scientific reports

Check for updates

OPEN

A robust energy management system for Korean green islands project

Lilia Tightiz & Joon Yoo✉

Penetration enhancement of renewable energy sources is a core component of Korean green-island microgrid projects. This approach calls for a robust energy management system to control the stochastic behavior of renewable energy sources. Therefore, in this paper, we put forward a novel reinforcement learning-driven optimization solution for the convex problem arrangement of the Gasa island microgrid energy management as one of the prominent pilots of the Korean green islands project. We manage the convergence speed of the alternating direction method of multipliers solution for this convex problem by accurately estimating the penalty parameter with the soft actor-critic technique. However, in this arrangement, the soft actor-critic faces sparse reward hindrance, which we address here with the normalizing flow policy. Furthermore, we study the effect of demand response implementation in the Gasa island microgrid to reduce the diesel generator dependency of the microgrid and provide benefits, such as peak-shaving and gas emission reduction.

To combat global warming, substituting RESs for fossil fuels is more beneficial to the island regions of the earth. As global warming continues, oceans rise, and islands disappear[1]. On the other hand, energy storage systems (ESS) mitigate the random nature of RESs, allowing microgrid power networks to take over the island's power supply without relying on central power plants[2,3].

In Korea, with 471 residential islands, using a microgrid to electrify the islands is the most cost-effective solution. In this country, 171 residential islands electrify with stand-alone microgrids, whereas DGs supplied their power earlier[4]. According to the Korean government plan, 63.8 GW of RESs will be installed by 2030. This amount is 20% of national electricity generation and would result in a 37% reduction in CO_2 emissions[5]. Implementing the Green Islands project is a significant step in this direction. One of the seven pilot green islands is Gasa Island, which initially operated as a stand-alone microgrid powered by three DGs[6].

Due to the objective of the Korean green island project, Gasa Island has moved towards entirely relying on RESs with assistance from ESSs. In this self-sufficient microgrid, DGs were used only as backup power in case of contingency. However, it is impossible to achieve this goal without an intelligent, robust EMS unit that can coordinate energy supply resources performance to reduce greenhouse gases. Since the use of ESSs on this island, in addition to the compensation of RESs power absence, is the voltage and frequency regulation, large-scale ESSs are needed. A RES's primary contribution on this island is to reduce greenhouse gas emissions, not to reduce residents' electricity bills[4]. Meanwhile, introducing DR to lower the peak-average ratio reduces the need for further investment in ESSs and RESs besides minimizing DG utilization in emergencies. As a result, in this paper, we examine the energy management of Gasa Island as part of the Green Island project, which aims to implement economic dispatch as a baseline, and then develop EMS using DR programs to reduce DG usage and ultimately greenhouse gas emissions. According to these objectives and the constraints of high penetration of RESs in green islands, in addition to the uncertainty in loads, the Gasa island EMS problem is stochastic, high-dimensional, and sequential. The MG EMS arrangement should consider several factors based on the type of MG, including centralized or decentralized decision making, optimization methods, RESs output, and load uncertainty management. So far, MG EMS optimization has been carried out using a mathematical approach, meta-heuristics, and artificial intelligence (AI)[7]. These EMS scheduling may follow a centralized or decentralized manner. A real-time EMS for microgrid in a decentralized fashion is presented in[8]. This paper examined the MG voltage and frequency stability from a power electronics perspective. Economic model predictive control (EPMC) offered economic dispatch for residential microgrids in[9]. Authors in this paper addressed applying prediction error in RESs output in their model with stochastic MPC in the future attempt. However, it is costly and complex to implement MPC. However, uncertainty in RESs power production is crucial to be considered to provide

School of Computing, Gachon University, 1342 Seongnamdaero, Seongnam-si, Gyeonggi-do 13120, Korea. ✉email: joon.yoo@gachon.ac.kr

nature portfolio

an accurate EMS vastly in literature implemented by deploying probability distribution. Raghav et al.[10] hired a sparrow search algorithm to arrange EMS, considering forecasting errors of RESs output and load prediction. They compared their method with a wide range of other metaheuristic-based optimization methods. Uncertainty in PV and WT power prediction is considered in the optimization of microgrid EMS[11] by deploying different probability density functions. The authors of this paper used the quantum teaching learning-based (QTLBO) optimization method and compared the results with a real-coded genetic algorithm, differential evolution, and TLBO. This paper focuses on the robust arrangement of fast converged optimization methods to support online EMS. Another goal is to reduce DG roles to strengthen the CO_2-free island target through DR. Additionally, this paper deploys historical data of RESs output power and load in the modeling environment considering their randomness behavior.

ADMM is a widely used response to EMS arrangement as a complex problem since of deploying distributed computational approach to reduce the severity of high-dimensional characteristics of microgrid EMS problem[12]. ADMM solved the quadratic and non-quadratic format of the economic dispatch problem of the islanded micro-grid in[13]. The proposed algorithm showed high performance in balancing power consumption and generation of the island microgrid. Lyu et al. in[14] proposed a dual-consensus ADMM to provide less communication-dependent microgrid economic dispatch scheduling. ADMM was used to solve the microgrid economic dispatch considering ESS costs in[15] and[16]. However, a primary concern of hiring ADMM is the hyper-parameter dependency of its convergence to the optimum. Designing dynamic step-size for updating dual variables and converting ADMM to proximal gradient method are remedies has been offered to overcome the convergence hindrance of ADMM[17–20].

Recently, considerable literature has grown up around the theme of sequential prediction of deep learning[21] and in an augmented horizon deep reinforcement learning (DRL) to estimate the penalty parameter of ADMM[22,23]. DRL proved the fast convergence provision of ADMM by justifying penalty parameters in[22]. However, the utilized DRL method was TD3 which suffers from hyper-parameter dependency and worsens the large-scale vulnerability of solving convex problems such as microgrids EMS. Zeng et al.[23] solved the quadratic format of distributed optimal power flow with the help of Q-learning to estimate the penalty parameter of ADMM. By transferring the residual value variant to the reward function in each iteration, the authors in this paper solved the sparse reward problem of arranged reinforcement learning. Moreover, this paper hired reinforcement learning, i.e., Q-learning's discrete action space limited the more credited actions searching possibilities. Therefore, in this paper, we employ the state-of-the-art of SAC to offer a fast and accurate converged ADMM. In contrast to other methods that experience the local optimum trap after a certain number of iterations due to ineffective exploration, SAC supports continuous action space and adds stochasticity to the policy, providing excellent action exploration that persists until the last training process iterations[24]. We deploy the NFP technique to increase the density of the policy probability distribution to overcome the sparse reward issue. The contribution of this paper has four folds:

- Providing a convex problem arrangement of Gasa island EMS considering DR and load flow constraints to make profits for consumers and the utility grid.
- Arrangement of a SAC-based solution to estimate the penalty parameter of ADMM to support the high-dimension and complex problem of Gasa island EMS.
- Solving sparse reward hindrance with a less computational burden on the learning process of SAC algorithm by arranging high-density action space with the NFP approach.
- Exploration of less dependency on conventional generators and acquisition costs with DR implementation.

The remainder of this paper is organized as follows. By specifying the objective function and microgrid elements constraints, "Problem formulation" section formulates the problem. "Proposed method" section represents the solution method, and "Results and discussion" section investigates the novelty of the proposed solution by analysis of the results and compare with benchmark methods. Later on, "Conclusion" section discloses the most relevant conclusions of our work.

