

Research Papers

Real-time power scheduling for an isolated microgrid with renewable energy and energy storage system via a supervised-learning-based strategy

Truong Hoang Bao Huy ^a, Tien-Dat Le ^b, Pham Van Phu ^a, Seongkeun Park ^{c,*}, Daehee Kim ^{a,*}^a Department of Future Convergence Technology, Soochunhyang University, Asan-si, Chungcheongnam-do 31538, South Korea^b Department of Mobility Convergence Security, Soochunhyang University, Asan-si, Chungcheongnam-do 31538, South Korea^c Department of Smart Automobile, Soochunhyang University, Asan-si, Chungcheongnam-do 31538, South Korea

ARTICLE INFO

ABSTRACT

Keywords:

Energy storage system
Isolated microgrid
Power scheduling
Renewable energy
Supervised learning

In the future of decentralized energy systems, isolated microgrids integrated with renewable energy and energy storage systems (ESS) have emerged as critical solutions for areas beyond conventional grid connectivity. Optimal power scheduling is essential for the efficient operation, cost efficiency, and stability of isolated microgrids. Therefore, this study proposes a new supervised learning (SL) strategy for real-time optimal power scheduling of an isolated microgrid. The proposed approach is three-fold: First, a deterministic mixed-integer linear programming (MILP) model is established for the optimal power scheduling problem of an isolated microgrid to minimize operational costs. By harnessing historical data, this optimization model is solved by a dedicated MILP solver to obtain an expert dataset of optimal decisions in the isolated microgrids. Second, an SL strategy is deployed to learn and mimic optimal ESS charging/discharging decisions by training a dense residual neural network (ResNetD) on the obtained expert dataset. Finally, the well-trained ResNetD model is applied to provide near-optimal power scheduling decisions based on real-time information. The performance of the proposed method is validated using a comprehensive set of test scenarios and compared with the base case, myopic policy, and other well-known deep reinforcement learning. The results reveal that the SL method reduces operating costs by 5.95 % and the output of the diesel engine generator by 12.67 % compared to the base case. Moreover, the SL method provides high-quality solutions that closely approximate the ideal results with an average performance gap of 0.37 %. Therefore, the proposed method demonstrates its robust adaptability to the real-time conditions of an isolated microgrid environment.

1. Introduction

1.1. Background

The need for sustainable energy systems has increased owing to the increasing global energy demand, greenhouse gas emissions, and depletion of fossil fuels [1–4]. The power sector is transforming from a centralized grid to a decentralized scheme that utilizes distributed energy resources, such as renewable energy sources (RES), flexible demand management, and energy storage systems (ESS) [5,6]. One solution for addressing the challenges of centralized power grids is the deployment of isolated microgrids. These small-scale electricity systems are designed to operate independently from the upscale grid and typically consist of local power generation sources such as diesel engine generators (DEG),

solar photovoltaic (PV) panels, wind turbines, ESSs, and controllable loads [7]. Isolated microgrids are becoming increasingly popular as a cost-effective means of providing electricity to remote or off-grid locations where extending upscale grids may not be feasible. However, isolated microgrids face significant challenges due to the intermittent and uncertain nature of RESs [8,9]. Accordingly, the day-ahead and real-time scheduling results may differ due to prediction errors in RES output power and load demand [10–12]. To operate isolated microgrids optimally, an optimal power scheduling strategy is essential to take advantage of available energy sources, reduce the impact of RES intermittency, maintain stable power sources, and reduce operating costs.

1.2. Literature review

In recent years, numerous studies have focused on power scheduling

* Corresponding authors.

E-mail addresses: trhbhuy@sch.ac.kr (T.H.B. Huy), ltdat@sch.ac.kr (T.-D. Le), pvpnu@sch.ac.kr (P.V. Phu), keiny@sch.ac.kr (S. Park), daeheekim@sch.ac.kr (D. Kim).

Nomenclature	
Abbreviations	
DEG	diesel engine generators
DR	demand response
EMS	energy management system
ESS	energy storage system
FL	flexible load
IL	inflexible load
OM	operation and maintenance
Pv	photovoltaic
WG	wind generator
Indexes (Sets)	
c (C)	set of loads
g (G)	set of power sources
m (M)	set of flexible customers
i (N)	set of buses
(i, j) E	set of lines
t (Γ)	set of time intervals
Parameters	
N_B	number of buses
N_L	number of lines
$\underline{p}^{DEG}, \bar{p}^{DEG}$	minimum and maximum dispatchable powers of DEG (kW)
$\bar{p}^{ESS,ch}, \bar{p}^{ESS,dch}$	rated charging and discharging power of ESS (kW)
$\underline{p}_t^{FL,m}, \bar{p}_t^{FL,m}$	minimum and maximum load demand to be served by m^{th} flexible consumer (kW)
p_t^{IL}, q_t^{IL}	active and reactive power demand of inflexible consumer at each time interval t (kW)
p_t^{PV}	active power generated by solar PV panel at time interval t (kW)
\bar{p}^{PV}	rated active power of solar PV panel (kW)
\hat{p}_t^{PV}	available power of solar PV panel at time interval t (kW)
p_t^{WG}	active power generated by wind turbine at time interval t (kW)
\bar{p}^{WG}	rated active power of wind turbine (kW)
R_{DEG}	ramp limit of the DEG (kW)
r_{ij}, x_{ij}	resistance and reactance of line (i, j) (p.u.)
T	optimization horizon (h)
α^{WG}, β^{WG}	coefficients of wind turbine ($\text{kW} \cdot (\text{m/s})^{-3}$, -)
γ_t	wind speed at time interval t (m/s)
$\underline{\gamma}^{WG}, \bar{\gamma}^{WG}, \bar{\gamma}^{WG,*}$	cut-in, rated, cut-out wind speeds of wind turbine (m/s)
$\underline{\epsilon}_{ESS}, \bar{\epsilon}_{ESS}$	minimum and maximum allowable capacity of ESS (kWh)
$\eta_{ESS,ch}, \eta_{ESS,dch}$	charging and discharging efficiency of ESS (p.u.)
η^{PV}	conversion efficiency of solar PV panel (p.u.)
η^{WG}	conversion efficiency of wind turbine (p.u.)
θ_t	ambient temperature at time interval t (°C)
λ^m	cost of unserved energy for m^{th} flexible customer (\$/kWh)
$\mu^{PV}, \mu^{WG}, \mu^{ESS}$	operation and maintenance costs of PV, WG, and ESS (\$/kWh or \$/kWh ²)
v_t	solar irradiance at time interval t (kW/m ²)
$\omega_1^{DEG}, \omega_2^{DEG}, \omega_3^{DEG}$	cost coefficients of the DEG (\$/h, \$/kWh, \$/kWh ²)
$\Delta\tau$	time interval (h)
Variables	
$I_{ij t}$	complex current on the line (i, j) at time interval t (p.u.)
$l_{ij t}$	squared magnitude of complex current on line (i, j) at time interval t (p.u.)
$P_{ij t}, Q_{ij t}$	active and reactive power flowing from buses i to j at time interval t (kW)
$p_{j t}^c, q_{j t}^c$	active and reactive power consumption at bus j at time interval t (kW)
p_t^{DEG}, q_t^{DEG}	active and reactive power generated by the DEG at time interval t (kW)
$p_t^{ESS,ch}, p_t^{ESS,dch}$	charging and discharging power of ESS at time interval t (kW)
$p_t^{FL,m}, q_t^{FL,m}$	active and reactive power consumed by m^{th} flexible consumer after curtailment at time interval t (kW)
$p_{j t}^g, q_{j t}^g$	active and reactive power generation at bus j at time interval t (kW)
q_t^{PV}	reactive power generated by solar PV panel at time interval t (kW)
q_t^{WG}	reactive power generated by wind turbine at time interval t (kW)
u_t^{DEG}	operation status of DEG (binary)
$u_t^{ESS,ch}, u_t^{ESS,dch}$	charging and discharging status of ESS at time interval t (binary)
$v_{i t}$	squared voltage magnitude at bus i at time interval t (p.u.)
$V_{i t}$	complex voltage at bus i at time interval t (p.u.)
y_i, \bar{y}_i	lower and upper bounds on voltage magnitude at bus i (p.u.)
ϵ_t^{ESS}	SOC of ESS at time interval t (kWh)

and energy management systems (EMS) in isolated microgrids. Owing to the independent operation of the upscale grid and the unpredictable nature of RESs and load demand, maintaining load balance is critical for isolated microgrids. Thus, many approaches have been proposed for day-ahead scheduling of isolated microgrids, such as stochastic programming, robust optimization, information gap decision theory (IGDT), model predictive control (MPC), and heuristic methods. In [13], the authors developed a stochastic optimization model for day-ahead energy management of an off-grid rural microgrid. Their problem was determined as a mixed-integer nonlinear programming (MINLP), incorporating wind and PV power, pumped storage, and incentive-based demand response (DR). Moradi et al. [14] proposed optimal power scheduling for a stand-alone microgrid, considering the uncertainties of load demands and RESs. They aimed to improve the system performance while reducing fuel costs and emissions by implementing advanced dynamic programming (ADP) to plan ESSs and resource generation. In [15], the authors proposed a stochastic programming framework for the unit commitment problem of a hybrid microgrid in isolated mode. An autoregressive moving average was employed to forecast the load

demands, and the spinning reserve selection was optimized to minimize the operational costs while solving for 24-hour-ahead scheduling. In [16], a bi-level EMS was developed to enhance the interconnected operations of networked microgrids. The model introduced a new pricing scheme for power exchange and utilized scenario-based analysis with scenario reduction methods to address uncertainties. In [17], a day-ahead economic dispatch model was proposed as an MILP optimization framework with the aim of reducing energy consumption in a stand-alone water-energy microgrid system. Nourollahi et al. [18] applied a hybrid stochastic robust optimization framework to achieve optimal microgrid scheduling in both normal and resilient modes. The uncertainties were modeled in grid price, wind and PV power, and electrical loads, and DR programs were incorporated to improve the microgrid resilience. Jordehi [19] proposed an MILP model for unit commitment in combined heat and power microgrids. In [20], a risk-aware IGDT-based EMS was proposed for combined cooling, heating, and power microgrids with battery charging stations. Rana et al. [21] introduced a new peak-load-shaving algorithm for hybrid PV-ESS systems in isolated microgrids. In [22], a stochastic IGDT method was

proposed for energy management of an isolated microgrid. The model effectively accounted for uncertainties related to RESs and load demand, robustly handled component failures, and integrated operational costs as artificial uncertainties. Ferahtia et al. [23] proposed an energy management strategy for a hybrid microgrid featuring a PV generator, fuel cell system, and ESS. The strategy utilized a bald eagle search algorithm to optimize system efficiency, minimize operational costs, and safeguard the battery against over-charging and over-discharging. Tostado-Vélez et al. [24] proposed an interval-based method for the optimal scheduling of an isolated microgrid incorporating DR programs and green hydrogen storage. The model employed MILP and logical rules for hydrogen generation while handling uncertainties with an interval formulation, supporting both pessimistic and optimistic strategies. Manzano et al. [25] proposed an economic MPC system for microgrid operation that utilized the actual load demand and RESs data to enhance economic performance through a predictive approach. Zhang et al. [26] introduced a particle swarm optimization approach to optimize dispatching in an isolated microgrid while considering both economic and environmental costs as objective functions. Asri et al. [27] presented a new MILP model for shared ESSs in island microgrids serving remote communities. This model enables consumers to share their ESSs with their nearest neighbors. In [28], the authors proposed an IGDT model for an off-grid system utilizing wind, bio-waste, stationary, and mobile storage to ensure a full RES supply, optimizing construction and maintenance costs. Although significant progress has been made in energy management for isolated microgrids, these methods are mainly applied for day-ahead scheduling or long-term planning, and they rely heavily on modeling and forecasting uncertainties. The scheduling results are significantly impacted by the accuracy of the mathematical or predictive models, which presents challenges in forecasting or understanding the statistical distribution information of uncertainties [29]. Forecast errors of uncertainties can force actual operating decisions to be re-optimized in the intraday online optimization process according to updated short-term forecast information. Therefore, these approaches face challenges in developing effective energy management models that can adapt to unpredictable changes in real-time situations.

The limitations of relying solely on mathematical models have sparked interest in implementing reinforcement learning (RL) and deep reinforcement learning (DRL) algorithms for energy management in stand-alone microgrids. These algorithms effectively handle stochastic decision-making problems by reducing reliance on modeling or forecasting uncertainties. In the RL and DRL algorithms, an agent learns to make a sequence of decisions by interacting with a given environment. The agent receives feedback in the form of rewards and seeks a policy that maximizes the total expected reward over time [30]. Using a trial-and-error approach, an agent can adapt its decision-making policy based on past experiences and receive rewards, providing a reliable mechanism for handling uncertainties and probabilistic outcomes [31]. In [32], the authors proposed an EMS that used adaptive power pinch analysis and Q-learning to address the uncertainties of load demand and RESs in a stand-alone microgrid. Li et al. [33] presented a new approach called prioritized experience replay automated reinforcement learning (PER-AutoRL) for optimal scheduling in isolated microgrid management. This method utilized a deep deterministic policy gradient (DDPG)-based forecasting model to address error accumulation in multi-step forecasting. In [34], the authors developed a forecast-driven optimization strategy for an isolated hydrogen microgrid. They utilized BiLSTM-CNN for 24-hour wind and load predictions and implemented a deep Q-network (DQN) to determine the optimal scheduling, considering uncertainty and storage degradation. Khawaja et al. [35] applied Q-learning to develop an EMS for a stand-alone microgrid with PV panels and an ESS to minimize operational costs. The above-cited studies primarily focused on applying the RL and DRL algorithms for day-ahead scheduling. In recent years, online power scheduling for isolated microgrids has been attracting interest. Kofinas et al. [36] proposed a distributed multi-agent system for a stand-alone solar microgrid using

fuzzy Q-learning in real-time, wherein agents represent microgrid components and coordinate behavior through shared state variables for efficient control. Totaro et al. [37] proposed a model-based RL algorithm for controlling long-term isolated microgrids. They formulated the control problem as a Markov decision process (MDP) and utilized proximal policy optimization (PPO) as the learning algorithm. Lei et al. [38] developed two DRL algorithms, namely, the finite-horizon recurrent deterministic policy gradient (FH-RDPG) and finite-horizon deep deterministic policy gradient (FH-DDPG), to control the power dispatch in IoT-enabled isolated microgrids. Xia et al. [39] proposed a multi-agent DRL algorithm to solve the economic frequency control problem in isolated networked microgrids. During the offline training stage, a soft actor-critic (SAC) algorithm was employed, where each agent was responsible for controlling the ESS and generator in each microgrid. In [40], a distributed quantum multi-agent deep meta-deterministic policy gradient (DQMA-DMDPG) algorithm was designed to aid in the area autonomy energy management of a multi-area microgrid to minimize the total generation costs and reduce frequency errors. In [41], the SAC algorithm was utilized for island microgrid energy management to regulate the stochastic behavior of RESs. Their study aimed to decrease the reliance on DEG, manage peak loads, and minimize gas emissions. Although the RL and DRL algorithms are valuable tools for making stochastic decisions, they have certain limitations that can affect the power scheduling problem. Since RL algorithms typically represent policy or value functions using basic functions or tables, they may struggle to handle continuous states and actions [42]. DRL algorithms can consume significant computational resources, particularly when training extensive neural networks. Furthermore, the agent may have difficulty adapting to new scenarios. Overfitting during training can also hinder the performance in unseen scenarios, thus limiting its applicability in real-world problems [29].

