵ decreases linearly from 1.0 to 0.05 with a decay factor of 0.995 per episode. The discount factor was set to $\gamma = 0.90$, whereas the replay memory contained 10,000 transitions and mini-batches of size 64 were sampled during training. The target network was updated every 50 iterations. A duelling DDQN architecture was adopted, consisting of separate value and advantage streams. The Q-network included two fully connected hidden layers with 128 and 64 neurons, respectively, using ReLU activation functions. The Adam optimiser was used for training, and the loss function was defined in ref. (53).

The proposed forecasting and planning models were implemented in MATLAB software and executed on a workstation equipped with an Intel Core i7 CPU, an NVIDIA RTX 3060 GPU (12 GB) and 32 GB of RAM. Owing to the low-dimensional action space (four sizing variables) and the deterministic structure of the planning environment, the models exhibited fast convergence. The LSTM forecasting network required approximately 10 s for training, whereas the DDQN-based planning algorithm converged in roughly 2 minute. This computational efficiency is one of the advantages of the proposed approach, making the model suitable for practical microgrid applications where rapid retraining may be required.

4.3 | Numerical Results

The results of the proposed model under RTP are compared to the following existing models:

- Simple operation under RTP: the BES is charged and discharged based on the availability of RES power after supplying the load. Hence, the RTP variation is not considered to affect the BES operation.
- Flat tariff: the purchasing and selling electricity prices are fixed at $0.48/kWh and $0.17/kWh, respectively [5]. The

renewable energy output is sold to the grid once it is higher than the load and charge request of BES.

- Time-of-use tariff: The energy price for off-peak (12a.m.–5p.m.) and peak times (6–11p.m.) are selected as \$0.31/kWh and \$0.58/kWh for purchasing, whereas the selling price is \$0.17/kWh [31]. BES is not allowed to be charged through the main grid at peak time.
- Grid only under RTP: there is no PV, WT and BES, and hence, the entire energy is supplied by the main grid under the RTP tariff.

The total NPC (\$) and LCOE ¢/kWh results of the proposed model and other existing models are indicated in Figure 4. As shown in the figure, the proposed model has achieved the lowest NPC and LCOE compared to other models. The LCOE of the proposed model is 36.26 ¢/kWh and its total NPC is \$145,811. The economic results of the grid-only system are almost double of the other models. Among the existing models, the simple EMS has closer results to the proposed model. To deeply investigate why the proposed model is more economic, all the electricity trades and components' economic results are discussed. Table 3 lists the NPC of electricity trading (NPC_{tra}) between the microgrid and the main grid obtained for the proposed model and the other existing models. It also lists the annual supply of charge of electricity (AC_C), and annual costs of purchased and sold electricity from/to the main grid, AC_P and AC_S. The annual supply of charge is constant and the same for all the considered models. Our proposed model achieves \$37,042 profit for the residential microgrid whereas the profit is \$35,289 for the simple EMS. The NPC_{tra} of electricity trading is positive for the other models which means that the annual purchased cost plus the annual supply of charge is higher than the cost of electricity sold to the grid.

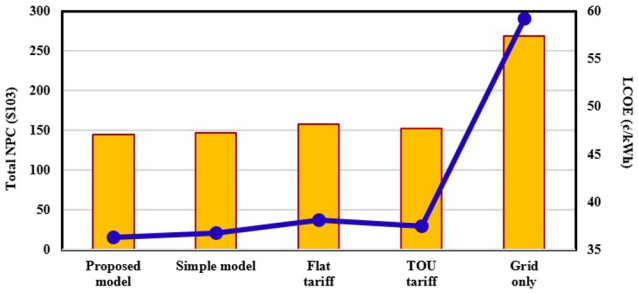

FIGURE 4 | Total NPC and LCOE obtained by the proposed and existing models.