Problem formulation

Microgrid objective function.
In this paper, we consider two scenarios for the EMS arrangement of the Gasa island microgrid. The first scenario includes photovoltaic cells (PV), wind turbines (WT), DGs, ESSs, and loads, as shown in Fig. 1. In the second scenario, we schedule the DR for the residential load to decrease the peak-average ratio. This approach reduces DG consumption and carbon emission production, making the green island a more practical objective. Consequently, we define the objective function of the Gasa island microgrid as minimizing power generation costs for the first scenario, and we enhance that by attaching minimization of power consumption cost for consumers through DR implementation. We formulate the objective function of the Gasa island microgrid, considering two scenarios as follows.

$$min\left(\sum_{t=1}^{T}\sum_{i=1}^{N}(c_i^g P_{i,t}^g + c_i^s u_t^s) + \sum_{t=1}^{T}\sum_{j=1}^{M} P_{t,j}^L price_{t,j} \right),$$ (1)

where i: The number of generation units; j: The number of loads; T: The time period of optimization; c^g: The cost of power generation (KRW); P^g: The amount of power generation (kW); c^s: The cost of start-up and shut-down of conventional power generation units (KRW); u^s: Conventional power generation units on/off status; P^L: The amount of power consumption (kW); $price$: The price of load power consumption (KRW).
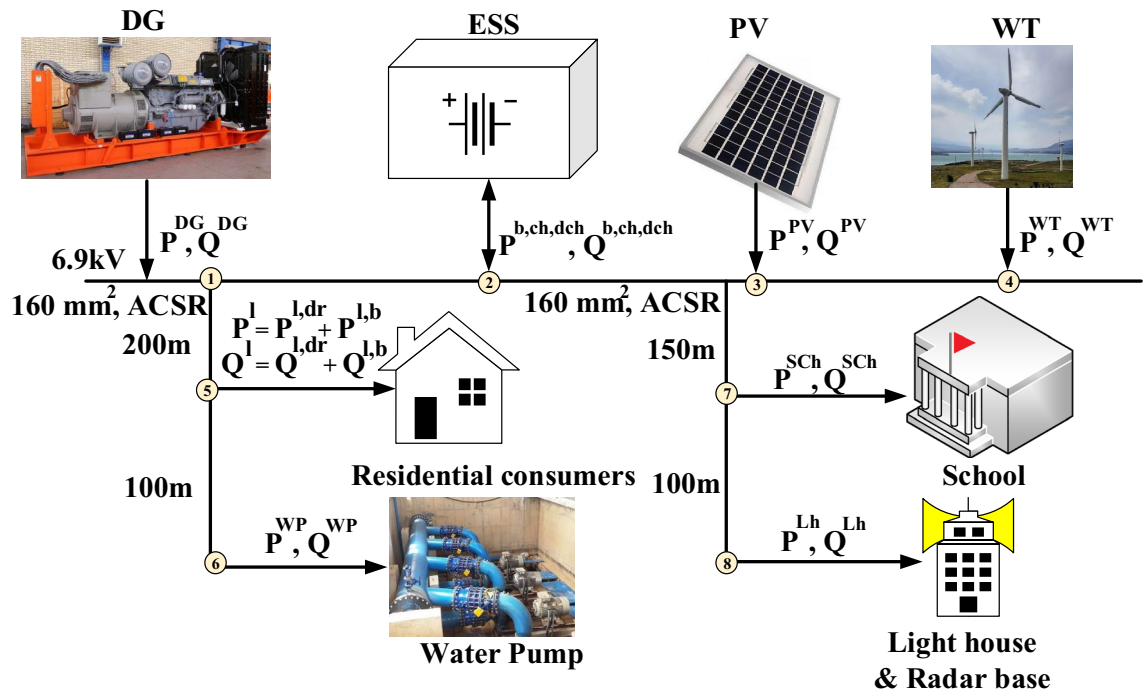
Figure 1. Gasa island microgrid structure.

The first term of (1) calculates the cost of power generation, and the second part is the power consumption expenses. According to Fig. 1, C^g can be applied to power generation units that include PV, WT, DG, and ESS in discharging mode. The cost of start-up and shut-down applies to conventional units, which is DG in this study. Since we plan to reduce peak load through implementing DR, consumption cost minimization is meaningful for this program participant loads.

Microgrid elements modeling and constraints. The main objective of CO_2 reduction in green island advent encourages priority of RESs in supplying loads. Therefore, we decline the cost of RESs power generation. The only RESs constraint is the maximum amount of power that can produce.

$$0 \leq P^{PV}(t) \leq P^{PV}_{max}, \tag{2}$$

$$0 \leq Q^{PV}(t) \leq Q^{PV}_{max}, \tag{3}$$

$$0 \leq P^{WT}(t) \leq P^{WT}_{max}, \tag{4}$$

$$0 \leq Q^{WT}(t) \leq Q^{WT}_{max}, \tag{5}$$

where P^{PV}_{max} and Q^{PV}_{max} are the maximum active/reactive power generation of PVs. P^{WT}_{max} and Q^{WT}_{max} are WTs' highest amount of active/reactive power production.

DG is the only conventional generator in the Gasa island microgrid. In addition to its limitations concerning the amount of energy generated, the DG faces constraints regarding its working duration and variation in power output as follows.

$$P^{DG}_{min} \leq P_{DG} \leq P^{DG}_{max}, \tag{6}$$

$$\left(P^{DG}(t)\right)^2 + \left(Q^{DG}(t)\right)^2 \leq \left(S^{DG}_{max}\right)^2(t), \tag{7}$$

$$\left(T^{DG}_{on}(t) - T^{DG}_{up}\right)\left(u^{DG}(t) - u^{DG}(t+1)\right) \geq 0, \tag{8}$$

$$\left(T^{DG}_{off}(t) - T^{DG}_{down}\right)\left(u^{DG}(t+1) - u^{DG}(t)\right) \geq 0, \tag{9}$$

where T^{DG}_{up} and T^{DG}_{down} determine minimum up and down time, respectively. u^{DG} is a binary value that shows the DG is on or off. The DG's power generation cost is calculated according to (10).

TOU plan	Spring and Autumn		Summer		Winter	
	Time of day	Average price (KRW/kWh)	Time of day	Average price (KRW/kWh)	Time of day	Average price (KRW/kWh)
On-peak	10–12 a.m. and 13–17 p.m.	91.9	10–12 a.m. and 13–17 p.m.	157.7	10–12 a.m. and 17–20 p.m. and 22–23 p.m.	137.3
Off-peak	0–9 a.m. and 23–24 p.m	55.4	0–9 a.m. and 23–24 p.m	55.4	0–9 a.m. and 23–24 p.m	62.3
Mid-peak	Others	70.2	Others	101	Others	98

Table 1. TOU price (p_t^{TOU}) (KRW/kWh) based on KEPCO regulations.

$$c^{DG}(t) = a_1 + a_2 P^{DG} + a_3 \left(P^{DG}\right)^2, \tag{10}$$

where a_1, a_2, and a_3 are factors for the fuel cost of DG. We also consider start-up cost of DG ($c^{s,DG}$) as a fixed number of 10 KRW.

The followings represent the main constraints that attach to the ESSs performance.

$$SoC(t) = SoC(t-1) + \left(P^{b,ch}\eta^{ch} - \frac{P^{b,dch}}{\eta^{dch}}\right)\tau, \tag{11}$$

$$SoC_{min} \leq SoC(t) \leq SoC_{max}, \tag{12}$$

$$-P_{min}^{b,dch} \leq P^b(t) \leq P_{max}^{b,ch}, \tag{13}$$

where η^{ch} and η^{dch} are ESS charging and discharging efficiency. SoC_{min} and SoC_{max} determine up and down limits of state of charge (SoC), and τ is the time slot. This study estimates the battery degradation cost based on (14).

$$c^b(t) = \sigma\tau\left(P^b(t) + E^b(t)\eta^{leakage}\right), \tag{14}$$

$$\sigma = c^{b,inv}/(E^{b,rated}n^{ch,dch}), \tag{15}$$

where $E^{b,rated}$ and $E^{b(t)}$ are rated stored energy and available energy of battery at time t, respectively. $\eta^{leakage}$ determines leakage loss, σ is a coefficient to calculate battery degradation during its lifespan, $c^{b,inv}$ is the initial investment to provide battery, and $n^{ch,dch}$ is the number of full charge and discharge cycles.