1.3. Research gaps and motivations

Table 1 summarizes recent research studies on real-time power scheduling for isolated microgrids, which shows that this topic has attracted much interest. Based on the literature review, the previous studies still have certain drawbacks, which can be outlined as follows:

- Conventional methodologies such as stochastic optimization, robust optimization, interval optimization, IGDT, and heuristic methods for day-ahead scheduling rely significantly on the modeling and forecasting uncertain parameters. If these predictions are inaccurate or the uncertainty distribution information is flawed, suboptimal solutions can be obtained. Accordingly, the predetermined schedule may need to be adjusted or even re-dispatched.
- Several researchers have attempted to solve power scheduling problems in isolated microgrids using RL and DRL algorithms.

Table 1

Summary of the reviewed references and the current study on real-time power scheduling for isolated microgrids.

Reference	DEG	PV	WG	DR program	AC power flow modeling	Approach
[25]	✓	✓	✓	–	–	MPC
[36]	✓	✓	–	–	–	fuzzy Q-Learning
[37]	✓	✓	–	–	–	PPO
[38]	✓	✓	–	–	–	FH-DDPG, FH-RDPG
[39]	✓	✓	–	–	–	SAC
[40]	✓	–	✓	–	–	DQMA-DMDPG
[41]	✓	✓	✓	–	✓	SAC
This study	✓	✓	✓	✓	✓	Supervised learning (RestNetD)

However, these algorithms typically learn from scratch through a trial-and-error process, which may require a high computational time. In addition, they may encounter problems of slow convergence, non-convergence, or instability.

- A real-time power scheduling problem for microgrids requires modeling the microgrid components and power flow in power networks. Many previous studies have ignored the power flow model for simplicity and only considered the simple supply and demand balance model. AC power flow is necessary to accurately represent reactive power requirements, nodal voltage limits, and line transmission limits in microgrids. Since the power flow model contains nonlinear constraints, some research works [13,16,18,41] applied DC power flow or formulated the optimization problem as MINLP problems. Mathematically, global optimality may not be guaranteed in the MINLP formulation.

Based on the above research gaps, the following research questions can be posed: (1) how to develop a real-time power scheduling framework for isolated microgrids that adapts to operating conditions without relying heavily on forecasting and uncertain parameter modeling?; (2) how to propose a new data-driven approach to learn and mimic expert decision-making in the optimal power scheduling of isolated microgrids, minimizing reliance on trial-and-error processes inherent in RL and DRL algorithms?; (3) how to model the power flow in isolated microgrids to ensure accuracy in representing AC power flows while maintaining problem linearity and ensuring global optimality in the solutions?

1.4. Research contributions

To fill these research gaps, this study aims to develop a new supervised learning (SL) strategy for the real-time optimal power scheduling of an isolated microgrid integrated with RESs and ESS to minimize operational costs, as shown in Fig. 1. This method is based on directly learning a policy from expert demonstrations inspired by the behavioral cloning (BC) approach. To accomplish this, the power scheduling problem of an isolated microgrid is initially defined as an MILP optimization framework to minimize operational costs, in which DistFlow

equations are applied to model AC power flows in the power network. Given the historical data, the original power scheduling problem is solved as a deterministic day-ahead problem using a dedicated MILP solver to obtain an expert dataset of optimal decisions in the isolated microgrids. Subsequently, the proposed method learns and mimics optimal ESS charging/discharging decisions by training a dense residual neural network (ResNetD) [43] on the obtained expert dataset. A well-trained ResNetD is applied to provide near-optimal power scheduling decisions in real-time scenarios. Based on the literature review, this study is the first to develop a real-time power scheduling strategy for an isolated microgrid based on an SL approach. The contributions of this research are as follows:

- The optimal power scheduling of an isolated microgrid is formulated as an MILP optimization model that integrates detailed modeling of microgrid components and AC power flow constraints under DR programs. Based on MILP formulation, the global optimality and feasibility of solutions can be guaranteed. The original optimization problem is then converted into a sequence of multi-step optimization sub-problems which are solved sequentially over the optimization horizon.
- A SL method is developed for real-time power scheduling in the isolated microgrid, wherein a ResNetD model is trained on state-action pairs extracted from solving the MILP optimization problem. In this way, the proposed method has an advantage over DRL methods by learning from an expert dataset instead of learning from scratch. This improves training efficiency while significantly reducing training time.
- Real-time power scheduling is performed in two phases at each time interval: (1) A well-trained ResNetD model is applied to determine the charging/discharging decisions of the ESS; (2) The remaining decision variables can be optimally determined by solving a one-step optimization problem. This approach significantly reduces the number of actions during training and reduces reliance on forecast information while fully satisfying microgrid and AC power flow constraints.
- The effectiveness of the proposed method is validated through a set of test scenarios using real-world data. The results show that the proposed method achieves operation costs close to the ideal results obtained under perfect information with an average performance error of 0.37 %. This implies that the proposed method effectively learns and imitates the MILP solver to make optimal decisions based on real-time information.
- The obtained results from the proposed method are compared to those from the myopic policy and other prominent DRL algorithms, including PPO, SAC, and DDPG. Comparative results showed that the proposed method provided better results than other methods, which confirms its superior performance and adaptability to the real-time power scheduling problem.

The remainder of this paper is organized as follows. Section 2 presents the system modeling and problem formulation. Section 3 describes the proposed methodology. Simulation results are presented in Section 4, and Section 5 summarizes the conclusions.

2. System modeling and problem formulation

Fig. 2 shows a schematic diagram of the proposed isolated microgrid, which is assumed not to be connected to the upscale grid. Thus, the electricity demand of the isolated microgrid is powered by DEG and renewable sources, including a solar PV panel and wind generator (WG). In addition, an ESS is employed to store the excess energy generated by renewable sources. This stored energy can be utilized during periods of low RES generation or peak demand. The isolated microgrid operates under DR programs and considers the following two types of consumers:

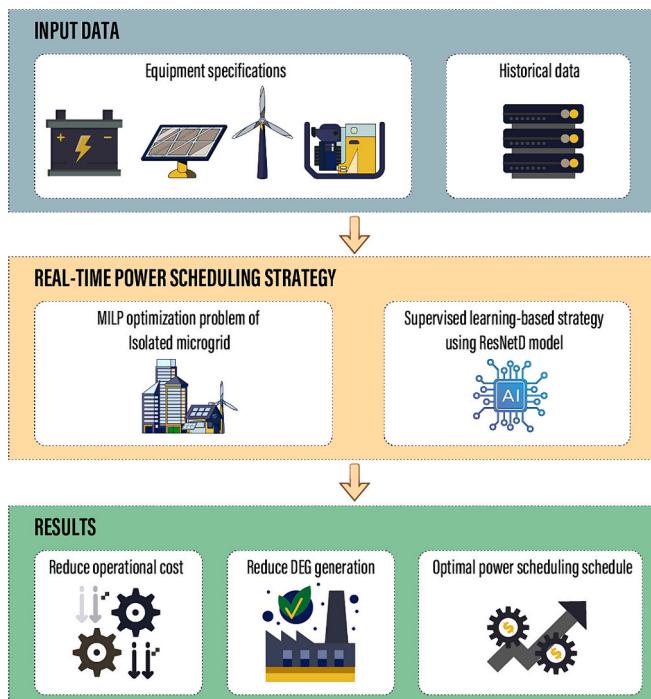


Fig. 1. The overview of the proposed framework.

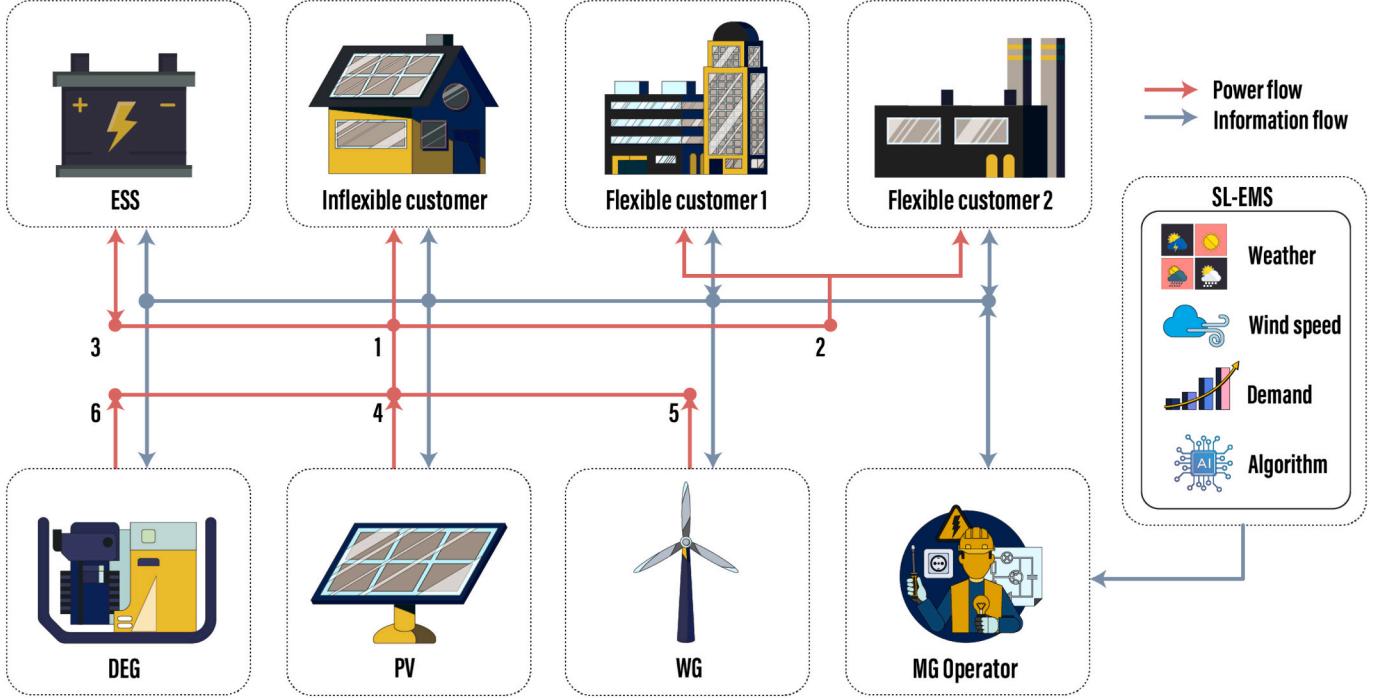


Fig. 2. Schematic diagram of the proposed isolated microgrid.

- Inflexible consumers: They do not participate in DR programs, and their expected demand must be fully served [22].
- Flexible consumers: Their expected demand can only be partially served, and they are compensated for each kilowatt hour of curtailed energy (\$/kWh) [22].

In an isolated microgrid, the operations of various electrical components and local customers are managed centrally. A communication network is deployed to connect the signals between the generators, consumers, ESS, and system operator. The system operator employs an EMS to gather real-time data and implements a real-time power scheduling strategy to minimize the operational costs of the isolated microgrid. The scheduling signals are then transmitted to the electrical components for optimal real-time operation.

2.1. System modeling

The power scheduling for the isolated microgrid is considered over a finite time horizon of 24 h ($T = 24$) and a time interval of 1 h ($\Delta\tau = 1$), where set $\Gamma = \{\Delta\tau, 2\Delta\tau, \dots, T\}$. Detailed mathematical formulations of each component of the isolated microgrid are presented in the following subsections.

2.1.1. DEG modeling

The power generated by the DEG is represented by the operational limits in (1) and the ramp restrictions in (2) [22,44]:

$$\underline{p}_t^{DEG} \cdot u_t^{DEG} \leq p_t^{DEG} \leq \bar{p}_t^{DEG} \cdot u_t^{DEG}; \quad \forall t \in \Gamma \quad (1)$$

$$p_{t-1}^{DEG} - R^{DEG} \leq p_t^{DEG} \leq p_{t+1}^{DEG} + R^{DEG}; \quad \forall t \in \Gamma \setminus \{t > 1\} \quad (2)$$

where p_t^{DEG} is the active power generated by the DEG at interval t ; \underline{p}_t^{DEG} and \bar{p}_t^{DEG} are the minimum and maximum dispatchable power of the DEG, respectively; u_t^{DEG} is the binary variable indicating the operation status of the DEG at interval t , and R^{DEG} is the ramp limit of the DEG.

2.1.2. Solar PV panel modeling

For a solar PV array, the available power of the PV system (\hat{p}_t^{PV}) can be modeled as a function of solar irradiance and ambient temperature as follows [22]:

$$\hat{p}_t^{PV} = \bar{p}^{PV} \cdot [0.25 \cdot v_t + 0.03 \cdot v_t \cdot \theta_t + (1.01 - 1.13 \cdot \eta^{PV}) \cdot v_t^2]; \quad \forall t \in \Gamma \quad (3)$$

where \bar{p}^{PV} and η^{PV} are the rated active power and conversion efficiency of the solar PV panel, respectively; v_t is the solar irradiance at interval t , and θ_t is the ambient temperature at interval t .

The active power output of the solar PV panel must not exceed the peak power, as expressed in (4).

$$p_t^{PV} = \begin{cases} \bar{p}^{PV} & \text{if } \hat{p}_t^{PV} > \bar{p}^{PV} \\ \hat{p}_t^{PV} & \text{otherwise} \end{cases}; \quad \forall t \in \Gamma \quad (4)$$

where p_t^{PV} is the active power generated by the PV panel at interval t , respectively.

2.1.3. WG modeling

Fig. 3 depicts a wind speed curve representing the correlation between the wind speed and the power generated by the wind turbine. The

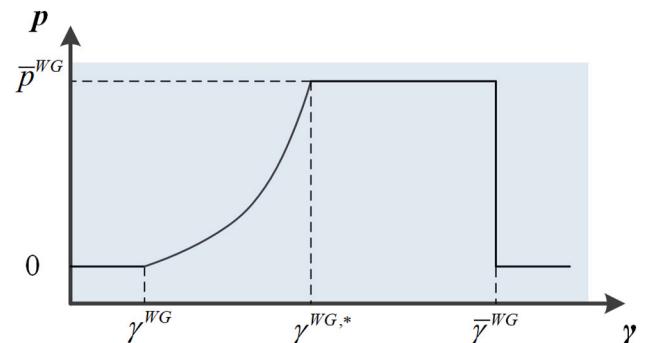


Fig. 3. Wind-power curve of wind turbine.

active power output of a WG can be defined as a function of wind speed as follows [24,44]:

$$p_t^{WG} = \eta^{WG} \cdot \begin{cases} 0, & \text{if } \gamma_t < \underline{\gamma}^{WG} \\ \alpha^{WG} \cdot \gamma_t^3 - \beta^{WG} \cdot \bar{p}^{WG}, & \text{if } \underline{\gamma}^{WG} \leq \gamma_t \leq \gamma^{WG,*} \\ \bar{p}^{WG}, & \text{if } \gamma^{WG,*} < \gamma_t \leq \bar{\gamma}^{WG}; \\ 0, & \text{if } \gamma_t > \bar{\gamma}^{WG} \end{cases} \quad \forall t \in \Gamma \quad (5)$$

where p_t^{WG} is the active power generated by the wind turbine at interval t ; \bar{p}^{WG} and η^{WG} are the rated active power and conversion efficiency of the wind turbine, respectively; γ_t is the wind speed at interval t ; $\underline{\gamma}^{WG}$, $\gamma^{WG,*}$, and $\bar{\gamma}^{WG}$ are the cut-in, rated, and cut-out wind speeds of the wind turbine, respectively; α^{WG} and β^{WG} are coefficients of the wind turbine.