The optimised microgrid with TOU tariff has the lowest revenue (\$1163) from selling electricity to the main grid. In the grid-only system, since all the energy is supplied by the main grid, then the NPC_{tra} is the highest among the models. The amount of purchased/sold electricity would be discussed in Table 4. Table 5 indicates the economic results of the microgrid's components. The NPC of components (NPC_{com}) for the simple sizing model is slightly higher than that of the proposed model because the NPC of BES is higher for the simple sizing model. The proposed model has the lowest NPC of BES among the other models. The model with flat tariff has the highest NPC of components. This is while that the NPC_{com} is the lowest for the model with TOU. The results of components' NPC will be more understandable when we discuss the capacity of components in the next table. Table 6 shows the optimised capacity of the microgrid components (i.e., PV, WT, BES and IVT). As shown in the table, the capacity of PV, WT and IVT is obtained the same for the proposed model and the simple EMS model. However, the capacity of BES is 1 kWh greater for the simple EMS. Although, the components have the same capacities in these two models, but the NPC_{tra} and the total NPC were lower for the proposed model as presented in Table 3. This is mainly because the developed EMS in this study works more efficiently than the simple EMS and it achieves lower cost of operation.

The model based on flat tariff has higher PV capacity compared to other models. This is while the WT capacity is almost half of the

TABLE 3 | Economic results of electricity trading between the microgrid and the main grid.

Model	$\mathbb{NPC}_{tra}$ (\$)	AC_P (\$)	AC_S (\$)	AC_C (\$)
Proposed model RTP	−37,042	7026	−15,279	3248
Simple sizing RTP	−35,289	6391	−14,407	3248
Flat tariff	18,432	6891	−7649	3248
TOU tariff	67,185	6992	−1163	3248
Grid only RTP	269,095	265,847	0	3248

TABLE 4 | Operational results of the studied models for the residential microgrid.

Model	TREG (MWh)	TCB (MWh)	TDB (MWh)	TSEG (MWh)	TPEG (MWh)	TCE (MWh)
Proposed model RTP	122.02	7.69	6.92	69.31	14.21	4.82
Simple sizing RTP	122.02	9.97	8.99	66.83	12.15	5.00
Flat tariff	93.67	11.15	10.06	44.99	14.36	0.59
TOU tariff	52.74	11.09	10.03	11.07	20.73	0
Grid only RTP	0	0	0	0	61.35	0

Abbreviations: TCB: total charging of BES; TCE: total curtailed electricity; TDB: total discharging of BES; TPEG: total purchased electricity from the grid; TREG: total renewable electricity generation and TSEG: total sold electricity to the grid.

TABLE 5 | Economic results of microgrid's components.

Model	$\mathbb{NPC}_{com}$ ($)	$\mathbb{NPC}_{PV}$ ($)	$\mathbb{NPC}_{WT}$ ($)	$\mathbb{NPC}_{ES}$ ($)	$\mathbb{NPC}_{IV}$ ($)
Proposed model RTP	182,853	40,336	76,721	15,796	50,000
Simple sizing RTP	183,301	40,336	76,721	16,247	50,000
Flat tariff	140,173	42,404	40,617	17,149	40,000
TOU tariff	85,092	26,891	18,052	17,149	23,000
Grid only RTP	0	0	0	0	0

TABLE 6 | The optimised capacities of microgrid components.

System	PV (kW)	WT (kW)	BES (kWh)	IVT (kW)
Proposed model RTP	39	34	35	50
Simple EMS under RTP	39	34	36	50
Flat tariff	41	18	38	40
TOU tariff	26	8	38	23
Grid only under RTP	0	0	0	0

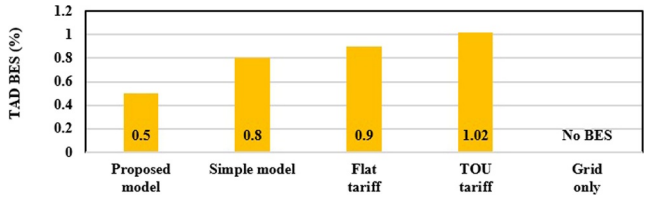

FIGURE 5 | Total NPC and LCOE obtained by the proposed and existing models.