Gasa island includes 164 households[25]. To model the residential load, we hire the Enertalk open dataset that consists of the per appliance load consumption of 22 houses in Korea[26]. This dataset provides commonly deployed appliances' power consumption data of a Korean household, including refrigerator, Kimchi refrigerator, water purifier, rice cooker, washing machine, and TV. We aggregated the power consumption of appliances by around 13% based on[27] to estimate the heating and cooling system. We implemented DR on residential consumers to decrease the peak-average ratio. Therefore, we categorize the Enertalk appliance records into non-controllable, shiftable, and reducible loads. TV, Kimchi refrigerator, refrigerator, and rice cooker are non-controllable loads. The washing machine and heating and cooling system are shiftable and reducible loads, respectively. Therefore the electricity consumption of the heating and cooling system during the DR program based on inside building temperature ($Temp^{in}$) has the following restrictions[28,29].

$$\Delta Temp^{in}(\tau) = \frac{1}{\beta\gamma}\left(Temp^{out}(t-1) - \beta P^{H\&C}(t)\right), \tag{16}$$

$$Temp_{min}^{in} \leq Temp^{in}(t) \leq Temp_{max}^{in}, \tag{17}$$

$$P_{min}^{H\&C} \leq P^{H\&C}(t) \leq P_{max}^{H\&C}, \tag{18}$$

where $\Delta Temp^{in}(\tau)$ is the inside building temperature deviation during time step τ. $Temp_{min}^{in}$ and $Temp_{max}^{in}$ are the inside building maximum and minimum desirable temperature. β and γ are the building thermal capacitance (kWh/ °C) and reactance (°C/kW), respectively. $P^{H\&C}(t)$ is the power usage of the heating or cooling system in time t.

To calculate the cost of load, we consider the time of use (TOU) price (p_t^{TOU}) according to Table 1[30].

To provide a trade-off between utility profit in decreasing peak-average ratio with higher TOU in peak hours and consumer comfort, we add the anxiety rate term ($\mathcal{A}_{rate}$) to the cost function to respect consumers' preferred power consumption rate.

In the case of the heating and cooling system, the DR affects the customer comfort where the inside temperature ($Temp^{in}(t)$) exceeds limitations that are defined in (17)[31].

$$\mathcal{A}_{rate}^{H\&C} = \begin{cases} 0, & Temp_{min}^{in} \leq Temp^{in}(t) \leq Temp_{max}^{in}, \\ Ke^{K(Temp^{in} - Temp_{max}^{in})}, & Temp^{in} > Temp_{max}^{in}, \\ Ke^{K(Temp^{in} - Temp_{min}^{in})}, & Temp^{in} < Temp_{min}^{in}, \end{cases} \tag{19}$$

where K is the anxiety rate coefficient.

Additionally, according to the customers' preferred time of using the washing machine, the anxiety rate of the washing machine ($\mathcal{A}_{rate}^{WM}$) participating in DR is determined as follows.

$$\mathcal{A}_{rate}^{WM} = \begin{cases} 0, & h_s < h_{WM} < h_f, \\ \sigma^{h_s - \zeta}, & h_{WM} < h_s, h_{WM} > h_f, \end{cases} \tag{20}$$

where h_s and h_f are the lower and upper preferred times of using the washing machine, respectively. σ and ζ determine penalty coefficients of shifting washing machines out of consumers' preferred time.

We deploy the water station load profile from[32] and simulate school daily power consumption from[33]. The amount of lighthouse and radar base power consumption of Gasa island in the average daily load profile decline in this paper. During power dispatch scheduling, we consider load flow constraints for each branch of the Gasa island microgrid as follows.

$$P_i^g(t) - P_i^l(t) = \sum_{i=1}^{n} V_i(t)V_j(t)\big(G_{i,j}Cos\theta_{i,j} + B_{i,j}Sin\theta_{i,j}\big), \tag{21}$$

$$Q_i^g(t) - Q_i^l(t) = \sum_{i=1}^{n} V_i V_j\big(G_{i,j}Sin\theta_{i,j} - B_{i,j}Cos\theta_{i,j}\big), \tag{22}$$

where,

$$P_i^g(t) = P_i^{DG}(t) + P_i^{PV}(t) + P_i^{WT}(t) + P_i^{b,ch}(t) - P_i^{b,dch}(t), \tag{23}$$

$$P_i^l(t) = P_i^{SCh}(t) + P_i^{WP}(t) + P_i^{Lh}(t) + P_i^{l,b} + P_i^{l,dr}, \tag{24}$$

$$P_i^{l,dr}(t) = P_i^{H\&C}(t) + P_i^{WM}(t) - \Delta P_i^{H\&C,dr}(t) - \Delta P_i^{WM,dr}(t), \tag{25}$$

$$Q_i^g(t) = Q_i^{DG}(t) + Q_i^{PV}(t) + Q_i^{WT}(t) + Q_i^{b,ch}(t) - Q_i^{b,dch}(t), \tag{26}$$

$$Q_i^l(t) = Q_i^{SCh}(t) + Q_i^{WP}(t) + Q_i^{Lh}(t) + Q_i^{l,b} + Q_i^{l,dr}, \tag{27}$$

$$Q_i^{l,dr}(t) = Q_i^{H\&C}(t) + Q_i^{WM}(t) - \Delta Q_i^{H\&C,dr}(t) - \Delta Q_i^{WM,dr}(t), \tag{28}$$

$$V_{min}^i \leq V^i(t) \leq V_{max}^i, \tag{29}$$

where i and j are bus indexes. $P^{i,j}(t)$ and $Q^{i,j}(t)$ are active and reactive power flow of lines between i and j buses at time t. V_{min}^i and V_{max}^i are lower and upper voltage limitations of each bus. P^{SCh}, P^{WP}, and P^{Lh} are the school, water pump station, and lighthouse active power consumption, respectively. $P^{l,b}$ denotes non-DR participants' residential load, and $P^{l,dr}$ is DR participants' residential load. $\Delta P^{H\&C,dr}$, and $\Delta P^{WM,dr}$ are the amount of heating and cooling systems and washing machines' active power consumption that contributes to DR. We modify the objective function in (1) based on Gasa island microgrid constraints as follows.

$$Min\left(\sum_{t=1}^{T}\left(c^{DG}P_t^{DG} + c^b P_t^b + c^{s,DG}u_t^{DG}\right) + \sum_{t=1}^{T} P_t^L p_t^{TOU}\right)$$
$$= Min\left(\sum_{t=1}^{T}\left(a_1 + a_2 P^{DG} + a_3(P^{DG})^2 + \alpha\,(SoC(t) - SoC(t-1))^2 P_t^b + c^{s,DG}u_t^{DG}\right)\right. \tag{30}$$
$$\left. + \sum_{t=1}^{T}\left(P_t^{SCh} + P_t^{WP} + P_t^{Lh} + P_t^{H\&C} + P_t^{WM} - \Delta P_t^{H\&C}\mathcal{A}_{rate}^{H\&C,dr} - \Delta P_t^{WM}\mathcal{A}_{rate}^{WM,dr}\right)p_t^{TOU}\right),$$

subject to (2)–(9), (12), (13), (16)–(29). This formulation arranges a quadratic form of convex problem, which is solved in the following section based on ADMM technique.

Proposed method
ADMM. ADMM is a popular and reliable approach for convex quadratic programming. The conventional ADMM method attempts to solve the following problem[34]:

$$Min f(x) + g(x'), \qquad x \in \mathbb{R}^n, x' \in \mathbb{R}^{n'}, \tag{31}$$

subject to

$$Ax + A'x' = d, \qquad A \in \mathbb{R}^{p \times n}, A' \in \mathbb{R}^{p \times n'}, d \in \mathbb{R}^p, \tag{32}$$

where x and x' are the optimization variables, f and g are coefficient matrices, d is the sequence, and f(x) and $g(x')$ are convex functions of optimization variables.

We arrange the Lagrangian function for the convex problem (31) as follows.

$$L(x, x', \lambda) = f(x) + g(x') + \lambda^T (Ax + A'x' - d) + \frac{\rho}{2} \|Ax + A'x' - d\|^2, \ \lambda \in \mathbb{R}^{n''}, \tag{33}$$

where λ is a relevant vector of multiplier for constraints (32), and ρ is the penalty factor.