2.1.4. ESS modeling

ESSs play a crucial role in maintaining a reliable and consistent power supply within an isolated microgrid. The charging and discharging power of an ESS are limited by the rated power, as expressed in (6) and (7), respectively. Moreover, constraint (8) forces charging and discharging modes to be mutually exclusive [45].

$$0 \leq p_t^{ESS,ch} \leq \bar{p}^{ESS,ch} \cdot u_t^{ESS,ch}, \quad \forall t \in \Gamma \quad (6)$$

$$0 \leq p_t^{ESS,dch} \leq \bar{p}^{ESS,dch} \cdot u_t^{ESS,dch}; \quad \forall t \in \Gamma \quad (7)$$

$$0 \leq u_t^{ESS,ch} + u_t^{ESS,dch} \leq 1; \quad \forall t \in \Gamma \quad (8)$$

where $p_t^{ESS,ch}$ and $p_t^{ESS,dch}$ is the charging and discharging power of the ESS at interval t , respectively; $\bar{p}^{ESS,ch}$ and $\bar{p}^{ESS,dch}$ are the rated charging and discharging power of the ESS, respectively; $u_t^{ESS,ch}$ and $u_t^{ESS,dch}$ are the binary variables representing the charging and discharging status of the ESS at interval t , respectively.

The state of charge (SOC) of the ESS is modeled as a function of the power exchange with the system, as given in (9). Moreover, constraint (10) shows that the SOC of the ESS is limited by the depth-of-discharge and rated capacity [44,46].

$$\epsilon_t^{ESS} = \epsilon_{t-1}^{ESS} + \Delta\tau \cdot \left[\eta^{ESS,ch} \cdot P_t^{ESS,ch} - \frac{P_t^{ESS,dch}}{\eta^{ESS,dch}} \right]; \quad \forall t \in \Gamma \quad (9)$$

$$\underline{\epsilon}^{ESS} \leq \epsilon_t^{ESS} \leq \bar{\epsilon}^{ESS}; \quad \forall t \in \Gamma \quad (10)$$

where ϵ_t^{ESS} is the SOC of the ESS at interval t ; $\eta^{ESS,ch}$ and $\eta^{ESS,dch}$ is the charging and discharging efficiency of the ESS, respectively; $\underline{\epsilon}^{ESS}$ and $\bar{\epsilon}^{ESS}$ are the minimum and maximum allowable capacity of the ESS, respectively.

It is assumed that the SOC of the ESS reaches the maximum value at the first and last intervals of the scheduling cycle as follows [22]:

$$\epsilon_1^{ESS} = \epsilon_T^{ESS} = \bar{\epsilon}^{ESS} \quad (11)$$

2.1.5. Inflexible and flexible consumers modeling

DR programs are mechanisms to adjust the power demand instead of adjusting the supply, which can be categorized into two main types: price-based DR programs and incentive-based DR programs. Price-based DR programs influence customer electricity usage by varying the price of electricity over time. This strategy encourages consumers to reduce or shift their electricity usage during peak demand periods to off-peak times when electricity is cheaper and more abundant. Meanwhile, incentive-based DR programs provide financial incentives to customers to encourage them to reduce their electricity consumption voluntarily during periods of high demand or when grid reliability is at risk. This study deploys incentive-based DR programs in isolated microgrid modeling.

In this way, two types of consumers are considered in the isolated microgrid: inflexible and flexible customers. For the inflexible consumer, the active and reactive power demand at each time interval t are represented by p_t^{IL} and q_t^{IL} , which must be fully met by the operator. Meanwhile, flexible consumers are willing to reduce a certain amount of their expected demand according to incentive-based DR programs [47]. Thus, compensation (\$/kWh) is paid to customers if the operator does not fully satisfy their expected demand. However, the curtailable power must be within a certain limit. Thus, flexible consumers are modeled as follows:

$$p_t^{FL,m} \leq p_t^{FL,m} \leq \bar{p}_t^{FL,m}; \quad \forall t \in \Gamma \wedge m \in M \quad (12)$$

where $p_t^{FL,m}$ and $\bar{p}_t^{FL,m}$ are the minimum and maximum load demands of the m^{th} flexible consumer that must be served at interval t , respectively; $p_t^{FL,m}$ is the load demand of the m^{th} flexible consumer after curtailment at interval t , and M is the set of flexible customers. This study considers one inflexible customer and two flexible customers.

2.2. MILP optimization framework for the proposed isolated microgrid

2.2.1. Power flow constraints

This study applies the DistFlow equations to model steady-state power flows in the power network. The proposed power flow model fully considers active and reactive power (AC), branch current limits, and node voltage limits while they essentially maintain the linearity of the equations. Therefore, DistFlow equations are widely applied in small-scale radial distribution networks, and their accuracy and reliability have been thoroughly validated in numerous studies [48,49]. Given a power distribution network where $N = \{1, 2, \dots, N_B\}$ is a set of buses and $E = \{1, 2, \dots, N_L\}$ is a set of lines, the DistFlow equations can be expressed as follows [48,50]:

$$P_{ij|t} = \left(P_{j|t}^c - P_{j|t}^g \right) + r_{ij} l_{ij|t} + \sum_{kj=k} P_{jk|t}; \quad \forall j \in N \wedge (i,j) \in E \wedge t \in \Gamma \wedge g \in G \wedge c \in C \quad (13)$$

$$Q_{ij|t} = \left(Q_{j|t}^c - Q_{j|t}^g \right) + x_{ij} l_{ij|t} + \sum_{kj=k} Q_{jk|t}; \quad \forall j \in N \wedge (i,j) \in E \wedge t \in \Gamma \wedge g \in G \wedge c \in C \quad (14)$$

$$v_{j|t} = v_{i|t} - 2(r_{ij} P_{ij|t} + x_{ij} Q_{ij|t}) + (l_{ij}^2 + x_{ij}^2) l_{ij}; \quad \forall i, j \in N \wedge (i,j) \in E \wedge t \in \Gamma \quad (15)$$

$$v_i \leq |V_{i|t}|^2 \leq \bar{v}_i; \quad \forall i \in N \wedge t \in \Gamma \quad (16)$$

$$l_{ij|t} = \frac{P_{ij|t}^2 + Q_{ij|t}^2}{v_{i|t}}; \quad \forall i \in N \wedge (i,j) \in E \wedge t \in \Gamma \quad (17)$$

where N_B is the number of buses, N_L is the number of lines, $P_{ij|t}$ and $Q_{ij|t}$ are the active and reactive power flowing from buses i to j at time interval t , respectively; r_{ij} and x_{ij} are the resistance and reactance of line (i, j) , respectively; $p_{j|t}^g$ and $q_{j|t}^g$ are the active and reactive power generation at bus j at time interval t , respectively; $p_{j|t}^c$ and $q_{j|t}^c$ are the active and reactive power consumption at bus j at time interval t , respectively; set $G = \{DEG, PV, WG, ESS^{dch}\}$ and set $C = \{IL, FL1, FL2, ESS^{ch}\}$; $v_{i|t}$ is the squared voltage magnitude at bus i at time interval t , and $l_{ij|t}$ is the squared magnitude of complex current on line (i, j) at time interval t . In the power flow model, $v_{i|t} = |V_{i|t}|^2$ and $l_{ij|t} = |I_{ij|t}|^2$, where $V_{i|t}$ and $I_{ij|t}$ are the complex voltage at bus i and the complex current on the line (i, j) at time interval t , respectively. The voltage magnitudes must satisfy Eq. (16), where \underline{v}_i and \bar{v}_i are given lower and upper bounds on voltage magnitudes, respectively. To formulate the optimization problem as a

convex problem, nonlinear quadratic equalities in (17) are relaxed as the following second-order cone constraint [51]:

$$\left\| \begin{array}{l} 2P_{ij|t} \\ 2Q_{ij|t} \\ l_{ij|t} - v_{ij|t} \end{array} \right\| \leq l_{ij|t} + v_{ij|t}; \quad \forall i \in N \wedge (i,j) \in E \wedge t \in \Gamma \quad (18)$$

2.2.2. Operational cost

In this study, the operational cost of an isolated microgrid at time interval t is defined as follows:

$$F_t = F_t^{DEG} + F_t^{OM} + F_t^{DR}; \quad \forall t \in \Gamma \quad (19)$$

In (19), the first term represents the fuel and maintenance costs associated with the DEG. The DEG costs can be defined as a quadratic function of the power generated by the DEG as follows [15,22]:

$$F_t^{DEG} = \Delta\tau \cdot \left[u_t^{DEG} \cdot \omega_1^{DEG} + p_t^{DEG} \cdot \omega_2^{DEG} + (p_t^{DEG})^2 \cdot \omega_3^{DEG} \right]; \quad \forall t \in \Gamma \quad (20)$$

where ω_1^{DEG} , ω_2^{DEG} , and ω_3^{DEG} are the cost coefficients of the DEG.

The second term in (19) represents the costs associated with the operation and maintenance of the PV, WG, and ESS. The cost of renewable generators can be considered proportional to the generated power. Meanwhile, the costs associated with an ESS can be formulated as a quadratic function of the ESS charging/discharging power. The second term in (19) is expressed as follows [22,52]:

$$F_t^{OM} = \Delta\tau \cdot \left[\mu^{PV} \cdot p_t^{PV} + \mu^{WG} \cdot p_t^{WG} + \mu^{ESS} \cdot (p_t^{ESS,i})^2 \right]; \quad \forall t \in \Gamma \wedge i \in \{ch; dch\} \quad (21)$$

where μ^{PV} , μ^{WG} , and μ^{ESS} are the operation and maintenance costs related to the PV, WG, and ESS, respectively. To preserve the MILP formulation of the model, the quadratic terms in (20) and (21) are linearized using a piecewise linear approximation, as presented in Appendix A.

As previously discussed, flexible consumers receive compensation when the operator reduces their expected load demands. The cost of applying DR programs is considered in the last terms of (19), which can be determined as follows [22]:

$$F_t^{DR} = \Delta\tau \cdot \sum_{m \in M} \{ \lambda^m \cdot (\bar{p}_t^{FL,m} - p_t^{FL,m}) \}; \quad \forall t \in \Gamma \quad (22)$$

where λ^m is the cost of the DR programs for the m^{th} flexible customer.

2.2.3. The optimization problem

In this study, the optimal power scheduling problem involves determining the optimal schedules of the DEG, RESs, load curtailments, and ESS to meet the energy demand of the isolated microgrid at any given time while considering the DR program and various system constraints. The main objective of the optimization problem is to minimize the operational cost of the isolated microgrid. To this end, solar PV and wind energies are prioritized for maximum utilization due to their operational costs and environmental benefits. The DEG serves as a dispatchable source that can be controlled to fill the gaps when renewable generation is insufficient. The charging and discharging schedules of the ESS are optimized to store excess renewable energy and discharge it during times of insufficient generation or peak demand. Moreover, an incentive-based DR program is implemented to adjust the load profile according to the availability of generation sources. Accordingly, incentives are paid to encourage flexible customers in load curtailment during peak times or when renewable generation is low. Mathematically, the optimal power scheduling problem is stated as follows:

$$P1 : \min_{x_t, \forall t \in \Gamma} F = \sum_{t=1}^T F_t \quad (23)$$

$$\text{s.t. : (1) - (22)}$$

where the vector of decision variables is outlined as follows:

$$x_t = \left\{ \begin{array}{l} p_t^{DEG}, p_t^{ESS,ch}, p_t^{ESS,dch}, p_t^{FL,m}, \epsilon_t^{ESS} \\ u_t^{DEG}, u_t^{ESS,ch}, u_t^{ESS,dch}, q_t^{DEG}, q_t^{PV} \\ q_t^{WG}, q_t^{FL,m}, P_{ij|t}, Q_{ij|t}, v_{ij|t}, l_{ij|t} \end{array} \right\}; \quad \forall i \in N \wedge (i,j) \in E \wedge t \in \Gamma \wedge m \in M \quad (24)$$

where q_t^{DEG} , q_t^{PV} , and q_t^{WG} are reactive power generated by the DEG, PV panel, and wind turbine, respectively; and $q_t^{FL,m}$ is the reactive power consumed by the m^{th} flexible load.

The optimization problem is subjected to the various constraints of the isolated microgrid. The power and ramp limitations of the DEG are given in (1) and (2), respectively. The maximum power generations of the solar PV panel and wind turbine are modeled in (3), (4), and (5). The ESS charging/discharging constraints are defined in (6), (7), and (8). The SOC of the ESS is modeled by (9) and limited by (10), while the initial and final SOCs are fixed in (11). The flexible customer modeling is defined by (12) considering the DR program. Finally, AC power flow constraints are considered in (13)–(18) for a comprehensive modeling. All constraints must be satisfied for all time intervals.

3. The proposed methodology

In problem **P1**, there are several difficulties owing to uncertain future information, such as load demand and renewable generation, which cannot be accurately known in advance. Conventionally, the values found by a given predictor for these uncertain variables can be used to solve this problem. However, forecasting errors are inevitable and may affect optimal results [42,53–55]. Therefore, a new real-time power scheduling is developed using a supervised learning (SL) strategy to minimize the operational costs of the isolated microgrid. This strategy formulates the optimal power scheduling as a sequential decision problem in an unknown environment with uncertainties, wherein the decision variables are partially converted to action variables and determined online based on real-time information in the state variables. The state vector at time interval t (s_t) contains state variables that characterize the current state of the isolated microgrid, including the current time interval (t), power output of the solar PV panel (p_t^{PV}), power output of the wind turbine (p_t^{WG}), load demand of the inflexible consumer (p_t^{IL}), expected load demands of flexible customers (p_t^{FL1} and p_t^{FL2}), and SOC of the ESS at the previous interval (ϵ_{t-1}^{ESS}). Thus, the state vector can be defined as follows:

$$s_t = [t, p_t^{PV}, p_t^{WG}, p_t^{IL}, \bar{p}_t^{FL1}, \bar{p}_t^{FL2}, \epsilon_{t-1}^{ESS}] \quad (25)$$

A simple approach is to employ the optimization decision variables in (24) as action variables. However, it is more challenging for SL models to learn all actions when the number of actions is large over a long horizon [42]. Moreover, the learned actions may not satisfy all constraints. A more logical approach is to reduce the number of actions by choosing a characteristic variable of the power scheduling problem, which is the charging/discharging power of the ESS. Accordingly, the decision variables in (24) are defined into two steps. In the first step, the charging/discharging power of the ESS at each time interval t is determined directly by the SL model based on the state observed in (25), which is assumed as follows:

$$\begin{aligned} p_t^{ESS,ch} &= \hat{p}_t^{ESS,ch} \\ p_t^{ESS,dch} &= \hat{p}_t^{ESS,dch} \\ u_t^{ESS,ch} &= \hat{u}_t^{ESS,ch} \\ u_t^{ESS,dch} &= \hat{u}_t^{ESS,dch} \\ \epsilon_t^{ESS} &= \hat{\epsilon}_t^{ESS} \end{aligned} \quad (26)$$

where $\hat{p}_t^{ESS,ch}$, $\hat{p}_t^{ESS,dch}$, $\hat{u}_t^{ESS,ch}$, $\hat{u}_t^{ESS,dch}$, and $\hat{\epsilon}_t^{ESS}$ are predicted values of

charging power, discharging power, charging status, discharging status, and SOC of the ESS, respectively.