one for the proposed model. The WT capacity is the lowest (8 kW) for the model based on TOU. This means that the WT does not achieve high profitability, compared to PV, by flat and TOU tariffs. However, the results show that higher capacity of BES is achieved for the flat and TOU tariffs which means that the BES maybe more compatible with those models. However, it should be noted that those systems have achieved much higher operation cost as indicated in Table 3. Since the total capacity of PV, WT and BES is higher for the proposed and simple EMS models, a higher IVT capacity (50 kW) is obtained for those systems compared to the Flat and TOU. The annual operational results of the studied models for the residential microgrid are shown in Table 4. This table indicates the total renewable electricity generation (TREG) by PV and WT, total charging of BES (TCB), total discharging of BES (TDB), total purchased electricity from the grid (TPEG), total sold electricity to the grid (TSEG) and total curtailed electricity (TCE) of PV and WT for all studied models. It is notable that the total electricity demand of the grid-tied microgrid is 61.35 MWh/year. As shown in the table, the TREG of the proposed and simple EMS models is 122.02 MWh/year. It is important to compare the proposed model with the simple sizing at first. Most of the generated power by PV and WT has been sold to the main grid, 69.31 MWh by the proposed model and 66.83 MWh by the simple EMS model. Hence, as obvious, the TSEG is increased by the proposed EMS. Indeed, the BES has better controlled in the proposed model.

This is the main reason for lower TCE for the proposed model compared to the simple EMS. Although the proposed model has higher TPEG (14.21 MWh), most of the energy is purchased during the low-price hours based on the developed methodology. The model based on flat tariff has high values of TCB and TDB compared to the RTP models. This is mainly because of higher capacity of BES and PV in that model. Although the flat model has 0.59 MWh curtailed energy, it is zero for the model based on TOU because of lower capacity of PV and WT. In the TOU model, a high portion of the load is purchased from the main grid (20.73 MWh) compared to the other models. Figure 5 demonstrates the total annual degradation (TAD) of BES for each model. As indicated in the figure, the proposed model has achieved the lowest TAD (0.5%/year) compared to other models. This is mainly because of the developed EMS in this study. The charging and discharging of BES are controlled based on the average of the day-ahead RTP tariff with comparison to the real-time tariff. This has decreased the charging/discharging of BES in the microgrid, and hence, lower cycles have been achieved which means less battery degradation. The model based on the TOU has the highest battery degradation due to increasing charging/discharging based on the low and high electricity prices during the day.

4.4 | Effect of Day-Ahead Forecasting Error for RTP

The MAE of the day-ahead electricity price was obtained as 9.8% by the LSTM algorithm. Considering the uncertainty in the day-ahead forecast error by forecasting, the effect of different forecast errors on the sizing problem is investigated. Figure 6 shows the total NPC and LCOE results of the proposed model for different day-ahead forecast errors of the electricity price. The minimum NPC and LCOE are shown to be obtained for forecast errors less than 2.5%, which means that high prediction accuracy is required to achieve near-optimal economic performance. The operation rules are changed on the basis of the predicted day-ahead electricity price. On the other hand, it is shown that from 2.5% to 15%, the NPC and LCOE of the microgrid increase progressively due to the accumulation of suboptimal operational

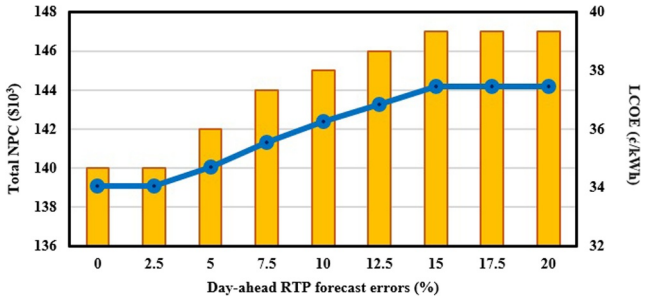

FIGURE 6 | Effects of day-ahead forecast errors of electricity price on total NPC and LCOE of the microgrid.