ADMM iteratively updates optimization variables according to the following.

$$x_{k+1} = \underset{x}{argmin}\, L_\rho(x, x'_k, \lambda_k), \tag{34}$$

$$x'_{k+1} = \underset{x'}{argmin}\, L_\rho(x_k, x', \lambda_k), \tag{35}$$

$$\lambda^{k+1} = \lambda_k + \rho(Ax_{k+1} + A'x'_{k+1} - d), \tag{36}$$

This iterative procedure will converge when the primal and dual residuals of ADMM techniques meet their thresholds ϵ_p and ϵ_d, respectively, as follows.

$$\| r_k^p \|_2 \le \epsilon_p, \tag{37}$$

$$\| r_k^d \|_2 \le \epsilon_d, \tag{38}$$

where $\| r_k^p \|$ and $\| r_k^p \|$ represent the primal and dual residuals of the ADMM technique, calculated according to (39) and (40), respectively.

$$\| r_k^p \| = Ax_k + A'x'_k - d, \tag{39}$$

$$\| r_k^d \| = 2\rho A^T A'(x'_k - x'_{k-1}). \tag{40}$$

We deploy the decompose technique introduced in[23] to solve our ADMM problem. To this end, we consider two penalty parameter vectors as follows.

$$\rho_{PQ} \in \{\rho_1^{PQ}, \rho_2^{PQ}, \ldots, \rho_{n_{PQ}}^{PQ}\}, \tag{41}$$

$$\rho_{V\theta} \in \{\rho_1^{V\theta}, \rho_2^{V\theta}, \ldots, \rho_{n_{V\theta}}^{V\theta}\}, \tag{42}$$

where ρ_{PQ} is the penalty parameter for active and reactive power and $\rho_{V\theta}$ denotes voltage and phase angle penalty parameter. n_{PQ} and $n_{V\theta}$ are the number of constraints for each pair of active and reactive power and voltage and phase angle, respectively.

Since penalty parameters are determinant factors in ADMM dual and primal in each iteration to converge according to (37) and (38), the sequential characteristic of these parameters calculation encourages using the DRL method to estimate penalty parameters.

SAC algorithm with enhanced exploration. SAC is an actor-critic DRL method with sample efficiency specification of its off-policy approach[24]. SAC works in both discrete and continuous environments. The main characteristic of SAC is stochastic-based policy optimization by adding entropy to policy. This characteristic gives advantages of productive exploration and, consequently, a higher convergence rate to off-policy methods, such as deep deterministic policy gradient (DDPG) and twin delayed DDPG (TD3). Its sample efficiency due to learning from experience saved in the replay buffer is superior to on-policy techniques such as proximal policy optimization (PPO) and trust region policy optimization (TRPO). The entropy term ($\mathcal{H}(.)$) is defined according to (29). The portion of entropy in the learning policy is determined by temperature parameter α, which is decreased during the learning iteration as follows.

$$\pi^* = argmax_\pi \sum_t \left(\mathbb{E}_{(s_t,a_t) \sim \rho_\pi} [r(s_t, a_t) + \alpha \mathcal{H}(\pi(a_t|s_t))] \right), \tag{43}$$

$$H(\pi(a_t|s_t)) = \mathbb{E}_{a \sim \pi(.|s)} \left[-log(\pi(a|s)) \right]. \tag{44}$$

The entropy term will update the Bellman equation of the value network training process according to (45).

6

$$V(s_t) = \mathbb{E}_{a \sim \pi_\phi(a_t|s_t)} \big[Q_\theta(s_t, a) - \alpha log(\pi_\phi(a|s_t)) \big]. \tag{45}$$

The critic in SAC includes value function V_ψ and soft Q-function Q_θ. The actor contains policy network Q_ϕ. The policy network in SAC chooses action from Gaussian probability distribution according to the squashing function $f_\phi(\epsilon; s_t)$ as follows.

$$a_t = f_\phi(\epsilon_t, s_t) = tanh(\mu_\phi(s_t) + \sigma_\phi(s_t)\epsilon_t), \quad \epsilon \in \mathcal{N}(0, 1). \tag{46}$$

To use SAC for parameter estimation of the ADMM, we need to arrange the Markov decision process (MDP), including state, action, transition function, and reward. The state will be the decision variables of the ADMM problem, which are the Gasa island microgrid elements' active and reactive power as follows.

$$S(t) = \{P^g, Q^g, P^l, Q^l, u^{DG}, V_i, \theta_{i,j}\}. \tag{47}$$

SAC will predict the suitable penalty parameter from the continuous action space according to (48), (49).

$$100 \leq \rho_{PQ,i} \leq 10^4, \quad 1 \leq i \leq n_{PQ}, \tag{48}$$

$$30 \leq \rho_{V\theta,j} \leq 3000, \quad 1 \leq j \leq n_{V\theta}. \tag{49}$$

The reward function is defined as follow.

$$r(s_t, a_t) = \begin{cases} 200, & if\ (37),(38), \\ 0, & otherwise. \end{cases} \tag{50}$$

However, this reward function is a sparse reward. The NFP is a trick that is used in this study to empower stability and provision of efficient action space exploration to defeat the sparse reward issues[35]. A set of invertible functions establishes normalizing flows. With the change of probability distribution variables, normalizing flows sequentially transform a distribution to a more density distribution as follows[36].

$$z_N = f_1 \circ f_2 \circ \cdots \circ f_{N-1} \circ f_N(z_0), \tag{51}$$

where z_0 is the base distribution and z_N is the final flow. The density of continuous variable the z_N parametrized by ϕ is as follows.

$$lnp_\phi(z_N) = ln\ p(z_0) - \sum_{n=1}^{N} ln\ det|(\partial f_N / \partial z_{N-1})|, \tag{52}$$

One of the simplest methods to determine invertible function f is RealNVP. We deployed the method introduced in[37] to combine SAC and normalizing flows. We will reparametrize the Gaussian distribution of action selection with RealNVP invertible transformation as follows.

$$\pi(a|s) = tanh(f_1(\mu_\phi(s) + \sigma_\phi(s)\epsilon) \circ f_N), \tag{53}$$

where $\pi(a|s) = z_0$ and the log density of action is as follows.

$$log\pi(a|s) = log\ p(z_0) - \sum_{n=1}^{N} log\ det|(\partial f_N / \partial z_{N-1})| - \sum_{n=1}^{N} log(1 - tanh((z_{N_n})^2), \tag{54}$$

Therefore, the normal SAC will be modified by adding a gradient step on the normalized flow layers during the ϕ setting.

The proposed technique to determine penalty parameter of ADMM with the contribution of SAC and NFP is represented in Fig. 2.

Results and discussion

We deploy our proposed algorithm to optimize the Gasa island microgrid EMS in this section. The optimization takes place hourly and can be extended to shorter intervals based on available data. Figure 3 depicts the load profile for the island power consumers based on "Microgrid elements modeling and constraints" section. Since the residential load profile is for August, we choose the WT and PV output power taken from[38] and the output temperature[39] according to this period, as shown in Fig. 4. We simulate the microgrid and implement a centralized ADMM solution as a baseline and our optimization solution with the CVXPY package[40] to use OSQP solver[41] in the Python environment. The simulations are accomplished on a PC with Intel(R) Core (TM) i5-10400F CPU @ 2.90GHz.