In the second step, the one-step optimization problem at each time interval t is performed to determine the remaining decisions as follows:

$$\begin{aligned} P2 : \quad & \min_{x'_t} F_t \\ \text{s.t. : } & (1) - (5), (12) - (22), (26) \end{aligned} \quad (27)$$

where x'_t includes the remaining decision variables in (24) as follows:

$$x'_t = \left\{ p_t^{DEG}, p_t^{FL,m}, u_t^{DEG}, q_t^{DEG}, q_t^{PV}, q_t^{WG}, q_t^{FL,m}, P_{ij|t}, Q_{ij|t}, v_{i|t}, l_{ij|t} \right\}; \quad \forall i \in N \wedge (i,j) \in E \wedge t \in \Gamma \wedge m \in M \quad (28)$$

The main advantage is that the action space processed by the SL model is vastly reduced. Moreover, most constraints can be handled efficiently by the MILP problem **P2**. After decomposing the problem as a sequential decision-making problem, the main aim is to build an SL model to make optimal decisions at each time interval to minimize the operational cost of the isolated microgrid. Thus, the first task is to generate an expert dataset consisting of state-action pairs, where the state vector represents the input information of the specific scenarios of the problem, and the action vector represents the corresponding optimal solutions obtained by solving the optimization problem explicitly. These state-action pairs serve as the training data for the model training. Accordingly, the proposed strategy proposes a dense residual neural network (ResNetD) model that can capture the mapping between state variables and corresponding optimal actions. After the model is sufficiently trained, it can be used to make predictions or generate solutions that approximate the optimal solutions, which can be applied in the decision-making process for new scenarios. Since this study considers a time interval of 1 h ($\Delta\tau = 1$), the proposed real-time scheduling can be considered as an hour-ahead power scheduling strategy. The general process of the developed SL-based real-time power scheduling framework for an isolated microgrid is presented in Algorithm 1 and Fig. 4. The following subsections provide a detailed description of the proposed methodology.

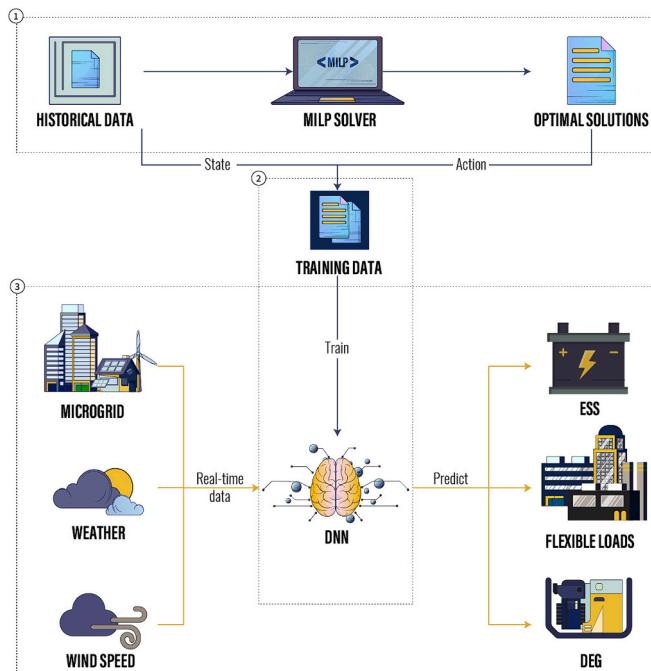


Fig. 4. Framework of the proposed methodology.

3.1. Stage 1: generate expert dataset by solving MILP problem on historical data

To generate an expert dataset, problem **P1** is formulated as a deterministic MILP optimization model and is solved exactly with precisely known input information. Numerous scenarios with accurate input information can be gathered in practical applications by collecting historical data or utilizing scenario generation. Each scenario corresponds to a power scheduling cycle in which the historical data of one scheduling cycle with T time intervals are considered. The data include the load demand of inflexible customer ($p_{1\tau}^L, p_{2\tau}^L, \dots, p_T^L$), expected load demand of flexible customer 1 ($\bar{p}_{1\tau}^{FL1}, \bar{p}_{2\tau}^{FL1}, \dots, \bar{p}_T^{FL1}$), expected load demand of flexible customer 2 ($\bar{p}_{1\tau}^{FL2}, \bar{p}_{2\tau}^{FL2}, \dots, \bar{p}_T^{FL2}$), solar irradiation ($v_{\Delta\tau}, v_{2\Delta\tau}, \dots, v_T$), ambient temperature ($\theta_{\Delta\tau}, \theta_{2\Delta\tau}, \dots, \theta_T$), and wind speed ($\gamma_{\Delta\tau}, \gamma_{2\Delta\tau}, \dots, \gamma_T$). These data are utilized as inputs for problem **P1**, and a dedicated MILP solver is employed to solve the problem as a day-ahead scheduling problem. The optimal values of the decision variables in (24) are obtained using the above process. Note that the optimal solution obtained by the dedicated MILP solver represents the ideal solution (i.e., the best possible solution), as problem **P1** is formulated as an MILP problem, and all the information during a scheduling cycle is known in advance, which is impossible in real operation.

From input data and optimal solutions for each scenario obtained by solving problem **P1**, an expert dataset of T state-action pairs $D = \{s_t, a_t\}_{t=1}^T$ can be formed, where the state vector s_t represents the observed information at each time interval t , and the action vector a_t represents optimal action corresponding to the observed information s_t . The state-action pair at each time interval t can be defined as follows:

$$s_t = [t, p_t^{net}, \epsilon_{t-1}^{ESS}]; \quad \forall t \in \Gamma \quad (29)$$

$$a_t = [p_t^{ESS}]; \quad \forall t \in \Gamma \quad (30)$$

where

$$p_t^{net} = p_t^L + \bar{p}_t^{FL1} + \bar{p}_t^{FL2} - p_t^{PV} - p_t^{WG}; \quad \forall t \in \Gamma \quad (31)$$

$$p_t^{ESS} = p_t^{ESS,ch} - p_t^{ESS,dch}; \quad \forall t \in \Gamma \quad (32)$$

The state vector s_t in (29) is rewritten from (25) by aggregating the load demand and the renewable generation at interval t into one single variable (i.e., net load), as presented in (31). As mentioned above, the action variable represents the charging/discharging power of the ESS, which is combined in a single variable using (32). This iterative process is implemented for all considered scenarios, and the resulting datasets of state-action pairs are combined, as indicated in lines 1–6 of Algorithm 1. Solving problem **P1** with N scenarios (each with T intervals) produces a dataset of NT state-action pairs, $D = \{s_i, a_i\}_{i=1}^{NT}$.

3.2. Stage 2: dense residual neural network (ResNetD) as a learner model

In the second stage, the objective is to learn a direct mapping π^* from a state in (29) to the corresponding optimal action in (30). Accordingly, a dataset of state-action pairs is utilized to train the model through SL. The inputs represent the state variables of the problem, whereas the outputs correspond to their optimal actions (i.e., optimal solutions). By learning from these labeled data, the model can generalize and make predictions for unseen inputs, resulting in optimized solutions in a data-driven manner. Using a dataset of NT state-action pairs obtained from solving historical scenarios with a MILP solver in Stage 1, an SL model π encodes a mapping $a = \pi(s; \theta)$ and learns the weight parameters θ through training on a dataset D . This process drives π towards the optimal policy π^* [42,53]. Technically, the problem is formulated as a regression problem with continuous inputs and outputs.

Residual neural network (ResNet), proposed first by He et al. [43], has become popular in deep learning. The ResNet architecture can make

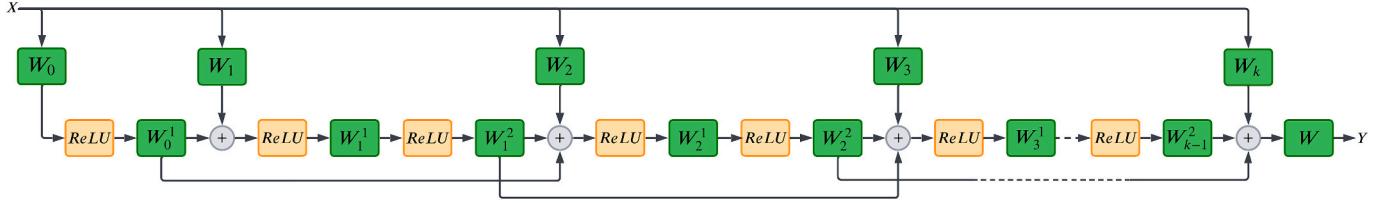


Fig. 5. The proposed ResNetD architecture with $K = 2$, where K is the number of hidden units W , X represents the input vector, and Y represents the output vector.

Table 2
Line parameters of the isolated microgrid.

Line	From bus	To bus	r (p.u.)	x (p.u.)
1	1	2	0.004100317	0.0064
2	1	3	0.007239685	0.007336076
3	1	4	0.007547918	0.022173236
4	4	5	0.003773959	0.011086618
5	4	6	0.004322245	0.004433667

training deeper networks feasible and efficient. It introduces the skip connection concept that allows the network to skip one or more layers. In traditional deep neural networks (DNNs), adding more layers can make training the network challenging due to problems of vanishing gradients and saturation training errors. However, in ResNets, the skip connections bypass a few layers and perform identity mapping, where the output of a layer is added to the output of a layer a few steps ahead. In this way, gradient flow can be preserved throughout the network, allowing for deeper architectures without degradation in performance.

This study proposes a dense residual neural network (ResNetD) as a learner model inspired by ResNet structures developed in previous studies [56–60]. In the proposed ResNetD architecture shown in Fig. 5, each block consists of two dense layers formed by merging the regular information flow, the output of previous blocks' dense layers, and a direct connection from the input through a dense layer [56]. The output of the last block is passed to the final prediction layer without an activation, which is typical for regression problems. The proposed ResNetD takes advantage of residual connections to combine both processed features and original input features at different stages, enhancing the gradient flow during training and potentially improving learning complex patterns in data. The ResNetD uses the rectified linear unit (ReLU) activation function and is trained using the Adam optimizer to minimize the mean square error (MSE) loss function. This workflow is outlined in lines 7–8 of Algorithm 1. The generalization ability of the ResNetD is expected to provide a real-time power scheduling tool capable of making near-optimal decisions in future scenarios.

3.3. Stage 3: real-time power scheduling

The real-time power scheduling strategy for an isolated microgrid is

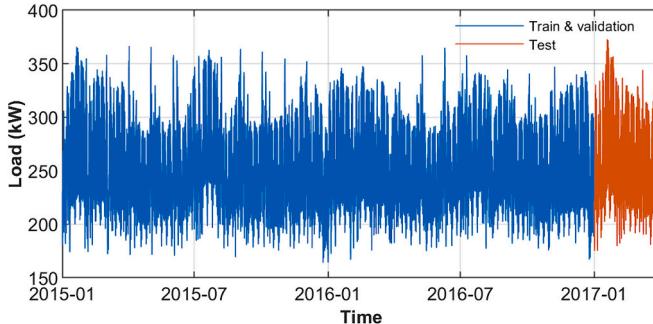


Fig. 6. Historical data of expected load demand of inflexible customer.

outlined in lines 9–17 of Algorithm 1 and is depicted graphically in Fig. 4 (Stage 3). Real-time power scheduling includes two main phases at each time interval. The first task is to find the optimal ESS action under the optimal mapping π^* , wherein the well-trained ResNetD is used for real-time decisions. The proposed strategy collects real-time data and determines state vector s_t in (29) at each interval t . Subsequently, the trained ResNetD π^* uses the state vector s_t as the input to predict the expected action $\hat{a}_t = \pi^*(s_t)$. The corresponding action of the ESS is defined by decomposing the expected action \hat{a}_t in (30). If the prediction is significantly biased or violates the rated power limits or over-charge/over-discharge of the ESS, the raw output estimated by the trained ResNetD is post-processed at each interval of the power scheduling cycle to fulfill their relevant bound constraints as follows:

$$\hat{p}_t^{ESS} = \begin{cases} (\bar{\epsilon}^{ESS} - \underline{\epsilon}_{t-1}^{ESS}) / (\eta^{ESS,ch} \cdot \Delta t); & \text{if } \hat{\epsilon}_t^{ESS} > \bar{\epsilon}_t^{ESS} \\ -(\underline{\epsilon}_{t-1}^{ESS} - \underline{\epsilon}_t^{ESS}) \cdot \eta^{ESS,dch} / \Delta t; & \text{if } \hat{\epsilon}_t^{ESS} < \underline{\epsilon}_t^{ESS}; \quad \forall t \in \Gamma \\ \hat{p}_t^{ESS}; & \text{otherwise} \end{cases} \quad (33)$$

$$\hat{p}_t^{ESS} = \begin{cases} \bar{p}^{ESS,ch}; & \text{if } \hat{p}_t^{ESS} > \bar{p}^{ESS,ch} \\ -\bar{p}^{ESS,dch}; & \text{if } \hat{p}_t^{ESS} < -\bar{p}^{ESS,dch}; \quad \forall t \in \Gamma \\ \hat{p}_t^{ESS}; & \text{otherwise} \end{cases} \quad (34)$$

Finally, the other variables, such as DEG and load demands of flexible customers, can be effectively obtained by solving the one-step optimization problem P2. Once the decision variables are fully determined, the system operator sends these values as a setpoint signal to operate the respective components of the isolated microgrid. This process is repeated at each time interval during the power scheduling cycle. Consequently, the original optimization problem (problem P1) is converted into a sequential decision-making problem decomposed into multiple one-step optimizations (one for each time interval).

4. Simulation results

The developed methodology is validated for a benchmark isolated microgrid with real-world data. The MILP formulation of the isolated microgrid is simulated in Python and solved using the Gurobi optimizer on a Mac Studio with an Apple M1 Max chip with 32 GB RAM. The simulations are performed over a 24-h time horizon with a time interval

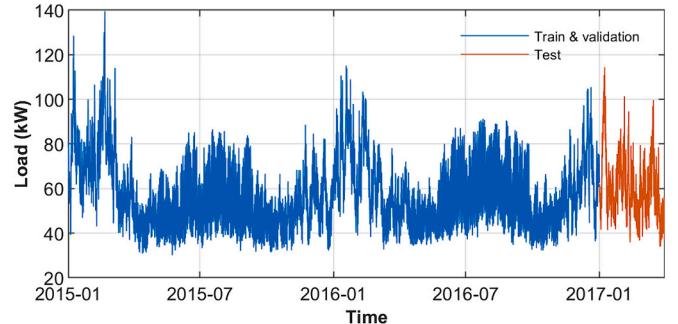


Fig. 7. Historical data of expected load demand of flexible customer 1.

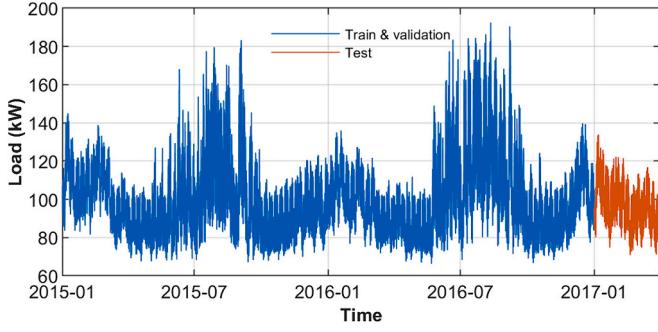


Fig. 8. Historical data of expected load demand of flexible customer 2.

of 60 min, resulting in 24 intervals for one power scheduling cycle.