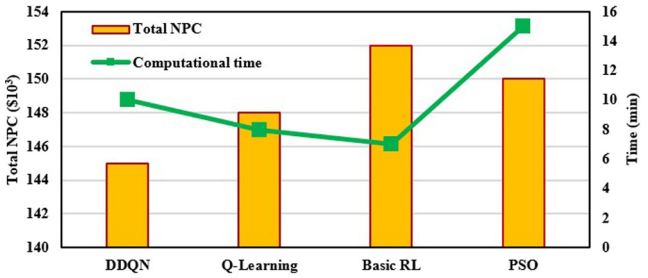

FIGURE 7 | Total NPC and computational time of different algorithms by solving the optimal sizing problem.

decisions described above. Interestingly, when the forecast error exceeds 15%, no further change in NPC and LCOE is observed, with both remaining at approximately $147,000 and 39.49 ¢/kWh, respectively. A 15% forecasting error therefore represents the worst-case condition within the investigated range as higher errors up to 20% yield identical results. Overall, the analysis confirms a strong link between forecasting accuracy and economic performance: lower forecast errors correlate with lower NPC and LCOE, whereas higher errors lead to progressively higher costs until reaching a saturation point. This emphasises the importance of accurate forecasting algorithms for achieving cost-optimal microgrid operation and sizing.

4.5 | Comparison With Other Algorithms

To demonstrate the effectiveness of the proposed DDQN algorithm, the total NPC and computational time are compared with two other machine learning methods and a benchmark metaheuristic methodology. A traditional Q-learning method and the basic RL algorithm are used for the optimal sizing problem. The PSO algorithm is selected as the metaheuristic method due to its successful implementation in power system sizing studies [4, 5]. Figure 7 shows the comparison of the methods. To overcome the issue of finding different solutions of different algorithms, the optimisation has been run for 20 times for each algorithm. Then, the best result (minimum cost) has been selected for each algorithm. As shown, the DDQN has achieved the lowest NPC compared to other methods which shows the efficacy of the DDQN method in finding the optimal solution. In terms of computational time, the Q-learning and basic RL algorithms need shorter time to achieve the results compared to DDQN. However, a 10-min computational time for the DDQN is quite reasonable for a long-term sizing study.

5 | Conclusion and Future Work

This study developed a practical and comprehensive model for the optimal sizing of a grid-tied microgrid under dynamic real-time pricing (RTP). The model incorporates all key practical parameters and solves the sizing problem using a double deep Q-network (DDQN) reinforcement learning method. The proposed approach was benchmarked against existing models and methods, demonstrating superior performance in terms of cost and efficiency.

The model achieved the lowest net present cost (NPC) of $145,811 and a levelised cost of electricity (LCOE) of 36.26 ¢/kWh, outperforming other sizing approaches. Although the component capacities were similar to those obtained through simple RTP-based sizing, the operational NPC was significantly lower, resulting in a reduced total NPC. Additionally, the model contributed to lower battery degradation, with an annual rate of just 0.5%, compared to higher degradation observed in other models, due to optimised charge/discharge cycles.

However, the performance of the model was sensitive to day-ahead forecast errors; forecast inaccuracies above 2.5% prevented the system from achieving the lowest NPC. Although the DDQN method required slightly more computational time, it consistently yielded better economic results.