The SAC algorithm's performance was compared with different hidden layer numbers and batch sizes to justify DNN parameters. Figure 5 shows how the learning process of DNN varies depending on hyperparameters. Based on the results, two layers with a batch size of 128 will result in a trade-off regarding learning convergence, stability, and complexity. It will not be advantageous to extend the network size to three layers and the batch size to 256 according to Fig. 5b. Table 2 depicts the microgrid element specifications and solution algorithm parameters.
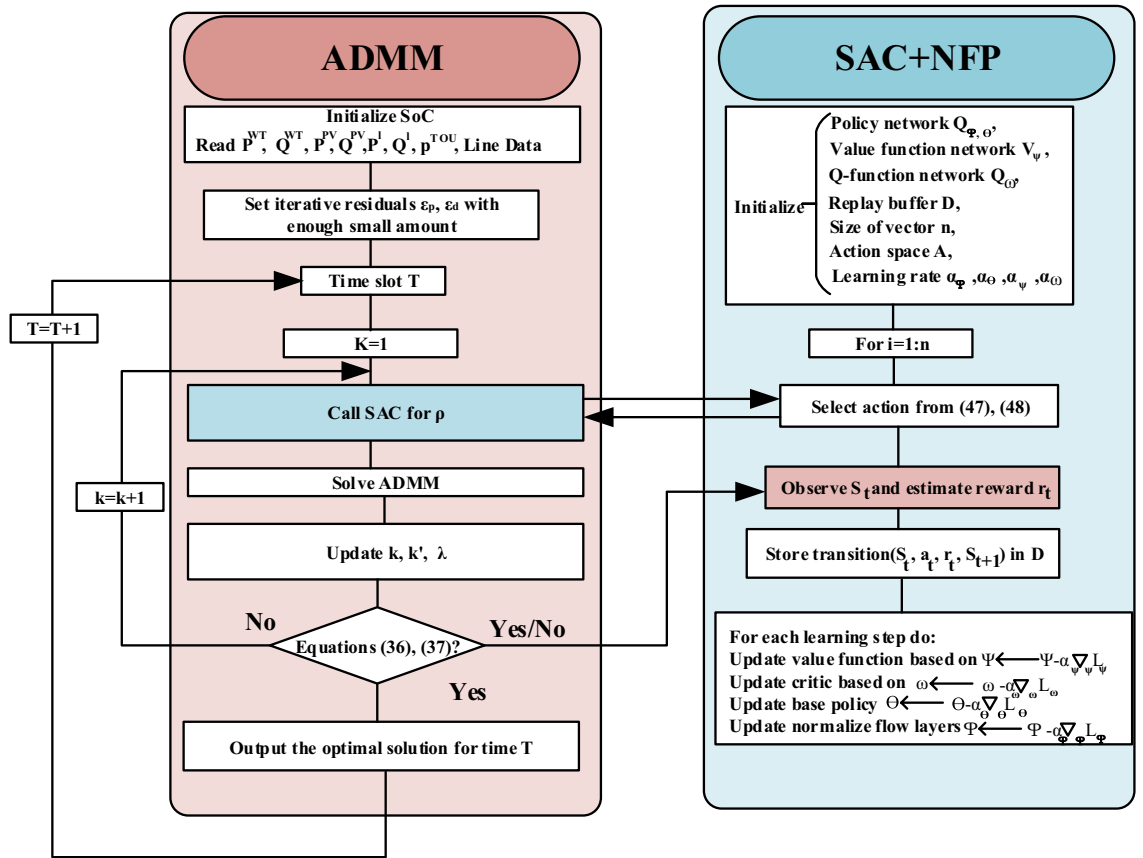
Figure 2. The proposed optimization algorithm for Gasa island EMS based on ADMM-NFP-SAC.

Figure 6 demonstrates the convergence speed comparison of the proposed algorithm and conventional ADMM. The effectiveness of our hired technique appears in this figure, where surplus convergence of vanilla ADMM based on the number of iterations. By learning the best policy, the RL agent speeds up convergence for dual and primal residuals to 300 and 500 iterations, respectively, while for normal ADMM, this process time is double.

Figure 7 shows the Gasa island microgrid economic power dispatches along with the SoC level of BESS with and without DR. This figure illustrates the effectiveness of DR planning in reducing DG utilization during the understudy day. Without DR, the DG starts to work with a higher power generating amount of 315 kW at 1 a.m. to compensate for the shortage of RESs power shown in Fig. 7a. However, DR deployment causes this amount to reduce to 210 kW, as can be seen in Fig. 7b. Additionally, without DR, DG, due to ramp-down time limitations, should stay working on 315 kW, although the amount of shortage power is lower than this amount of generation. Moreover, we also solved the QP arrangement of Gasa island with the CPLEX solver as an analytical method. Table 3 represents our proposed SAC-NF-ADMM to improve the performance of ADMM in case of operational cost for under study day.

Figure 8 shows the voltage magnitude of the whole Gasa island power network nodes. This figure delineates the voltage magnitude of the island power network stay in the allowed range between 0.99 and 1.01 p.u.

The effects of DR scheduling on the cooling system and washing machine are represented in Figs. 9 and 10. The washing machine takes part in the DR program by adjusting operating hours to off-peak and mid-peak with lower TOU pricing. The results of DR implementation show that the anxiety rates perfectly direct the optimization problem to keep the desired time of washing machines' working hours. As we discussed before, the washing machine's preferred time of working is between 4 and 12 a.m. Therefore, as can be seen in Fig. 9, the washing machine power usage between 9 and 5 p.m. with higher TOU transferred to other times of day in the desired range.

The optimization algorithm justifies indoor temperature in the desired temperature between 23 and 25 °C. The inside temperature tends to be higher between 10 a.m. and 5 p.m., where TOU is the highest amount, resulting in less power consumption for the cooling system compared to situations without DR schedules.

Figure 11 reveals the DR scheduling resulted in peak shaving of around 20% for residential load during peak hours. The green color shaded area in Fig. 11 shows peak shaving, while a red color shaded area signifies a transferred portion of peak load relating to washing machine usage. The most significant benefit of DR scheduling for consumers is dropping 21% in total consumer electricity bills, as illustrated in Fig. 12. After DR implementation, DG cost dropped by around 42%, which is another evidence of reduced DG usage and fewer gas emissions.

Microgrid elements	
DGs	
$P_{DG,min/max}$ (kW/step)	75/330
$P_{DG,ramp-down/up}$ (kW)	120
$T_{DG,up/down}$ (h)	2
Quadratic coefficients	
a_1	1.3
(₩/kW)	0.0304
(₩/kW^2)	0.00104
RESs	
$P_{PV,max}$ (kW)	314
$P_{WT,max}$ (kW)	400
BESS	
E_{bat} (MWh)	3
$P_{bat,min/max}$ (kW)	− 300/300
$SoC_{min/max}$ (%)	20/90
η_{bat} (%)	90
σ (KRW/MW)	100
$\eta^{leakage}$	3%/month
Load	
$Temp_{min}^{in}$ (°C)	23
$Temp_{max}^{in}$ (°C)	25
Building thermal capacitance β (kWh/°C)	0.8
Building thermal reactance γ (°C/kW)	− 0.02
Heating and cooling anxiety rate coefficient (K)	100
Lower/Upper prefered time (h_s/h_f)	4/12 (a.m.)
Washing machine anxiety rate coefficients (σ, ζ)	50
Algorithm parameters	
SAC	
Learning rate	0.0001
Discount factor (γ)	0.9
Replay buffer size	50,000
Number of training episodes	5,000
Mini-batch size	128
Number of hidden layer	2
$\alpha_\phi, \alpha_\theta, \alpha_\psi, \alpha_\omega$	0.0003
Activation function	ReLU
Optimizer	SGD
ADMM	
$\rho_{min/max}^{PQ}$	100/10^4
$\rho_{min/max}^{V\theta}$	30/3000
ϵ_p, ϵ_d	0.1, 0.001

Table 2. The Gasa island EMS solution simulation technical specifications and constraints.