4.1. Input data

The structure of the isolated microgrid under study is shown in Fig. 2. Table 2 gives relevant data for the power network, including the resistance and reactance parameters of the lines, as referred from the IEEE 13-bus Test Feeder [61]. For simplicity, it is assumed that the reactive power consumption accounts for a fixed proportion (namely 30 %) of the active power consumption, yielding power factors for each load of approximately 0.98. Moreover, the voltage magnitude limit is 0.95 p.u.–1.05 p.u.

Historical data on inflexible demand, flexible demands, solar irradiation, ambient temperature, and wind speed are considered to simulate the isolated microgrid model, as shown in Figs. 6–11. The inflexible demand is adapted from the actual load demand data in Spain. Additionally, simulations are conducted with two flexible consumers, each with different typical load profiles extracted and scaled from the power system data of the United States published by PJM Interconnection LLC [62]. The cost assigned to the unsupplied energy is set at 0.45 \$/kWh for flexible consumer 1 and 0.5 \$/kWh for flexible consumer 2. Under the DR programs, the expected demand for flexible loads could be reduced by up to 30 %. Solar irradiance and ambient temperature data are obtained from the European Commission in Madrid (Spain) [63]. Moreover, inflexible demand, flexible demands, solar radiation, and ambient temperature data are collected from January 2015 to March 2017, which amounts to 821 scenarios. To create a large wind speed dataset for 821 scenarios, Gaussian noise is injected into historical wind speed data from January 2020 to December 2020 provided by a wind energy company in Vietnam. Table 3 provides information regarding the DEG, PV, WG, and ESS which are sourced from references [15, 22, 44].

4.2. Ideal deterministic case

In the proposed methodology, Stage 1 involves formulating a deterministic MILP optimization model for an isolated microgrid, which

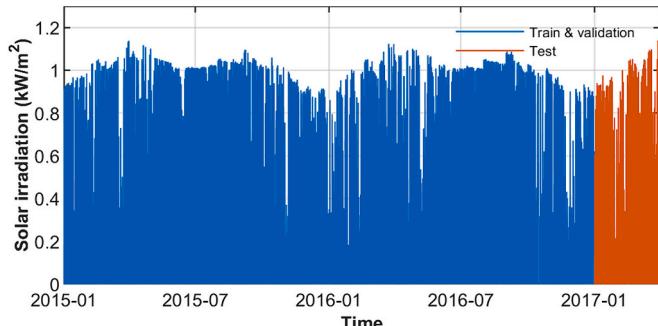


Fig. 9. Historical data of global solar irradiation.

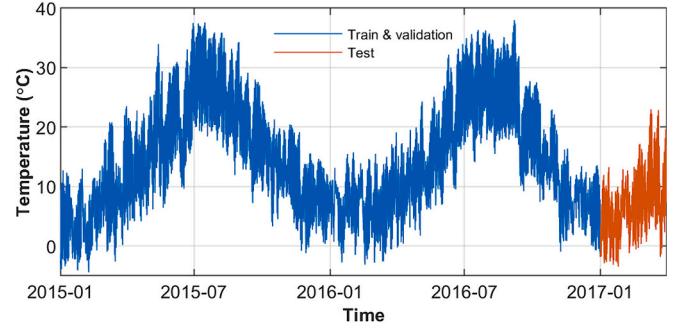


Fig. 10. Historical data of ambient temperature.

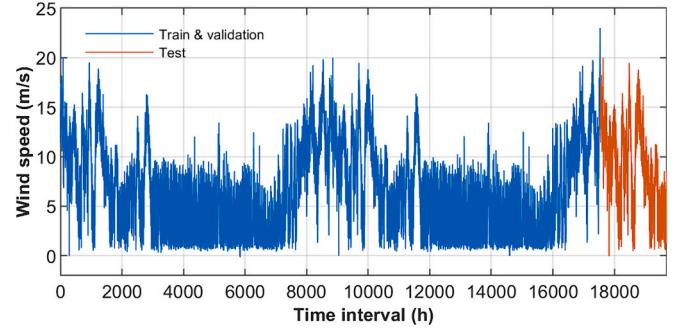


Fig. 11. Data of wind speed.

Table 3

Data of the DEG, PV, WG, and ESS.

Component	Parameter	Value
DEG	$\underline{P}^{DEG}; \bar{P}^{DEG}; R^{DEG}$ $\omega_1^{DEG}, \omega_2^{DEG}, \omega_3^{DEG}$	20; 600; 200 kW 1.3 \$/h; 0.0304 \$/kWh; 0.00104 \$/kWh ²
PV	\bar{P}^{PV} η^{PV} μ^{PV}	150 kW 0.167 p.u. 0.24 \$/kWh
WG	\bar{P}^{WG} η^{WG} $\underline{\gamma}^{WG}; \gamma^{WG,*}; \bar{\gamma}^{WG}$ $\alpha^{WG}; \beta^{WG}$ μ^{WG}	150 kW 0.88 p.u. 2; 11; 23 m/s 0.2268 kW·(m/s) ⁻³ ; 0.006 0.19 \$/kWh
ESS	$\bar{P}^{ESS,ch}; \bar{P}^{ESS,dch}$ $\underline{\epsilon}^{ESS}; \bar{\epsilon}^{ESS}$ $\eta^{ESS,ch}; \eta^{ESS,dch}$ μ^{ESS}	100 kW 40; 200 kWh 0.98 p.u. 10 ⁻⁶ \$/kWh ²

is known as the ideal theoretical method. This method assumes accurate information on all forecasted data, such as the expected load demands, solar irradiation, ambient temperature, and wind speed, which are gathered from historical data, as illustrated in Figs. 6–11. The MILP solver then uses this information to solve the MILP problem P1, thereby determining the optimal actions for flexible loads and the ESS to minimize the operational cost.

In Fig. 12, the scheduling results for a scenario randomly selected from the training scenario in Stage 1 are shown, clearly demonstrating the advantages of the optimal power scheduling developed for an isolated microgrid. The load demand is fully satisfied in the morning and at midday by utilizing the high renewable energy penetration and energy stored in the ESS. During the peak load period (16:00–23:00 h), the expected load demands from flexible consumers are only partly met by 73 %, followed by a decrease in the peak demand of 323.96 kW. The ESS is primarily discharged early in the morning when renewable generation is low (1:00–8:00 h), and charged during the hours of high renewable generation (10:00–12:00 h).

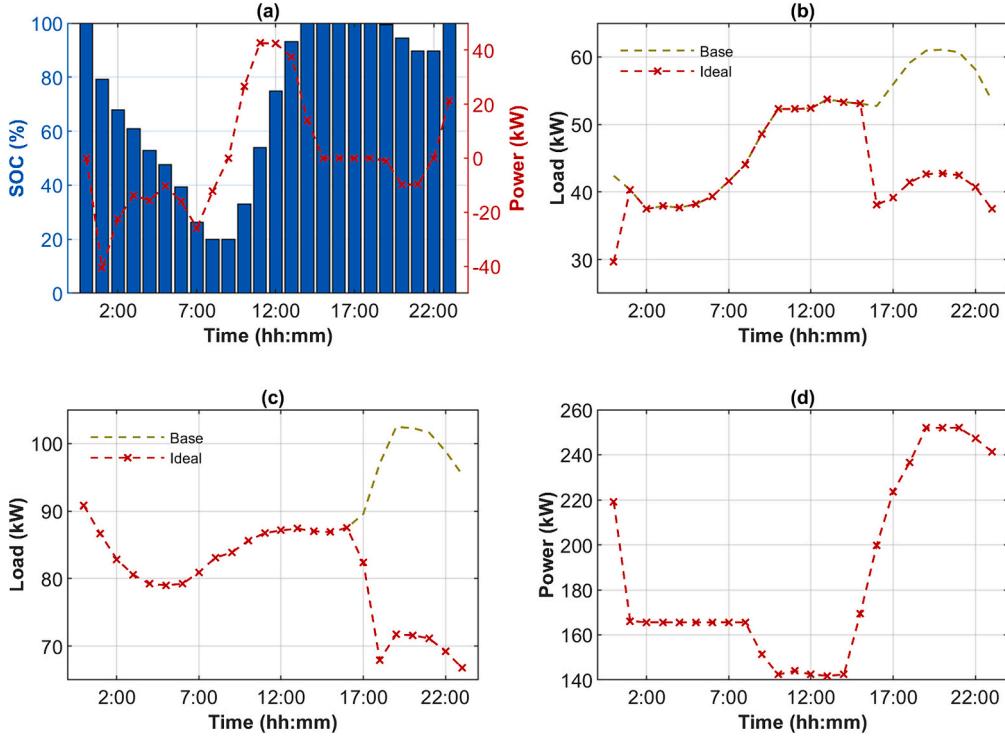


Fig. 12. Optimal results from the ideal theoretical method: (a) SOC and actions of ESS, (b) Flexible load 1, (c) Flexible load 2, (d) DEG.

In the base case, where the isolated microgrid fully provides the load demand using only DEG, PV, and WG (without using the ESS and DR programs), the operational cost is \$2107. Meanwhile, the ideal method achieves a significantly lower operational cost of \$2054 for this scenario. In the ideal method, the application of ESS and DR programs contributes to reducing operational costs by approximately \$53 while also decreasing the dependence on DEG. This highlights the crucial role of ESS in mitigating the impacts of RES intermittency and uncertainty in isolated microgrids. Additionally, all constraints on the flexible loads, DEG, and ESS during the power scheduling cycle are correctly ensured.

Based on the analysis of the results, it is evident that the ideal theoretical method is highly effective in achieving ideal power scheduling for an isolated microgrid. In essence, the MILP solver serves as an expert, and the optimal results obtained from the MILP solver provide perfect demonstrations of the proposed methodology. Notably, the optimal operational costs generated by the ideal method represent the lowest possible theoretical values, thus serving as a baseline reference for future comparisons.

4.3. Model training and convergence

Stage 2 involves training the ResNetD model to map states (29) to actions (30) using supervised training. Therefore, the ResNetD with four blocks containing two hidden layers is proposed, as described in Fig. 5. The hidden layers use the ReLU activation function, and the output layer employs a linear activation function. The ResNetD is trained using the Adam optimizer and the loss function of the MSE. The training process starts with an initial learning rate of 0.005 and implements learning rate decay and early stopping mechanisms to stabilize and accelerate the training process.

Based on historical weather, inflexible demand, and flexible demand data, 609 training scenarios are extracted, each corresponding to a daily power scheduling cycle of 24 intervals. Each scenario is solved with perfect information using the MILP solver. This resulted in the creation of 14,616 state-action pairs, which served as training samples for the ResNetD model. Additionally, the training progress of the ResNetD is

monitored using 122 holdout validation scenarios with 2928 state-action pairs. The loss curve for the training process of the ResNetD model is shown in Fig. 13, wherein the ResNetD training is completed in nearly 230 epochs, taking approximately 80 s. The insignificant difference between the training and validation losses demonstrates effective training of the ResNetD model. The generalization capability of a well-trained ResNetD is expected to enable it to perform near-optimal actions for flexible loads and ESS using real-time data.

4.4. Method performance

After offline training of the ResNetD model, it can be utilized to schedule an isolated microgrid in real time. The performance of the proposed SL method is tested across 90 scenarios. Figs. 6–11 show the data for the test scenarios. To apply real-time scheduling, the trained ResNetD model π^* is loaded. At each time interval t , the action a_t in Eq. (30) is determined by $\hat{a}_t = \pi^*(s_t)$.

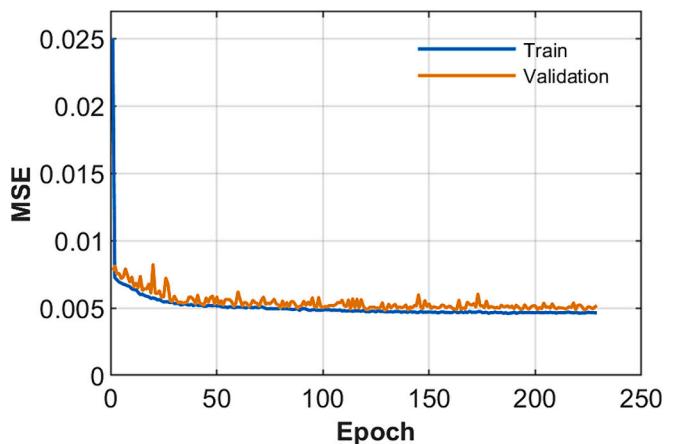


Fig. 13. Loss curve of the training process of the ResNetD model.

Table 4 presents the cumulative operational costs of all test scenarios obtained by the proposed SL method. The results of the ideal theoretical method and base case are also included in **Table 4**. In the base case, it is assumed that the isolated microgrid does not deploy ESS and DR programs, meaning that surplus energy from RESs cannot be stored, and load demands must be fully satisfied. Compared to the base case, the total operational cost is better when the SL method is applied, resulting in a 5.95 % improvement with the DR programs. Using the incentive-based DR program, the SL can satisfy 75.79 % of their load demands. Instead of using a DEG to meet the load demand fully, flexible consumers are compensated in a more attractive manner for curtailed energy [22]. Furthermore, the total DEG generation decreases from 663 MW (base case) to 579 MW, representing a reduction of 12.67 %. This reduction in DEG generation further contributes to reduced emissions. This demonstrates that ESS integration and flexible load scheduling provide economic and environmental benefits for an isolated microgrid. As shown in **Table 4**, the proposed SL achieves a total operational cost that is only 0.39 % higher than the ideal method. Moreover, its values of flexible demand satisfied and total DEG generation are also similar to the ideal results. **Fig. 14** depicts the comparison between the proposed SL method, base case, and ideal results in terms of operation costs of 90 test scenarios, which shows that the SL method (red dotted line) obtains better results than the base case (green line) and is very close to the ideal method (blue line) for most test scenarios.

For a more intuitive assessment, the power scheduling strategy of the proposed SL method is considered in a randomly selected scenario. **Fig. 15** depicts the input data of this scenario chosen for testing, where data on inflexible demand, two flexible demands, solar radiation, and ambient temperature are taken from March 9th, 2017, while the wind speed data was generated by injecting Gaussian noise to the wind speed data of March 9th, 2020.

The SL method is developed to operate without prior knowledge of future information, making it applicable to real-world applications. Meanwhile, the ideal method, which unrealistically assumes access to all future data, can be used as a benchmark to evaluate the efficacy of the SL method. **Fig. 16(a)** shows a comparison between the results obtained from the proposed SL method and the ideal result for the ESS operations. Furthermore, the SOCs of the ESS are determined and compared in **Fig. 16(b)**. The ESS actions in both methods show similar trends, although there are some time intervals where the differences are noticeable. It is worth noting that deviations in the decisions of the proposed method are acceptable owing to inevitable errors when training the ResNetD model. Subsequently, once the ESS action is determined at each interval, the flexible loads and DEG power can be defined by solving problem **P2**, as shown in **Fig. 17**. As shown in **Fig. 17** (a) and (b), the behaviors of the two flexible loads appear to be almost identical in most time intervals for both methods, except for the difference in the flexible load 1 at 11:00. It is essential to note that all the constraints imposed on the DEG, load demands, ESS, and AC power flow are fully satisfied.