Future work may include developing an experimental microgrid prototype to evaluate real-time operational performance and validate the effectiveness of the DDQN-based sizing approach under real-world conditions, including communication delays, forecast uncertainty and control variability. Further studies could also examine the role of energy aggregators, particularly how coordinated multihousehold or community-level scheduling under RTP affects optimal sizing, operational cost and system stability. In addition, the integration of electric vehicles with bidirectional charging, vehicle-to-grid (V2G), presents an important avenue for future investigation as the flexible storage capacity and mobility patterns of EVs may significantly influence the optimal sizing of distributed resources. Exploring these areas would enhance the robustness and scalability of the proposed model for broader practical deployment.

Nomenclature

Acronyms

BES	battery energy storage
CRF	capital recovery factor
DDQN	double deep Q-network
DOD	depth of discharge
DQN	Deep Q-Network
EMS	energy management system
LCOE	levelised cost of electricity
LSTM	long short-term memory

MAE	mean absolute error
ML	machine learning
NOCT	normal operating cell temperature
NPC	net present cost
PSO	particle swarm optimization
PV	photovoltaic
RCRF	real capital recovery factor
RL	reinforcement learning
RTP	real-time pricing
SOC	state-of-charge
TAD	total annual degradation
TCB	total charging of BES
TCE	total curtailed electricity
TOU	time-of-use
TPEG	total purchased electricity from the grid
TREG	total renewable electricity generation
TSEG	total sold electricity to the grid
WT	wind turbine

Parameters

Δt	Time interval (hr)
η_{IV}	Efficiency of inverter (%)
η_{PV}	Efficiency of solar PV (%)
AC_{C}	Annual supply of charge of electricity ($)
C_i^{cap}	Initial investment cost for component i ($)
C_i^{opm}	Annual operation and maintenance cost for component i ($)
C_i^{rep}	Replacement cost for component i ($)
C_{tra}	Annual electricity trading cost between the microgrid and the main grid ($)
d_{PV}	Degradation of photovoltaic system (%)
d_{WT}	Degradation of wind turbine (%)
L_i	Lifetime of component i (year)
L_{ES}^{ca}	Calendar lifetime of battery (year)
M_i	Remaining lifetime of component i (year)
P_{ES}	Nominal power of battery (kW)
P_{IV}	Nominal power of inverter (kW)
PD_i^{cap}	Present discounted value of capital cost for component i ($)
PD_i^{opm}	Present discounted value of operation and maintenance cost for component i ($)
PD_i^{rep}	Present discounted value of replacement cost for component i ($)
PD_i^{sal}	Present discounted value of salvage cost for component i ($)
R_i	Replacement year for component i (year)
$\eta_{ES}^{cg}/\eta_{ES}^{dc}$	Charging/discharging efficiency of battery (%)
P_{PV}^{rt}/P_{WT}^{rt}	Rated power of solar PV and wind turbine (kW)
RTP_P^{avg}/RTP_S^{avg}	Average of forecasted day-ahead purchase/sell real-time price (¢/kWh)
$SOC_{ES}^{min}/SOC_{ES}^{max}$	Minimum and maximum state of charge limits of battery (%)
e	Escalation rate (%)
g	Interest/discount rate (%)
l	Project lifetime (year)
u	Real interest rate (%)

Variables

N_i	Number of component i (decision variables)
N_{ES}, N_{IV}	Numbers of battery, inverter
N_{PV}, N_{WT}	Numbers of solar PV and wind turbine
AC_P/AC_S	Annual costs of purchased and sold electricity from/to the main grid ($)
d_{ES}^{tot}/d_{ES}	Total and each cycle degradations of battery (%)
E_M	Total electricity demand of the microgrid (MWh)
E_{ES}	Nominal energy of battery (kWh)
L_{ES}	Obtained real lifetime of battery (year)
$LCOE$	Levelised cost of electricity (¢/kWh)
NPC_{com}	Net present cost of the components ($)
NPC_{tot}	Total net present cost of the grid-tied microgrid ($)
NPC_{tra}	Net present cost of electricity trading between the microgrid and the main grid ($)
P_D	Curtailed power (kW)
P_M	Power demand of the microgrid (kW)
P_p/P_S	Purchased/sold power from/to the main grid (kW)
P_{ES}^{cg}/P_{ES}^{dc}	Charging/discharging power of battery (kW)
P_{ES}^{in}/P_{ES}^{ou}	Available input/output power of battery (kW)
P_{PV}	Generated power by 1-kW solar PV (kW)
P_{RE}	Actual generated power by renewable energy resources (kW)
P_{WT}	Generated power by 1-kW wind turbine (kW)
RTP_P/RTP_S	Real-time price for purchased/sold power from/to the main grid (¢/kWh)
SOC_{ES}	State of charge of battery (%)
P_P^{max}/P_S^{max}	Maximum purchased/sold power limit from/to the main grid (kW)