Here, we deployed the TOU policy to schedule DR scheduling currently utilized in the Korean power system. TOU is one of the price-based DR policies. It is also possible to implement DR on Gasa Island using other price-based DR models, such as real-time pricing (RTP) and critical peak pricing (CPP). Furthermore, incentive-based DR methods, including direct load control and emergency demand response programs (EDRP), can be used jointly with price-based DR with incentive payments from utilities to increase customer profits and encourage them to participate in DR. Therefore, in our future investigation, we will use a combination of price-based and incentive-based DR techniques to study their effect on CO_2 emission reduction. On the other hand, to completely meet the current situation of Gasa island, we deployed the WT and PV output power historical data. However, in our future attempt to consider uncertainties in WT and PV power generation, we deploy long short-term memory (LSTM) based solution to provide the RESs predictor.
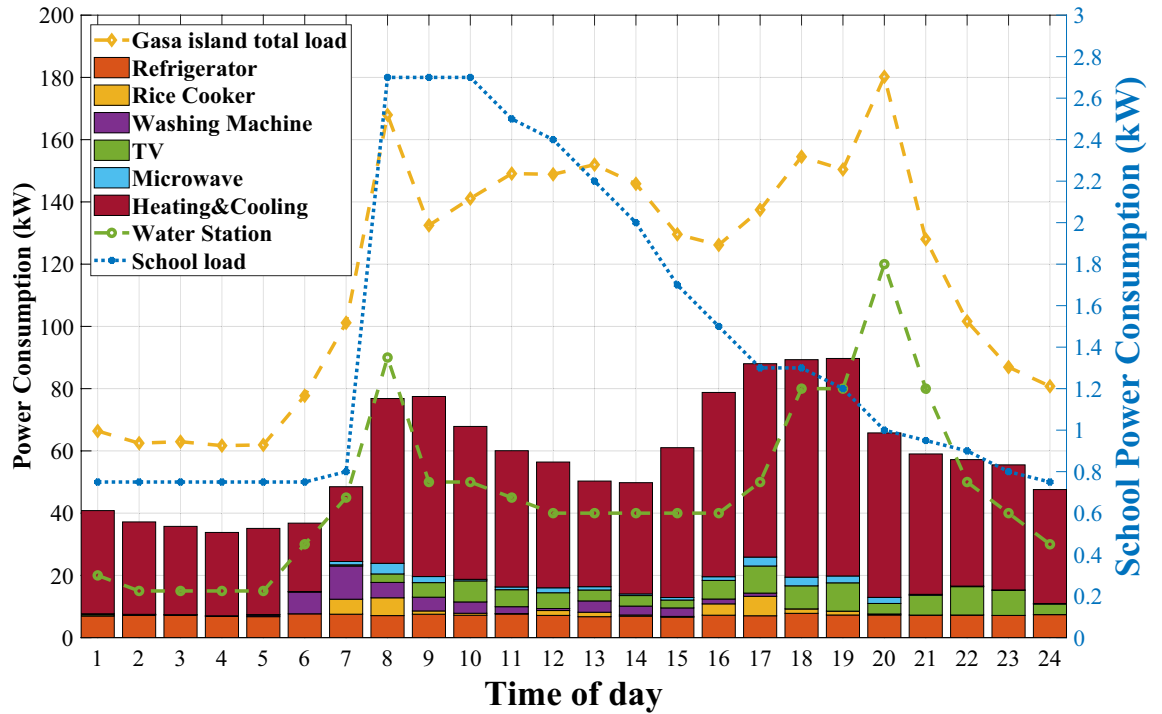
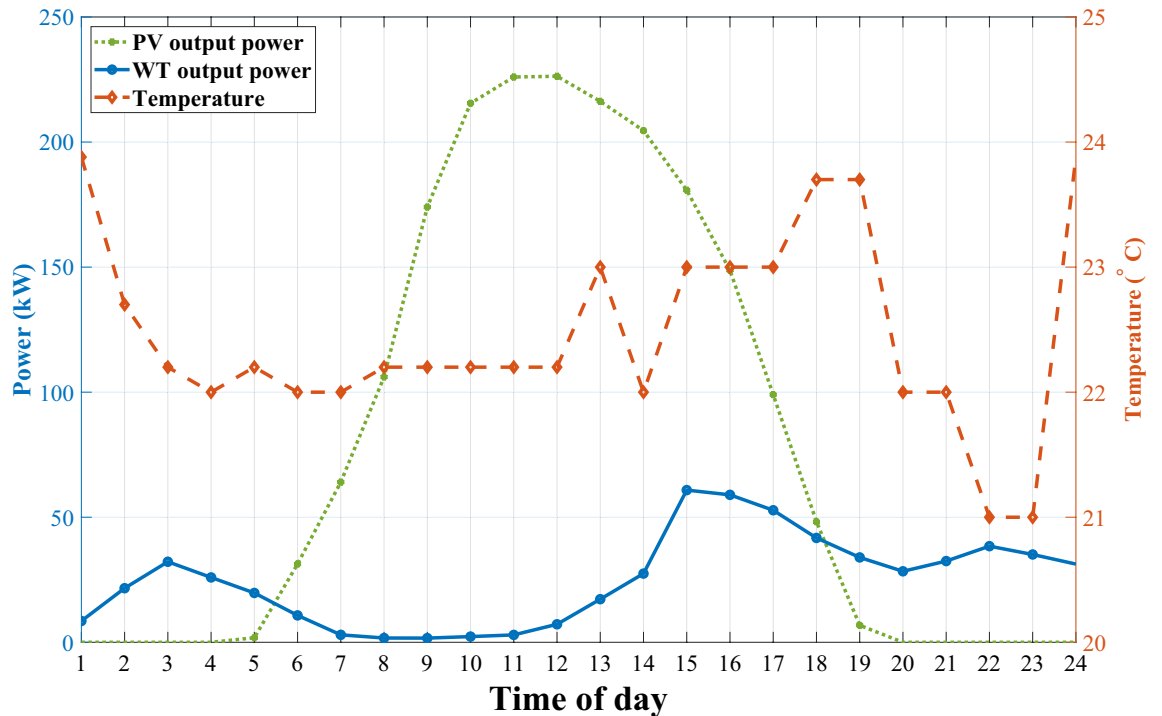**Figure 3.** The Gasa island load profile based on consumers.

Figure 4. The Gasa island RESs output power and daily temperature during August.

Conclusion

In this paper, we arranged the EMS unit for the Gasa island microgrid as one of the prominent Korean Green island project pilots. The proposed approach improves the main objective of this green island microgrid from the feasible framework for RESs utilization to a profitable microgrid for consumers with DR deployment. Additionally, our method resulted in less dependency on DGs with DR schedules. The ADMM-based solution for EMS provided a fast converged process of optimization by penalty parameter prediction with the state-of-the-art DRL method SAC. We released each iteration from the computational cost of transferring dual variable
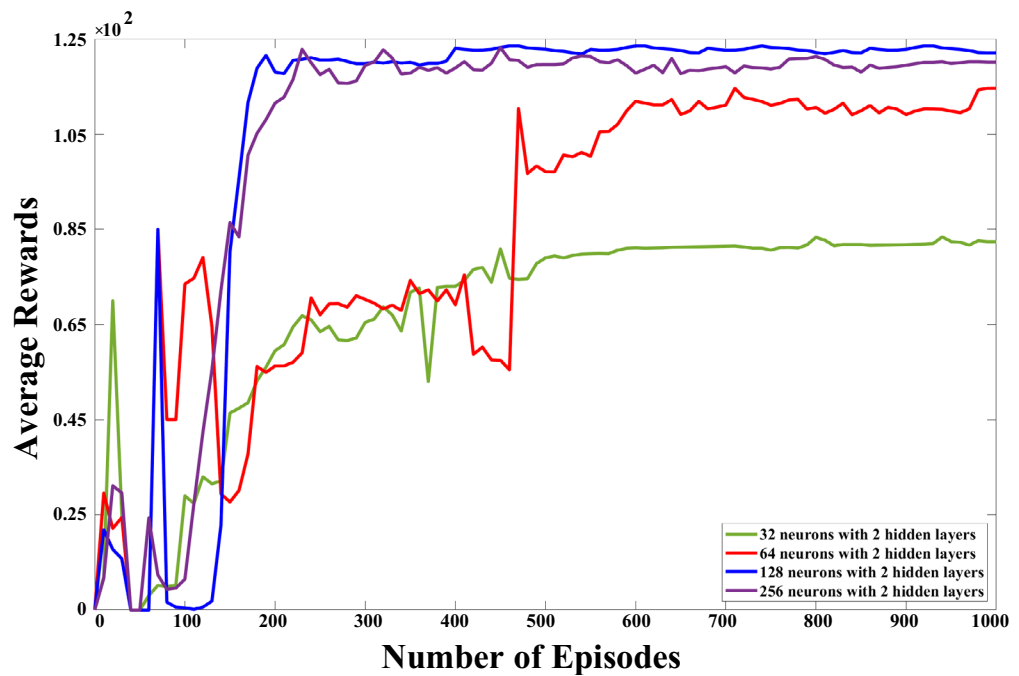
(a) SAC Learning process with 2 layers of DNN and different sizes of batch
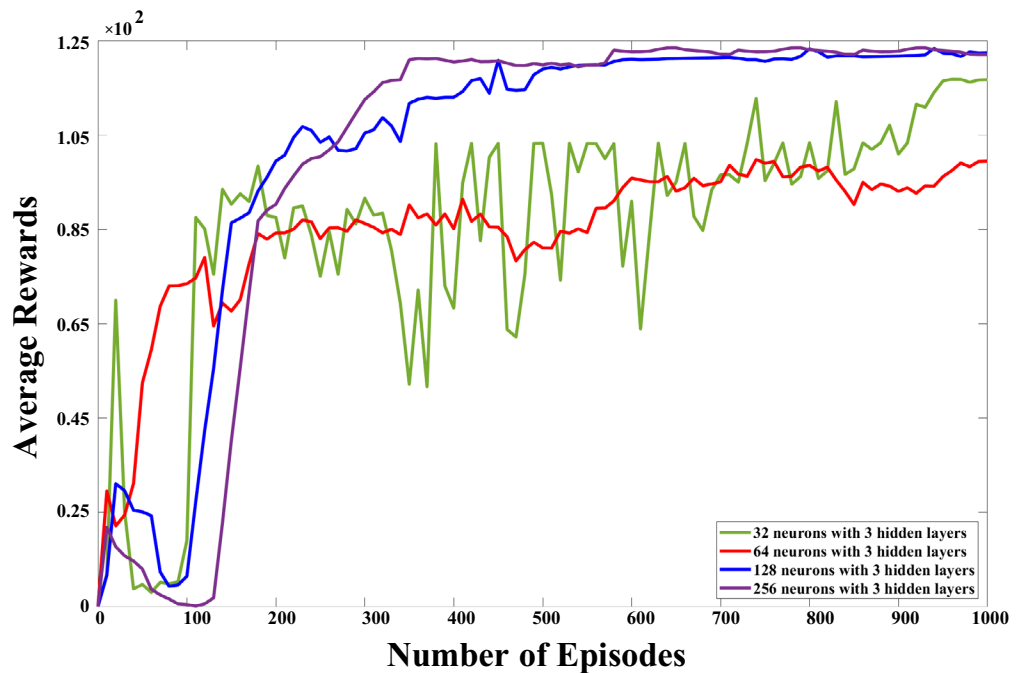
(b) SAC Learning process with 3 layers of DNN and different sizes of batch