For a quantitative comparison, the proposed SL method obtains an operational cost of \$2791 for the selected scenario, which is relatively close to the ideal result of \$2780. The trained ResNetD in the proposed method uses only the current state (i.e., only one interval) as the input to perform real-time power scheduling decisions. In contrast, the ideal method requires all future information to be available for all 24 intervals of the scheduling cycle; thus, the result of the ideal method is basically

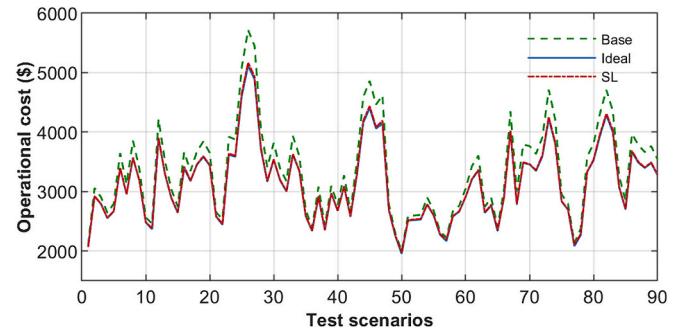


Fig. 14. Comparisons of the proposed SL, base case, and ideal method for operation costs of test scenarios.

the theoretical minimum that cannot be achieved in practice. It is evident from the comparisons that the proposed SL method is highly effective when applied to the isolated microgrid under study.

Fig. 18 shows the scheduling profile of the isolated microgrid for the selected scenario using the proposed method. As shown in **Figs. 16** and **18**, the ESS is deeply discharged to supply the system when the load demand increases and the power generated by the RESs remains low during 1:00–8:00 h. During these periods, load shedding is also performed to reduce DEG generation, as shown in **Fig. 17**. However, during periods of high renewable generation from wind and solar energy (11:00–16:00 h), the ESS is fully charged and flexible loads are entirely supplied. During peak load periods (17:00–23:00 h), the proposed strategy forces load shedding to balance demand and generation. The SL method successfully learns and mimics the decision-making processes of the MILP solver. Consequently, the proposed SL method offers an intelligent real-time power scheduling strategy that enables system operators to manage an isolated microgrid effectively based on the DR strategy.

4.5. Performance comparison

The performance of the proposed SL approach is further verified by comparing it with other extensively used methods. The comparison is conducted using 90 test scenarios. The comparable methods are briefly described as follows:

- **Ideal method:** Based on the exact known information of the entire scheduling cycle, the problem is solved as a deterministic MILP optimization model, and the result is the best possible solution, which serves as the baseline.
- **Myopic method:** An online technique that considers only the current time step, which means that problem **P1** is solved by the MILP solver at each time step (i.e., horizon $T = 1$) with real-time information. Thus, this method does not explicitly consider the future impact of current decisions.
- **Deep deterministic policy gradient (DDPG)** [64] is an actor-critic, model-free algorithm that merges the benefits of deep learning and deterministic policy gradient (DPG), where the actor learns to select actions, and the critic learns to evaluate the quality of those actions. It utilizes DNNs to approximate both the policy and the value function. DDPG optimizes policies based on the gradient of expected returns, efficiently navigating complex environments through a combination of policy-based and value-based methods.
- **Soft actor-critic (SAC)** [65] is an advanced RL algorithm that harmonizes stochastic policy optimization with DDPG-like methods through off-policy training. Its key characteristic is entropy regularization, encouraging the policy to optimize a balance between expected return and entropy, which quantifies policy randomness. This principle is deeply tied to the exploration-exploitation dilemma:

Table 4
Simulation results of the proposed SL, ideal method, and base case for test scenarios.

Indicator	Base case	Ideal method	The proposed SL
Operation cost (\$)	302,140	283,082	284,172
Flexible demand satisfied (%)	100	75.85	75.79
Total DEG generation (MW)	663	580	579

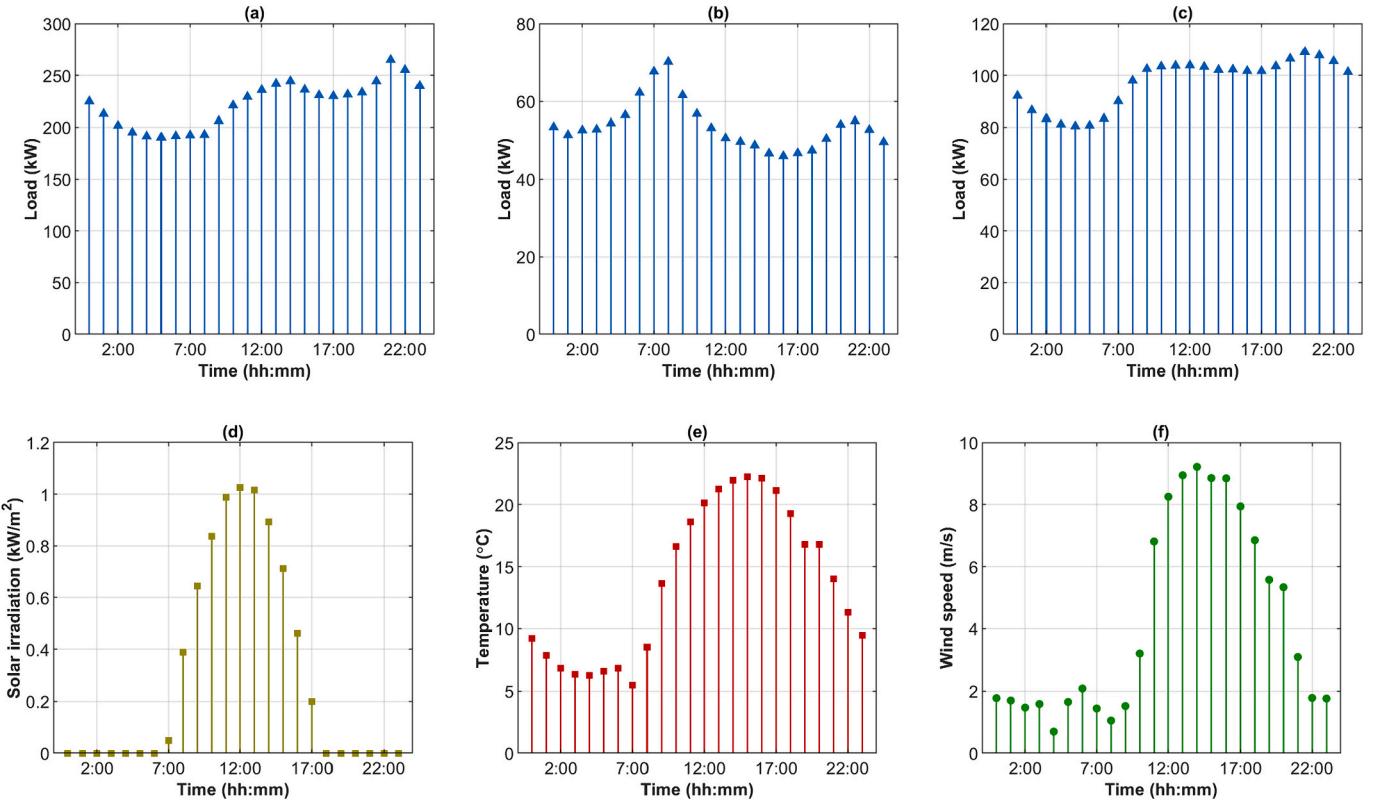


Fig. 15. Data for the selected test scenario: (a) Expected inflexible load demand, (b) Expected flexible load demand 1, (c) Expected flexible load demand 2, (d) Solar irradiation, (e) Ambient temperature, (f) Wind speed.

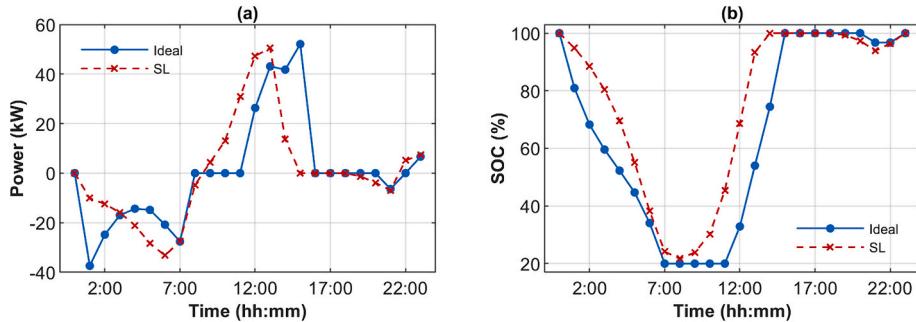


Fig. 16. Comparisons of the proposed SL and ideal result: (a) ESS charging/discharging power, (b) SOCs of ESS.

higher entropy fosters more exploration, enhancing future learning and averting premature convergence to suboptimal local optima. - Proximal policy optimization (PPO) [66] is a state-of-the-art DRL technique for optimizing stochastic control policies. It stands out for its simplicity and effectiveness, particularly in environments with high-dimensional observation and action spaces. PPO enhances learning stability and efficiency by using a clipped objective function, preventing drastic policy updates and balancing exploration and exploitation.

For a fair comparison, DDPG, PPO, and SAC are also trained to predict the best action of ESS using the same state and action as in (29) and (30). Thus, the steps to perform real-time power scheduling of these DRL methods are similar to the proposed SL as given in Section 3.3. The main difference between the proposed SL and DRL methods is the training process, where DDPG, PPO, and SAC learn based on trial and error, while the proposed SL learns from perfect demonstrations. All methods are tested with the same test scenarios in the same environment

of the isolated microgrid. The input data of the isolated microgrid components is given in Table 3. Data for 90 testing scenarios are depicted in orange color in Figs. 6–11.

In this study, a performance error for the n^{th} scenario is used to evaluate the performance of a method in real-time power scheduling, which can be given as follows:

$$J_{\text{gap}} = \frac{J_n - J_n^{\text{ideal}}}{J_n^{\text{ideal}}} \quad (35)$$

where J_n and J_n^{ideal} are the operational costs obtained by the proposed and ideal methods for the n^{th} scenario, respectively. Performance errors are also employed to evaluate the cumulative cost of all test scenarios.

Table 5 presents the results, including average and standard deviation values of performance errors obtained by the proposed method and myopic policy, PPO, SAC, and DDPG over 90 test scenarios. As shown in Table 5, the proposed SL method exhibits an average performance error of 0.37 %, which is significantly superior to the myopic policy (2.01 %),

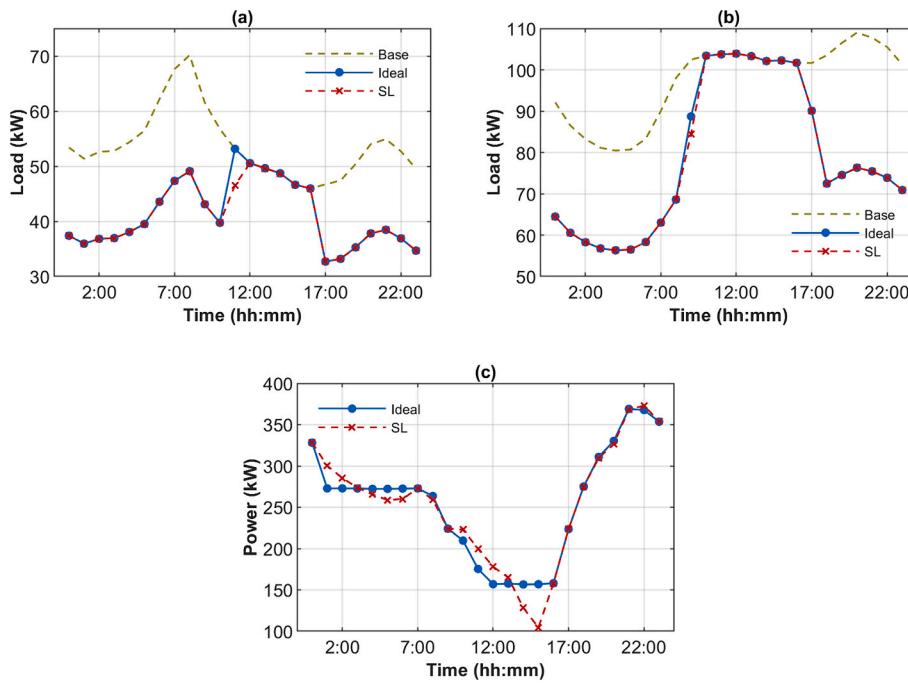


Fig. 17. Comparisons of the proposed SL and ideal result: (a) Flexible load 1, (b) Flexible load 2, (c) DEG.

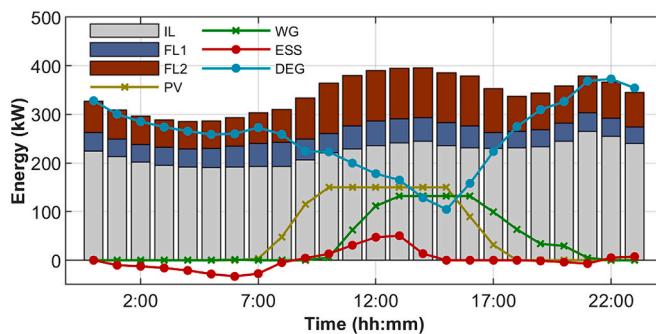


Fig. 18. Scheduling results of isolated microgrid using the proposed SL method.

Table 5
Comparisons of performance errors of different methods for test scenarios.

Method	Performance error for each scenario (%)		Performance error for cumulative operational cost (%)
	Average	Std.	
Myopic policy	2.01	0.51	2.00
PPO	1.65	0.55	1.64
SAC	1.50	0.47	1.53
DDPG	1.03	0.45	1.06
The proposed SL	0.37	0.27	0.39

PPO (1.65 %), SAC (1.50 %), and DDPG (1.03 %). It is also worth noting that the standard deviation value of the performance error of the proposed method is reasonably small and lower than those of other methods, which shows its effective generalization ability and robustness in performing real-time power scheduling for different test scenarios.

Table 5 also presents the performance errors for cumulative operational costs of all test scenarios obtained by different methods. The proposed SL, myopic policy, PPO, SAC, and DDPG obtain performance errors for cumulative costs of 0.39 %, 2.00 %, 1.64 %, 1.53 %, and 1.06

%, respectively. Accordingly, the SL method also outperforms the other methods in terms of cumulative operational cost. With an average performance error of 0.37 % and a performance error for cumulative costs of 0.39 %, the SL method achieves operational costs close to the optimal values obtained by the ideal method under perfect information. The results in Table 5 are further verified by comparing boxplots of performance errors of these methods in Fig. 19. The proposed method has the lowest mean performance error, indicated by the central line in the box. The interquartile range, expressed as the height of the box, is also relatively small for the SL method, showing less variability in performance error. Thus, SL has the best performance in terms of median performance error, while DDPG is the second best.

Based on these comparisons, it can be observed that the myopic policy underperforms the SL and other DRL methods. This is likely because the myopic policy focuses solely on optimizing operational costs at each time interval without explicitly considering the impact of current decisions on future cumulative operational costs. In contrast, the SL approach aims to mimic a reliable expert by learning from perfect demonstrations rather than relying solely on trial and error, as in the

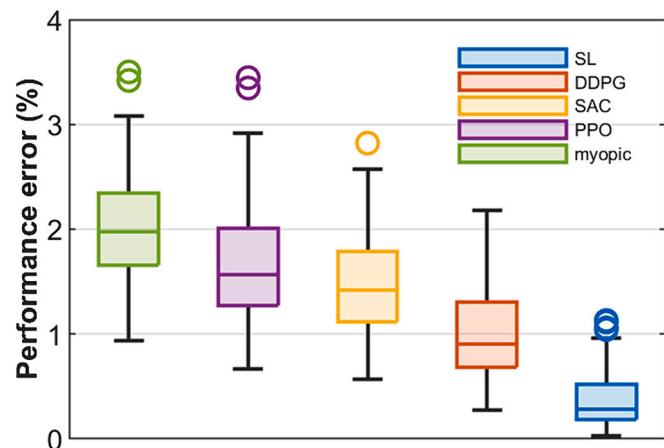


Fig. 19. Boxplots of performance errors of different methods.

case of the DRL methods. The SL method also leverages the generalization capability of the ResNetD to provide optimal actions for the regression problem based on the optimal decisions derived from the MILP solver. The proposed SL method outperforms the other methods across various test scenarios, indicating good generalizability, stability, and robustness in real-time power scheduling for the isolated microgrid under study.