Author Contributions

Rahmat Khezri: conceptualisation, data curation, formal analysis, investigation, methodology. **Peyman Razmi:** conceptualisation, writing – original draft, formal analysis. **Amin Mahmoudi:** formal analysis, writing – review and editing, project administration. **Mohammad Hassan Khooban:** formal analysis, writing – review and editing, project administration.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data from this study are available upon request.

References

1. R. Khezri, A. Mahmoudi, and H. Aki, "Optimal Planning of Solar Photovoltaic and Battery Storage Systems for Grid-Connected Residential Sector: Review, Challenges and New Perspectives," *Renewable and Sustainable Energy Reviews* 153 (January 2022): 111763, https://doi.org/10.1016/j.rser.2021.111763.

2. T. Ding, M. Qu, N. Amjady, F. Wang, R. Bo, and M. Shahidehpour, "Tracking Equilibrium Point Under Real-Time Price-Based Residential Demand Response," *IEEE Transactions on Smart Grid* 12, no. 3 (May 2021): 2736–2740, https://doi.org/10.1109/tsg.2020.3040084.

3. R. Khezri and A. Mahmoudi, "Review on the State-of-the-Art Multi-Objective Optimisation of Hybrid Standalone/Grid-Connected Energy Systems," *IET Generation, Transmission & Distribution* 14, no. 20 (October 2020): 4285–4300, https://doi.org/10.1049/iet-gtd.2020.0453.

4. S. Bandyopadhyay, G. R. C. Mouli, Z. Qin, L. R. Elizondo, and P. Bauer, "Techno-Economical Model Based Optimal Sizing of PV-Battery Systems for Microgrids," *IEEE Transactions on Sustainable Energy* 11, no. 3 (July 2020): 1657–1668, https://doi.org/10.1109/tste.2019.2936129.

5. R. Khezri, A. Mahmoudi, and M. H. Haque, "Impact of Optimal Sizing of Wind Turbine and Battery Energy Storage for a Grid-Connected Household With/Without an Electric Vehicle," *IEEE Transactions on Industrial Informatics* 18, no. 9 (Sept. 2022): 5838–5848, https://doi.org/10.1109/tii.2022.3140333.

6. R. Rigo-Mariani, B. Sareni, and X. Roboam, "Integrated Optimal Design of a Smart Microgrid With Storage," *IEEE Transactions on Smart Grid* 8, no. 4 (July 2017): 1762–1770, https://doi.org/10.1109/tsg.2015.2507131.

7. M. Aghamohamadi, A. Mahmoudi, and M. H. Haque, "Two-Stage Robust Sizing and Operation Co-Optimization for Residential PV-Battery Systems Considering the Uncertainty of PV and Load," *IEEE Transactions on Industrial Informatics* 17, no. 2 (February 2021): 1005–1017, https://doi.org/10.1109/tii.2020.2990682.

8. R. Atia and N. Yamada, "Sizing and Analysis of Renewable Energy and Battery Systems in Residential Microgrids," *IEEE Transactions on Smart Grid* 7, no. 3 (May 2016): 1204–1213, https://doi.org/10.1109/tsg.2016.2519541.