Figure 5. SAC hyperparameters justification.

variants by definition of an independent, constant reward to each converged iteration. However, this approach resulted in a sparse reward hindrance for the process of training the agent. To overcome this problem, we used NFP to increase the probability distribution of policies. The results showed the proposed ADMM converged 50% faster than vanilla ADMM. Additionally, the implemented DR scheduling on reducible and shiftable residential load decreased 20% of the peak load. Since in this paper we considered current situation of Gasa island microgrid network TOU which is utilized DR policy in Korean power system conisdered to implement DR. In our future
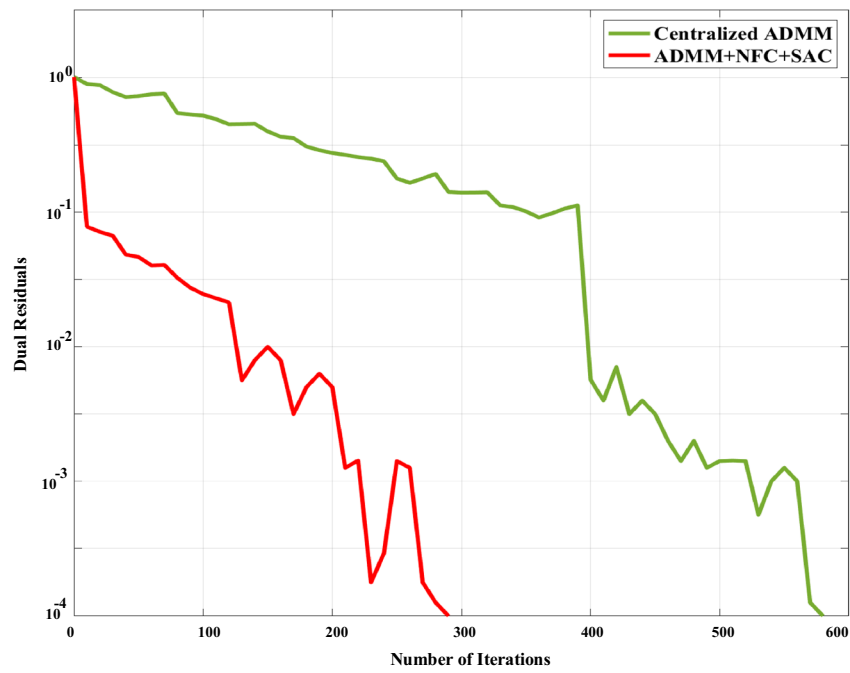
(a) Dual residual
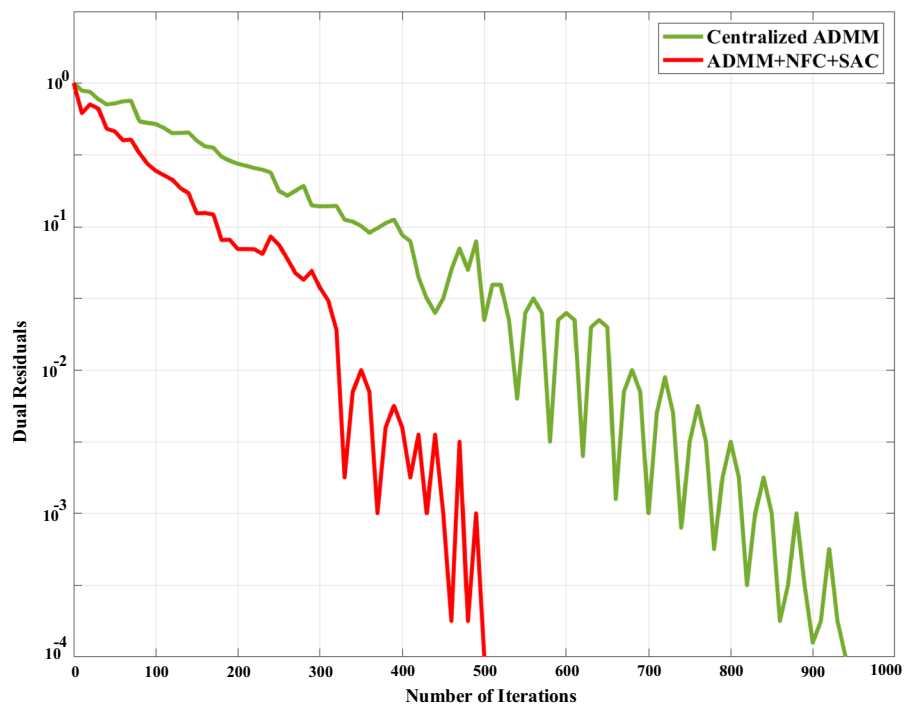
(b) Primal residual