Table 6 presents a comparison of training time and real-time execution time of the proposed method, myopic policy, PPO, SAC, and DDPG, in which the myopic policy does not implement offline training. Since the proposed SL directly trains the ResNetD model on the expert dataset of state-action pairs, the training time is significantly shorter than PPO, SAC, and DDPG while still obtaining better performance than these DRL methods. It should be emphasized that the training time of the proposed algorithm is mainly due to the MILP optimization to create the expert dataset (i.e., Stage 1 in Section 3). The ResNetD model training only takes about 80 s, as discussed in Section 4.3. Moreover, the average execution time of the proposed method for one single time-step scheduling is 0.1162 s, which is approximately the same as those of other methods. Thus, the proposed method can satisfy the timing requirements of real-time executions.

4.6. Stability and feasibility of the proposed method

To evaluate the performance stability of the proposed method, five training runs of the ResNetD model on the expert dataset are performed independently to obtain five different mappings. Afterward, each mapping is applied to perform real-time power scheduling on 90 test scenarios. Fig. 20 displays the boxplots of the performance errors of the SL over five different runs, which validate its stability and repeatability. The average performance errors of all five runs are relatively consistent (approximately 0.4 %). Therefore, the proposed SL method has a very stable performance in different runs.

To verify the feasibility of the decision making strategy in terms of power flow constraints, the relaxation gaps of the power flow are defined and plotted in Fig. 21, which can be given as follows [67]:

$$RG_{ij|t} = \left| \frac{\left(P_{ij|t}^2 + Q_{ij|t}^2 \right)}{v_{ij|t}} - l_{ij|t} \right| \quad \forall j \in N \wedge (i,j) \in E \wedge t \in \Gamma \quad (36)$$

Small values of relaxation gaps indicate better AC power flow feasibility [67]. Fig. 21 shows that the maximum gap over 90 test scenarios is less than 7×10^{-5} , confirming that power flow constraints are valid for all test scenarios. Thus, the proposed method provides a near-optimal solution while also meeting the feasibility of the AC power flow constraints.

4.7. Performance of the proposed method with different sets of state-action pairs

In this section, some extra trials are conducted to investigate the performance of the proposed method when using different sets of state-action pairs during training ResNetD. Accordingly, other decision variables are chosen as action variables as follows:

- (1) Approach #1: the variables related to two flexible loads are selected as characteristic variables, and the ESS charging/discharging power is calculated by solving the MILP optimization problem as follows:

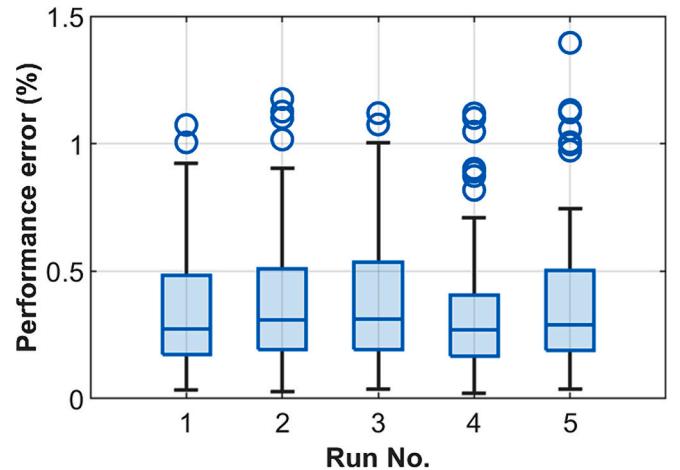


Fig. 20. Boxplots of performance errors of the proposed SL in five different runs.

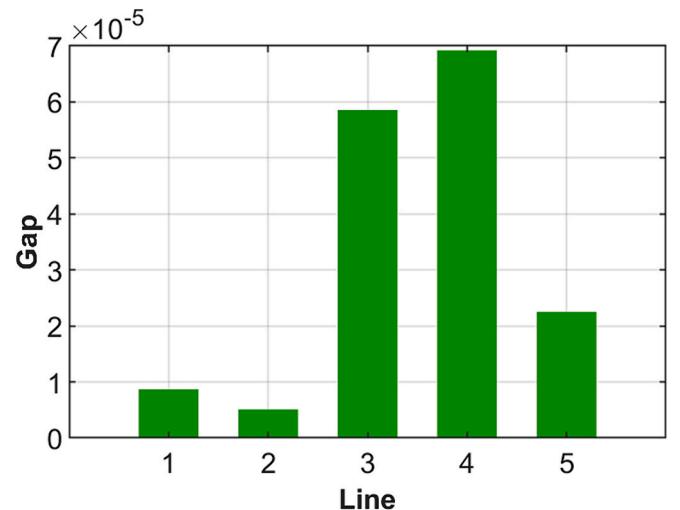


Fig. 21. The relaxation gaps of the power flow.

$$s_t = [t, p_t^{IL-RES}, \bar{p}_t^{FL1}, \bar{p}_t^{FL2}]; \quad \forall t \in \Gamma \quad (37)$$

$$a_t = [p_t^{FL1}, p_t^{FL2}]; \quad \forall t \in \Gamma \quad (38)$$

where

$$p_t^{IL-RES} = p_t^{IL} - p_t^{PV} - p_t^{WG}; \quad \forall t \in \Gamma \quad (39)$$

- (2) Approach #2: the variables related to two flexible loads and ESS charging/discharging power are all selected as characteristic variables in actions:

$$s_t = [t, p_t^{IL-RES}, \bar{p}_t^{FL1}, \bar{p}_t^{FL2}, \epsilon_{t-1}^{ESS}]; \quad \forall t \in \Gamma \quad (40)$$

$$a_t = [p_t^{FL1}, p_t^{FL2}, p_t^{ESS}]; \quad \forall t \in \Gamma \quad (41)$$

Table 6

Comparisons of training and execution times of different methods.

Method	The proposed SL	Myopic policy	PPO	SAC	DDPG
Training time (s)	2714	–	13,389	14,407	13,818
Execution time (s)	0.1162	0.1012	0.1241	0.1233	0.1262

Table 7

Comparisons of performance errors of the proposed SL with different sets of state-action pairs.

Approach	Performance error for each scenario (%)		Performance error for cumulative operational cost (%)
	Average	Std.	
Approach #1	2.19	0.50	2.16
Approach #2	0.59	0.40	0.63
The proposed approach	0.37	0.27	0.39

(3) The proposed approach: the state-action pairs are given as in (29) and (30) in the proposed study.

The ResNetD models are trained using different datasets of state-action pairs for the remaining two approaches (i.e., Approach #1 and Approach #2). The ResNetD models of all approaches are tested in the same environment, and their results are given in Table 7. From comparisons in Table 7, the proposed approach obtains the best results in terms of average performance errors for each scenario and performance errors for cumulative operational costs of all scenarios. Indeed, the average performance errors obtained by the proposed approach are much better than those of Approach #1 (2.19 %) and Approach #2 (0.59 %). Furthermore, Approach #1 has the worst performance, indicating that removing the ESS power from action variables significantly reduces the performance of the proposed model. It confirms that choosing the charging/discharging power of the ESS as the characteristic variable is the best approach compared to the two other approaches.

The proposed SL method is developed to be adaptable and scalable to new microgrid environments. The main advantage of the proposed methodology lies in its ability to learn from historical data and expert decisions to generate near-optimal power scheduling decisions in real-time, which can be particularly beneficial in managing the dynamic and uncertain nature of RESs and demand patterns in isolated microgrids. For new systems, it is vital to formulate the appropriate optimization problem and collect sufficient data to create an expert dataset of optimal decisions. For large and complex systems with more state and action variables, more complex ResNetD models with more hidden layers and neurons can be appropriately developed and fine-tuned to learn a function approximator to map the states to the corresponding optimal actions. In conclusion, the proposed SL method is flexible and can be applied to new isolated microgrids. Its ability to learn from data makes it a powerful tool for optimizing the operation of diverse microgrid configurations, enhancing their efficiency, reliability, and sustainability.

5. Conclusion

In this study, a real-time power scheduling strategy was developed for an isolated microgrid integrated with DEG, RESs, flexible loads, and ESS using an SL method. The optimal power scheduling problem was formulated and solved using a dedicated MILP solver that utilized historical data and considered system constraints to derive the optimal solutions. Subsequently, a ResNetD model was trained to learn and

mimic the optimal actions based on the real-time states, accelerating decision-making in real-time scenarios and minimizing the operational cost of the benchmark isolated microgrid. Based on real-world data simulations, the proposed SL method can provide near-optimal power scheduling decisions using real-time information. Compared with the base case, the SL approach decreased operational costs by 5.95 % and DEG generation by 12.67 %. The cumulative operational costs for all test scenarios obtained by the SL method were only 0.39 % higher than the ideal result, indicating that its performance was very close to that of the ideal solution. Furthermore, a comparative analysis demonstrates that the SL method outperforms the myopic policy, PPO, SAC, and DDPG, demonstrating its high performance and excellent generalizability in effectively handling unseen scenarios. Therefore, the proposed SL-based framework is an effective energy management tool for real-time optimal power scheduling of isolated microgrids.

Despite the effectiveness of the proposed SL, this study may also have certain limitations. Supervised training of the ResNetD model requires significant historical data, which may be difficult to collect for other large-scale systems. Furthermore, this study only focuses on the economic aspect of an isolated microgrid and assumes to ignore some technical aspects, such as the effects of battery degradation, and failures of grid components, to simplify the model. Thus, future studies should apply the SL method to complex and large-scale systems, such as multi-energy systems, multi-microgrid systems, and distribution networks. Moreover, the generalized performance of the proposed SL needs to be enhanced to improve its adaptability to unexpected situations, such as failures of DEG, RESs, ESS.

CRediT authorship contribution statement

Truong Hoang Bao Huy: Writing – original draft, Validation, Methodology, Conceptualization. **Tien-Dat Le:** Visualization, Validation, Investigation, Data curation. **Pham Van Phu:** Visualization, Methodology, Investigation, Formal analysis. **Seongkeun Park:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Daehee Kim:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2023-00245084), and supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023-00277255). This work was also supported by the Soonchunhyang Research Fund.

Appendix A. Linearization of quadratic terms

To linearize the quadratic and cubic terms, a piecewise linear approximation is applied, wherein the nonlinear function ψ is discretized in n grid points that determine $(n - 1)$ segments as follows [22,68]:

$$\psi = \langle \tilde{x}_i, \psi(\tilde{x}_i) \rangle; \quad \forall i \in \{2, 3, \dots, n\} \quad (A1)$$

Thus, the nonlinear term x can be replaced by the auxiliary variable z in the following equation [45]:

$$z = \sum_{i=2}^n \{\delta_i \cdot (K_i \cdot x + L_i)\} \quad (A2)$$

where δ is a binary SOS1, and K and L can be determined as follows [24]:

$$K_i = \frac{\psi(\tilde{x}_i) - \psi(\tilde{x}_{i-1})}{\tilde{x}_i - \tilde{x}_{i-1}}; \quad \forall i \in \{2, 3, \dots, n\} \quad (A3)$$

$$L_i = \psi(\tilde{x}_i) - K_i \cdot \tilde{x}_i; \quad \forall i \in \{2, 3, \dots, n\} \quad (A4)$$

To simultaneously prevent the activation of more than one piecewise section, constraint (A5) must be imposed as follows [45]:

$$\sum_{i=1}^{n-1} \{\delta_i \cdot \tilde{x}_i\} \leq x \leq \sum_{i=2}^n \{\delta_{i-1} \cdot \tilde{x}_i\} \quad (A5)$$

The products of the integer and continuous variables that appear in (A2) can be linearized by imposing the constraints (A6) and (A7) [22,24].

$$x - M \cdot (1 - \delta_i) \leq \omega_i \leq x + M \cdot (1 - \delta_i); \quad \forall i \in \{2, 3, \dots, n\} \quad (A6)$$

$$-M \cdot \delta_i \leq \omega_i \leq M \cdot \delta_i; \quad \forall i \in \{2, 3, \dots, n\} \quad (A7)$$