9. S. Kahrobaee, S. Asgarpoor, and W. Qiao, "Optimum Sizing of Distributed Generation and Storage Capacity in Smart Households," *IEEE Transactions on Smart Grid* 4, no. 4 (December 2013): 1791–1801, https://doi.org/10.1109/tsg.2013.2278783.

10. L. Zhou, Y. Zhang, X. Lin, C. Li, Z. Cai, and P. Yang, "Optimal Sizing of PV and BESS for a Smart Household Considering Different Price Mechanisms," *IEEE Access* 6 (June 2018): 41050–41059, https://doi.org/10.1109/access.2018.2845900.

11. J. Wu, Z. Wang, C. Wu, K. Wang, and Y. Yu, "A Data-Driven Storage Control Framework for Dynamic Pricing," *IEEE Transactions on Smart Grid* 12, no. 1 (January 2021): 737–750, https://doi.org/10.1109/tsg.2020.3012124.

12. D. Zhang, H. Zhu, H. Zhang, H. H. Goh, H. Liu, and T. Wu, "Multi-Objective Optimization for Smart Integrated Energy System Considering Demand Responses and Dynamic Prices," *IEEE Transactions on Smart Grid* 13, no. 2 (March 2022): 1100–1112, https://doi.org/10.1109/tsg.2021.3128547.

13. O. Erdinc, N. G. Paterakis, T. D. P. Mendes, A. G. Bakirtzis, and J. P. S. Catalão, "Smart Household Operation Considering Bi-Directional EV and ESS Utilization by Real-Time Pricing-Based DR," *IEEE Transactions on Smart Grid* 6, no. 3 (May 2015): 1281–1291, https://doi.org/10.1109/tsg.2014.2352650.

14. A. Mohsenian-Rad and A. Leon-Garcia, "Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments," *IEEE Transactions on Smart Grid* 1, no. 2 (Sept. 2010): 120–133, https://doi.org/10.1109/tsg.2010.2055903.

15. U. B. Tayab, K. N. Hasan, R. Shah, and S. Islam, "Optimum Battery Sizing, Scheduling and Demand Management for Microgrids Using Slime Mould Algorithm," *Journal of Energy Storage* 91 (2024): 112034, https://doi.org/10.1016/j.est.2024.112034.

16. R. Elazab, A. T. Abdelnaby, H. E. Keshta, and A. A. Ali, "Optimal Techno-Economic Feasibility Analysis of a Grid-Tied Microgrid Considering Demand Response Strategy," *Electric Power Systems Research* 224 (2023): 109768, https://doi.org/10.1016/j.epsr.2023.109768.

17. R. Elazab, A. T. Abdelnaby, and A. A. Ali, "A Comparative Study of Advanced Evolutionary Algorithms for Optimizing Microgrid Performance Under Dynamic Pricing Conditions," *Scientific Reports* 14, no. 1 (2024): 4548, https://doi.org/10.1038/s41598-024-54829-9.

18. A. Priyanka, B. Sharma, and M. Rizwan, "Size Optimization of Grid-Tied Hybrid Energy System by Employing Forecasted Meteorological Data," *Mapan* 39, no. 3 (2024): 739–750, https://doi.org/10.1007/s12647-024-00758-x.

19. F. A. Kassab, B. Celik, F. Locment, M. Sechilariu, S. Liaquat, and T. M. Hansen, "Optimal Sizing and Energy Management of a Microgrid: A Joint MILP Approach for Minimization of Energy Cost and Carbon Emission," *Renewable Energy* 224 (2024): 120186, https://doi.org/10.1016/j.renene.2024.120186.

20. S. Chakraborty, G. Modi, B. Singh, B. K. Panigrahi, and M. Z. Farooqi, "A NIHOC-GI-FLL Control With Seasonal ToU Tariff Based Economic Power Regulation for a Weak Grid Tied Optimally Sized SPV-BES System," *IEEE Transactions on Industry Applications* 61, no. 3 (May–June 2025): 4414–4427, https://doi.org/10.1109/tia.2025.3542725.