References

- [1] C. Gerbaulet, C. von Hirschhausen, C. Kemfert, C. Lorenz, P.-Y. Oei, European electricity sector decarbonization under different levels of foresight, *Renew. Energy* 141 (2019) 973–987, <https://doi.org/10.1016/j.renene.2019.02.099>.
- [2] T.H.B. Huy, H.T. Dinh, D. Kim, Multi-objective framework for a home energy management system with the integration of solar energy and an electric vehicle using an augmented ϵ -constraint method and lexicographic optimization, *Sustain. Cities Soc.* 88 (2023) 104289, <https://doi.org/10.1016/j.scs.2022.104289>.
- [3] Z. Qu, C. Xu, F. Yang, F. Ling, S. Pirouzi, Market clearing price-based energy management of grid-connected renewable energy hubs including flexible sources according to thermal, hydrogen, and compressed air storage systems, *J. Energy Storage* 69 (2023) 107981, <https://doi.org/10.1016/j.est.2023.107981>.
- [4] M. Norouzi, J. Aghaei, S. Pirouzi, T. Niknam, M. Fotuhi-Firuzabad, Flexibility pricing of integrated unit of electric spring and EVs parking in microgrids, *Energy* 239 (2022) 122080, <https://doi.org/10.1016/j.energy.2021.122080>.
- [5] S. Vergine, C. Álvarez-Arroyo, G. D'Amico, J.M. Escano, L. Alvarado-Barrios, Optimal management of a hybrid and isolated microgrid in a random setting, *Energy Rep.* 8 (2022) 9402–9419, <https://doi.org/10.1016/j.egyr.2022.07.044>.
- [6] S. Kakran, S. Chanana, Smart operations of smart grids integrated with distributed generation: a review, *Renew. Sustain. Energy Rev.* 81 (2018) 524–535, <https://doi.org/10.1016/j.rser.2017.07.045>.
- [7] H. Jiayi, J. Chuanwen, X. Rong, A review on distributed energy resources and MicroGrid, *Renew. Sustain. Energy Rev.* 12 (2008) 2472–2483, <https://doi.org/10.1016/j.rser.2007.06.004>.
- [8] X. Zhang, X. Yu, X. Ye, S. Pirouzi, Economic energy management of networked flexi-renewable energy hubs according to uncertainty modeling by the unscented transformation method, *Energy* 278 (2023) 128054, <https://doi.org/10.1016/j.energy.2023.128054>.
- [9] E. Akbari, S.F. Mousavi Shabestari, S. Pirouzi, M. Jadidoleslam, Network flexibility regulation by renewable energy hubs using flexibility pricing-based energy management, *Renew. Energy* 206 (2023) 295–308, <https://doi.org/10.1016/j.renene.2023.02.050>.
- [10] G. Pilant, S. Pirouzi, A. Azarhooshang, A. Rezaee Jordehi, A. Paeizi, M. Ghadamyari, Storage-integrated virtual power plants for resiliency enhancement of smart distribution systems, *J. Energy Storage* 55 (2022) 105563, <https://doi.org/10.1016/j.est.2022.105563>.
- [11] S. Pirouzi, Network-constrained unit commitment-based virtual power plant model in the day-ahead market according to energy management strategy, *IET Gener. Transm. Distrib.* 17 (2023) 4958–4974, <https://doi.org/10.1049/gtd2.13008>.
- [12] M. Norouzi, J. Aghaei, T. Niknam, S. Pirouzi, M. Lehtonen, Bi-level fuzzy stochastic-robust model for flexibility valorizing of renewable networked microgrids, *Sustain. Energy Grids Networks* 31 (2022) 100684, <https://doi.org/10.1016/j.segan.2022.100684>.
- [13] A. Ghasemi, M. Enayatzare, Optimal energy management of a renewable-based isolated microgrid with pumped-storage unit and demand response, *Renew. Energy* 123 (2018) 460–474, <https://doi.org/10.1016/j.renene.2018.02.072>.
- [14] H. Moradi, M. Esfahanian, A. Abtahi, A. Zilouchian, Optimization and energy management of a standalone hybrid microgrid in the presence of battery storage system, *Energy* 147 (2018) 226–238, <https://doi.org/10.1016/j.energy.2018.01.016>.
- [15] L. Alvarado-Barrios, Á. Rodríguez del Nozal, J. Boza Valerino, I. García Vera, J. L. Martínez-Ramos, Stochastic unit commitment in microgrids: influence of the load forecasting error and the availability of energy storage, *Renew. Energy* 146 (2020) 2060–2069, <https://doi.org/10.1016/j.renene.2019.08.032>.
- [16] S.E. Ahmadi, N. Rezaei, A new isolated renewable based multi microgrid optimal energy management system considering uncertainty and demand response, *Int. J. Electr. Power Energy Syst.* 118 (2020) 105760, <https://doi.org/10.1016/j.ijepes.2019.105760>.
- [17] F. Moazeni, J. Khazaei, Optimal operation of water-energy microgrids: a mixed integer linear programming formulation, *J. Clean. Prod.* 275 (2020) 122776, <https://doi.org/10.1016/j.jclepro.2020.122776>.
- [18] R. Nourollahi, P. Salyani, K. Zare, B. Mohammadi-Ivatloo, Resiliency-oriented optimal scheduling of microgrids in the presence of demand response programs using a hybrid stochastic-robust optimization approach, *Int. J. Electr. Power Energy Syst.* 128 (2021) 106723, <https://doi.org/10.1016/j.ijepes.2020.106723>.
- [19] A.R. Jordehi, Scheduling heat and power microgrids with storage systems, photovoltaic, wind, geothermal power units and solar heaters, *J. Energy Storage* 41 (2021) 102996, <https://doi.org/10.1016/j.est.2021.102996>.
- [20] A. Rezaee Jordehi, Information gap decision theory for operation of combined cooling, heat and power microgrids with battery charging stations, *Sustain. Cities Soc.* 74 (2021) 103164, <https://doi.org/10.1016/j.scs.2021.103164>.
- [21] M.M. Rana, M.F. Romlie, M.F. Abdullah, M. Uddin, M.R. Sarkar, A novel peak load shaving algorithm for isolated microgrid using hybrid PV-BESS system, *Energy* 234 (2021) 121157, <https://doi.org/10.1016/j.energy.2021.121157>.
- [22] M. Tostado-Vélez, S. Kamel, F. Aymen, A. Rezaee Jordehi, F. Jurado, A stochastic-IGDT model for energy management in isolated microgrids considering failures and demand response, *Appl. Energy* 317 (2022) 119162, <https://doi.org/10.1016/j.apenergy.2022.119162>.
- [23] S. Ferahita, H. Rezik, M.A. Abdelkareem, A.G. Olabi, Optimal techno-economic energy management strategy for building's microgrids based bald eagle search optimization algorithm, *Appl. Energy* 306 (2022) 118069, <https://doi.org/10.1016/j.apenergy.2021.118069>.
- [24] M. Tostado-Vélez, S. Kamel, H.M. Hasanien, R.A. Turky, F. Jurado, A mixed-integer-linear-logical programming interval-based model for optimal scheduling of isolated microgrids with green hydrogen-based storage considering demand response, *J. Energy Storage* 48 (2022) 104028, <https://doi.org/10.1016/j.est.2022.104028>.
- [25] J.M. Manzano, J.R. Salvador, J.B. Romaine, L. Alvarado-Barrios, Economic predictive control for isolated microgrids based on real world demand/renewable energy data and forecast errors, *Renew. Energy* 194 (2022) 647–658, <https://doi.org/10.1016/j.renene.2022.05.103>.
- [26] H. Zhang, G. Li, S. Wang, Optimization dispatching of isolated island microgrid based on improved particle swarm optimization algorithm, *Energy Rep.* 8 (2022) 420–428, <https://doi.org/10.1016/j.egyr.2022.10.199>.
- [27] R. Asri, H. Aki, D. Kodaira, Optimal operation of shared energy storage on islanded microgrid for remote communities, *Sustain. Energy Grids Networks* 35 (2023) 101104, <https://doi.org/10.1016/j.segan.2023.101104>.
- [28] M.R. Jokar, S. Shahmoradi, A.H. Mohammed, L.K. Foong, B.N. Le, S. Pirouzi, Stationary and mobile storages-based renewable off-grid system planning considering storage degradation cost based on information-gap decision theory optimization, *J. Energy Storage* 58 (2023) 106389, <https://doi.org/10.1016/j.est.2022.106389>.
- [29] T.H.B. Huy, H. Truong Dinh, D. Ngoc Vo, D. Kim, Real-time energy scheduling for home energy management systems with an energy storage system and electric vehicle based on a supervised-learning-based strategy, *Energ. Convers. Manage.* 292 (2023) 117340, <https://doi.org/10.1016/j.enconman.2023.117340>.
- [30] P. Tadepalli, D. Ok, Model-based average reward reinforcement learning, *Artif. Intell.* 100 (1998) 177–224, [https://doi.org/10.1016/S0004-3702\(98\)00002-2](https://doi.org/10.1016/S0004-3702(98)00002-2).
- [31] R.S. Sutton, A.G. Barto, *Reinforcement Learning: An Introduction*, A Bradford Book, Cambridge, MA, USA, 2018.

- [32] B.E. Nyong-Bassey, D. Giaouris, C. Patsios, S. Papadopoulos, A.I. Papadopoulos, S. Walker, S. Voutetakis, P. Seferlis, S. Gadoue, Reinforcement learning based adaptive power pinch analysis for energy management of stand-alone hybrid energy storage systems considering uncertainty, *Energy* 193 (2020) 116622, <https://doi.org/10.1016/j.energy.2019.116622>.
- [33] Y. Li, R. Wang, Z. Yang, Optimal scheduling of isolated microgrids using automated reinforcement learning-based multi-period forecasting, *IEEE Trans. Sustain. Energy* 13 (2022) 159–169, <https://doi.org/10.1109/TSTE.2021.3105529>.
- [34] W. Dong, H. Sun, C. Mei, Z. Li, J. Zhang, H. Yang, Forecast-driven stochastic optimization scheduling of an energy management system for an isolated hydrogen microgrid, *Energ. Convers. Manage.* 277 (2023) 116640, <https://doi.org/10.1016/j.enconman.2022.116640>.
- [35] Y. Khawaja, I. Qiqieh, J. Alzubi, O. Alzubi, A. Allahham, D. Giaouris, Design of cost-based sizing and energy management framework for standalone microgrid using reinforcement learning, *Sol. Energy* 251 (2023) 249–260, <https://doi.org/10.1016/j.solener.2023.01.027>.
- [36] P. Kofinas, A.I. Dounis, G.A. Vouros, Fuzzy Q-learning for multi-agent decentralized energy management in microgrids, *Appl. Energy* 219 (2018) 53–67, <https://doi.org/10.1016/j.apenergy.2018.03.017>.
- [37] S. Totaro, I. Boukas, A. Jonsson, B. Cornélusse, Lifelong control of off-grid microgrid with model-based reinforcement learning, *Energy* 232 (2021) 121035, <https://doi.org/10.1016/j.energy.2021.121035>.
- [38] L. Lei, Y. Tan, G. Dahlenburg, W. Xiang, K. Zheng, Dynamic energy dispatch based on deep reinforcement learning in IoT-driven smart isolated microgrids, *IEEE Internet Things J.* 8 (2021) 7938–7953, <https://doi.org/10.1109/JIOT.2020.3042007>.
- [39] Y. Xia, Y. Xu, Y. Wang, S. Mondal, S. Dasgupta, A.K. Gupta, G.M. Gupta, A safe policy learning-based method for decentralized and economic frequency control in isolated networked-microgrid systems, *IEEE Trans. Sustain. Energy* 13 (2022) 1982–1993, <https://doi.org/10.1109/TSTE.2022.3178415>.
- [40] J. Li, T. Zhou, H. Keke, H. Yu, H. Du, S. Liu, H. Cui, Distributed quantum multiagent deep meta reinforcement learning for area autonomy energy management of a multiarea microgrid, *Appl. Energy* 343 (2023) 121181, <https://doi.org/10.1016/j.apenergy.2023.121181>.
- [41] L. Tightiz, J. Yoo, A robust energy management system for Korean green islands project, *Sci. Rep.* 12 (2022) 22005, <https://doi.org/10.1038/s41598-022-25096-3>.
- [42] S. Gao, C. Xiang, M. Yu, K.T. Tan, T.H. Lee, Online optimal power scheduling of a microgrid via imitation learning, *IEEE Trans. Smart Grid* 13 (2022) 861–876, <https://doi.org/10.1109/TSG.2021.3122570>.
- [43] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, 2016, pp. 770–778. https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html. (Accessed 25 November 2023).
- [44] P. Arévalo, M. Tostado-Véliz, F. Jurado, A novel methodology for comprehensive planning of battery storage systems, *J. Energy Storage* 37 (2021) 102456, <https://doi.org/10.1016/j.est.2021.102456>.
- [45] M. Tostado-Véliz, S. Kamel, H.M. Hasanien, R.A. Turky, F. Jurado, Uncertainty-aware day-ahead scheduling of microgrids considering response fatigue: an IGDT approach, *Appl. Energy* 310 (2022) 118611, <https://doi.org/10.1016/j.apenergy.2022.118611>.
- [46] I. Alsaidan, A. Khodaei, W. Gao, A comprehensive battery energy storage optimal sizing model for microgrid applications, *IEEE Trans. Power Syst.* 33 (2018) 3968–3980, <https://doi.org/10.1109/TPWRS.2017.2769639>.
- [47] M.S. Javadi, M. Gough, A.E. Nezhad, S.F. Santos, M. Shafe-khah, J.P.S. Catalão, Pool trading model within a local energy community considering flexible loads, photovoltaic generation and energy storage systems, *Sustain. Cities Soc.* 79 (2022) 103747, <https://doi.org/10.1016/j.scs.2022.103747>.
- [48] M.E. Baran, F.F. Wu, Network reconfiguration in distribution systems for loss reduction and load balancing, *IEEE Trans. Power Deliv.* 4 (1989) 1401–1407, <https://doi.org/10.1109/61.25627>.
- [49] K. Turitsyn, P. Šulc, S. Backhaus, M. Chertkov, Distributed control of reactive power flow in a radial distribution circuit with high photovoltaic penetration, in: *IEEE PES General Meeting*, 2010, pp. 1–6, <https://doi.org/10.1109/PES.2010.5589663>.
- [50] S.H. Low, Convex relaxation of optimal power flow—part I: formulations and equivalence, *IEEE Trans Control Netw Syst* 1 (2014) 15–27, <https://doi.org/10.1109/TCNS.2014.2309732>.
- [51] M. Farivar, S.H. Low, Branch flow model: relaxations and convexification—part I, *IEEE Trans. Power Syst.* 28 (2013) 2554–2564, <https://doi.org/10.1109/TPWRS.2013.2255317>.
- [52] M. Tostado-Véliz, A.A. Ghadimi, M.R. Miveh, D. Sánchez-Lozano, A. Escamez, F. Jurado, A novel stochastic mixed-integer-linear-logical programming model for optimal coordination of hybrid storage systems in isolated microgrids considering demand response, *Batteries* 8 (2022) 198, <https://doi.org/10.3390/batteries8110198>.
- [53] V.-H. Bui, A. Hussain, H.-M. Kim, Double deep Q\$-learning-based distributed operation of battery energy storage system considering uncertainties, *IEEE Trans. Smart Grid* 11 (2020) 457–469, <https://doi.org/10.1109/TSG.2019.2924025>.
- [54] H. Shuai, J. Fang, X. Ai, Y. Tang, J. Wen, H. He, Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming, *IEEE Trans. Smart Grid* 10 (2019) 2440–2452, <https://doi.org/10.1109/TSG.2018.2798039>.
- [55] E. Foruzan, L.-K. Soh, S. Asgarpoor, Reinforcement learning approach for optimal distributed energy management in a microgrid, *IEEE Trans. Power Syst.* 33 (2018) 5749–5758, <https://doi.org/10.1109/TPWRS.2018.2823641>.
- [56] N. Bhusal, R.M. Shukla, M. Gautam, M. Benidris, S. Sengupta, Deep ensemble learning-based approach to real-time power system state estimation, *Int. J. Electr. Power Energy Syst.* 129 (2021) 106806, <https://doi.org/10.1016/j.ijepes.2021.106806>.
- [57] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, J. He, Short-term load forecasting with deep residual networks, *IEEE Trans. Smart Grid* 10 (2019) 3943–3952, <https://doi.org/10.1109/TSG.2018.2844307>.
- [58] M.S. Hanif, M. Bilal, Competitive residual neural network for image classification, *ICT Express* 6 (2020) 28–37, <https://doi.org/10.1016/j.icte.2019.06.001>.
- [59] L. Zhao, J. Wang, X. Li, Z. Tu, W. Zeng, Deep Convolutional Neural Networks with Merge-and-Run Mappings, 2021, <https://doi.org/10.48550/arXiv.1611.07718>.
- [60] L. Zhang, G. Wang, G.B. Giannakis, Real-time power system state estimation and forecasting via deep unrolled neural networks, *IEEE Trans. Signal Process.* 67 (2019) 4069–4077, <https://doi.org/10.1109/TSP.2019.2926023>.
- [61] Resources – IEEE PES Test Feeder, (n.d.). <https://cmte.ieee.org/pes-testfeeders/resources/> (accessed March 5, 2024).
- [62] Hourly Energy Consumption, (n.d.). <https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption> (accessed September 25, 2023).
- [63] JRC Photovoltaic Geographical Information System (PVGIS) - European Commission, (n.d.). https://re.jrc.ec.europa.eu/pvg_tools/en/tools.html (accessed June 7, 2022).
- [64] T.P. Lillicrap, J.J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, D. Wierstra, Continuous Control With Deep Reinforcement Learning, 2019, <https://doi.org/10.48550/arXiv.1509.02971>.
- [65] T. Haarnoja, A. Zhou, P. Abbeel, S. Levine, Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning With a Stochastic Actor, 2018, <https://doi.org/10.48550/arXiv.1801.01290>.
- [66] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal Policy Optimization Algorithms, 2017, <https://doi.org/10.48550/arXiv.1707.06347>.
- [67] H. Shuai, H. He, Online scheduling of a residential microgrid via Monte-Carlo tree search and a learned model, *IEEE Trans. Smart Grid* 12 (2021) 1073–1087, <https://doi.org/10.1109/TSG.2020.3035127>.
- [68] M. Tostado-Véliz, P. Arévalo, F. Jurado, A comprehensive electrical-gas-hydrogen microgrid model for energy management applications, *Energ. Convers. Manage.* 228 (2021) 113726, <https://doi.org/10.1016/j.enconman.2020.113726>.