21. R. Muduli and D. Das, "A New Analytical Technique for Obtaining the Optimal Sizing, Location, and Scheduling of BESS in a Grid-Connected Microgrid With Multiple Cases of Heavy DG Penetration," *International Journal of Ambient Energy* 45, no. 1 (2024): 2421995, https://doi.org/10.1080/01430750.2024.2421995.

22. A. Ahmad Khan, A. Faiz Minai, R. K. Godi, V. Shankar Sharma, H. Malik, and A. Afthanorhan, "Optimal Sizing, Techno-Economic Feasibility and Reliability Analysis of Hybrid Renewable Energy System: A Systematic Review of Energy Storage Systems' Integration," *IEEE Access* 13 (2025): 59198–59226, https://doi.org/10.1109/access.2025.3535520.

23. C. H. B. Apribowo, S. P. Hadi, F. D. Wijaya, M. I. B. Setyonegoro, and Sarjiya, "Optimal Sizing and Placement of Battery Energy Storage System for Maximum Variable Renewable Energy Penetration Considering Demand Response Flexibility: A Case in Lombok Power System, Indonesia," *Energy Conversion and Management X* 23 (2024): 100620, https://doi.org/10.1016/j.ecmx.2024.100620.

24. M. Holzbach, J. Fredy Franco Baquero, and M. Resener, "Optimal Sizing of a Hybrid Renewable Energy System for Auxiliary Services in Substations Through Genetic Algorithm and Variable Neighborhood Search," *IEEE Access* 13 (2025): 94740–94760, https://doi.org/10.1109/access.2025.3574106.

25. S. Hasanvand, M. Rafiei, M. Gheisarnejad, and M. Khooban, "Reliable Power Scheduling of an Emission-Free Ship: Multiobjective Deep Reinforcement Learning," *IEEE Trans. Transportation Electrification* 6, no. 2 (June 2020): 832–843, https://doi.org/10.1109/tte.2020.2983247.

26. R. Khezri, A. Mahmoudi, and H. Aki, "Resiliency-Oriented Optimal Planning for a Grid-Connected System With Renewable Resources and Battery Energy Storage," *IEEE Transactions on Industry Applications* 58, no. 2 (March–April 2022): 2471–2482, https://doi.org/10.1109/tia.2021.3133340.

27. W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid* 10, no. 1 (January 2019): 841–851, https://doi.org/10.1109/tsg.2017.2753802.

28. S. Merrington, R. Khezri, and A. Mahmoudi, "Optimal Sizing of Gridconnected Rooftop Photovoltaic and Battery Energy Storage for

Houses With Electric Vehicle," *IET Smart Grid* 6, no. 3 (2023): 297–311, https://doi.org/10.1049/stg2.12099.

29. X. Chen, G. Qu, Y. Tang, S. Low, and N. Li, "Reinforcement Learning for Selective Key Applications in Power Systems: Recent Advances and Future Challenges," *IEEE Transactions on Smart Grid* 13, no. 4 (July 2022): 2935–2958, https://doi.org/10.1109/tsg.2022.3154718.

30. B. Wang, Y. Li, W. Ming, and S. Wang, "Deep Reinforcement Learning Method for Demand Response Management of Interruptible Load," *IEEE Transactions on Smart Grid* 11, no. 4 (July 2020): 3146–3155, https://doi.org/10.1109/tsg.2020.2967430.

31. X. Pan, R. Khezri, A. Mahmoudi, and S. Muyeen, "Optimal Planning of Solar PV and Battery Storage With Energy Management Systems for Time-of-Use and Flat Electricity Tariffs," *IET Renewable Power Generation* 16, no. 6 (April 2022): 1206–1219, https://doi.org/10.1049/rpg2.12433.