Figure 6. Convergence rate of residuals with centralized ADMM and proposed method comparison.
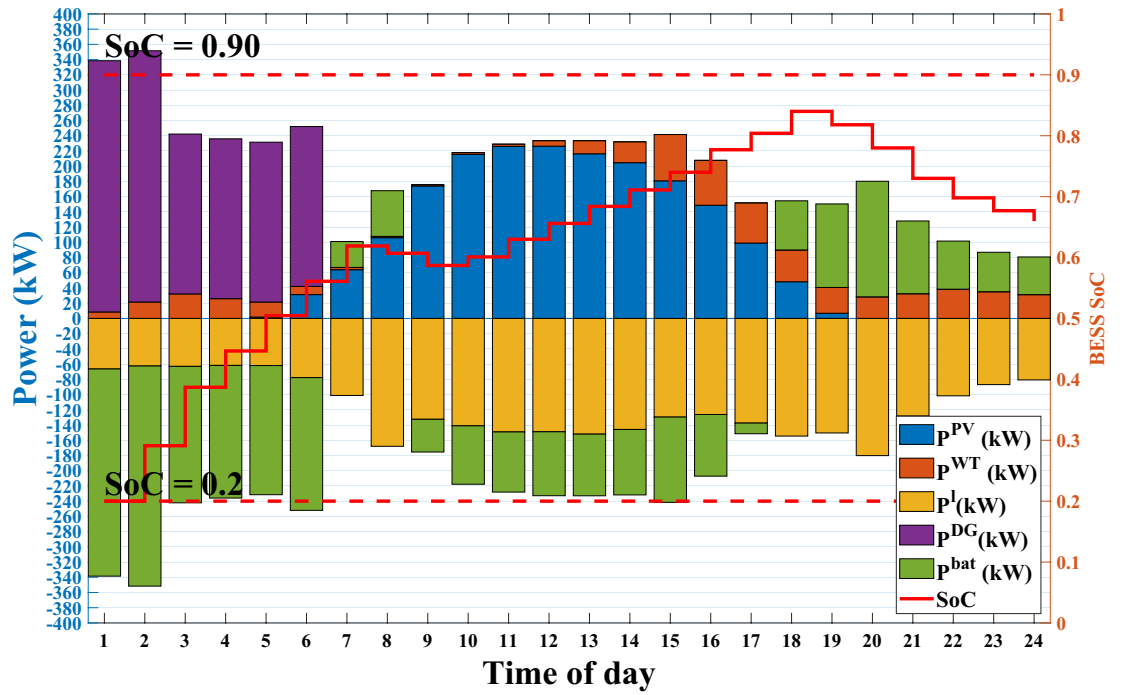
(a) Gasa island economic dispatch without DR
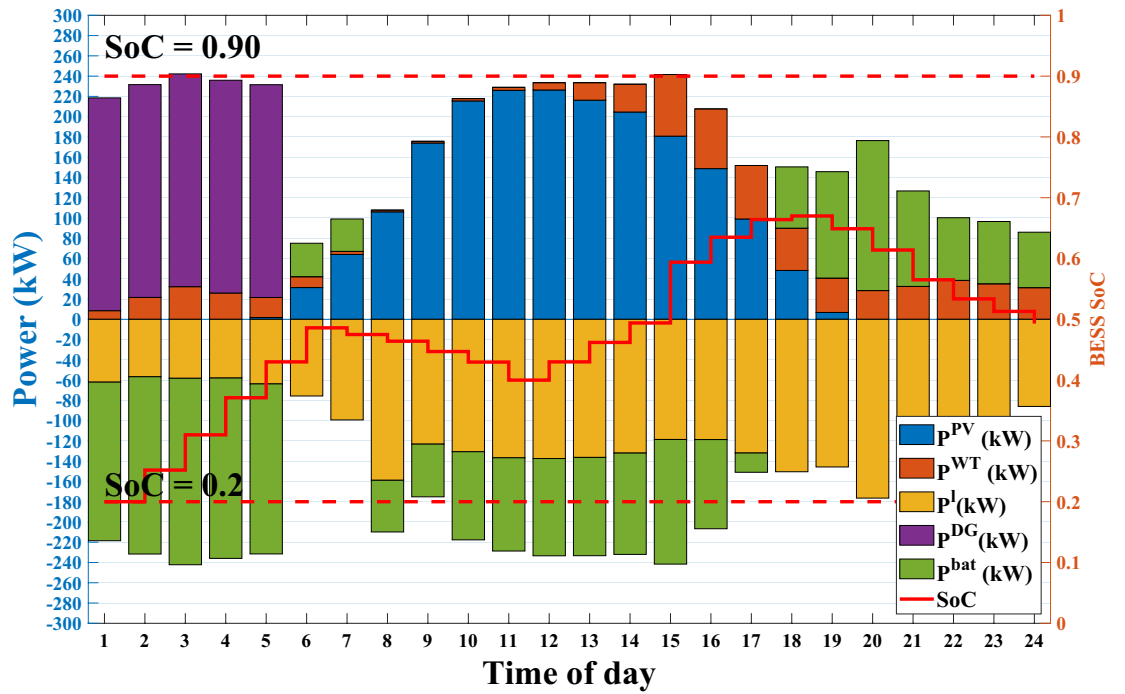
(b) Gasa island economic dispatch with DR

Figure 7. Economic dispatch in Gasa island microgrid with and without DR.
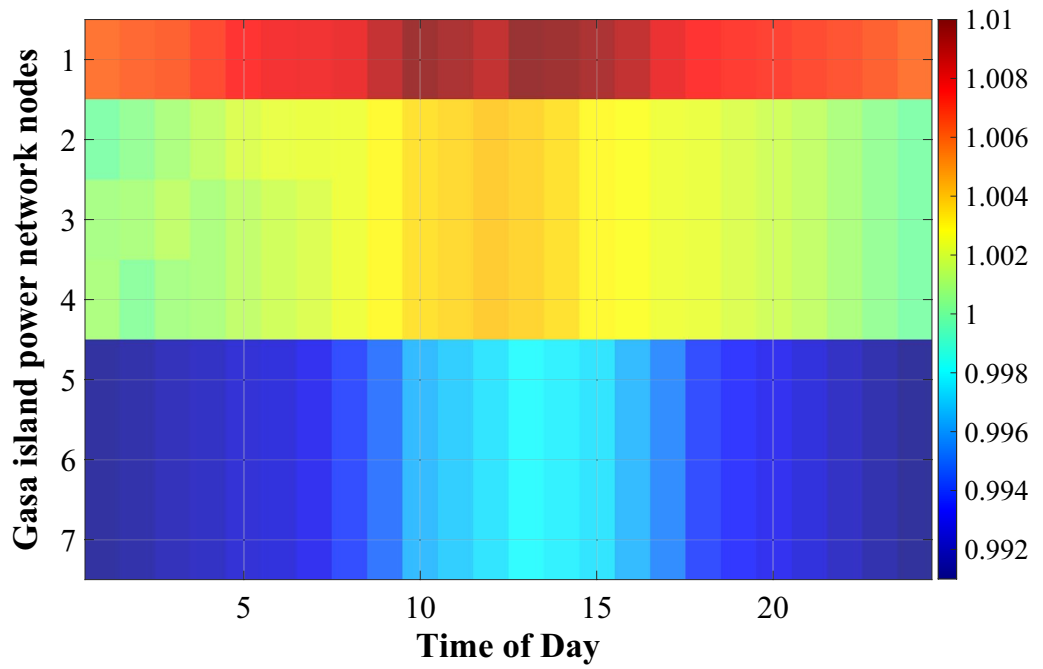
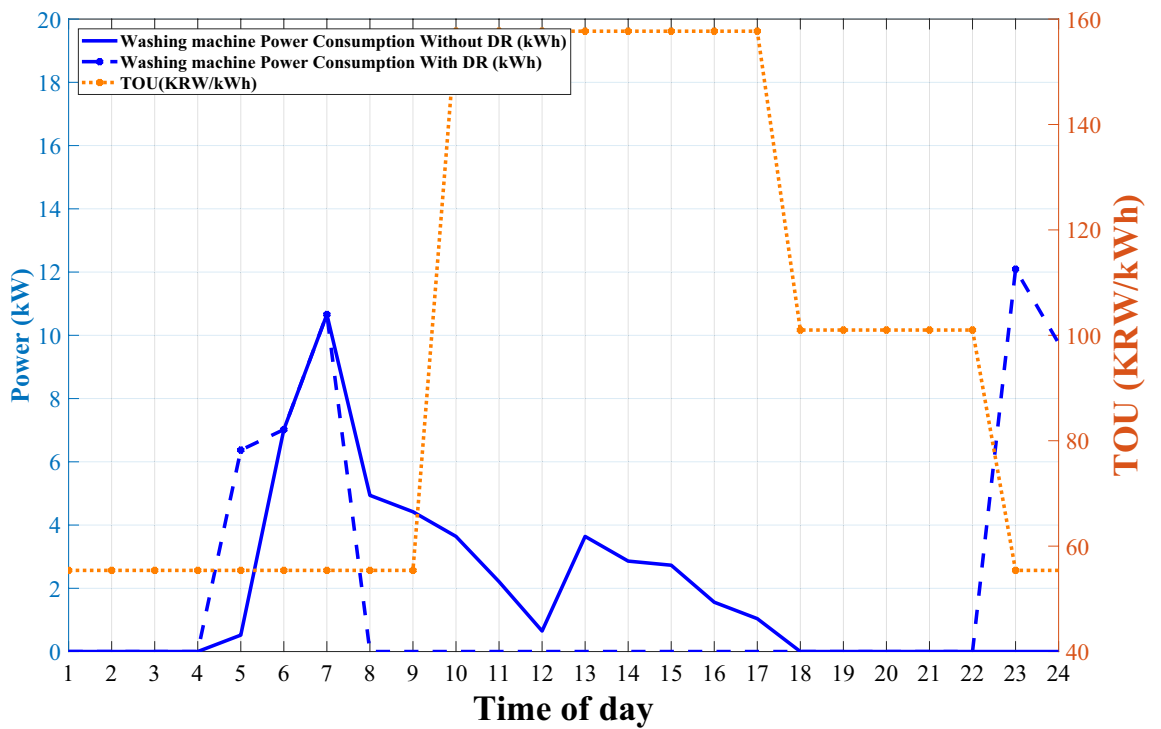**Figure 8.** Gasa islands power grid nodes voltage magnitude (p.u.).

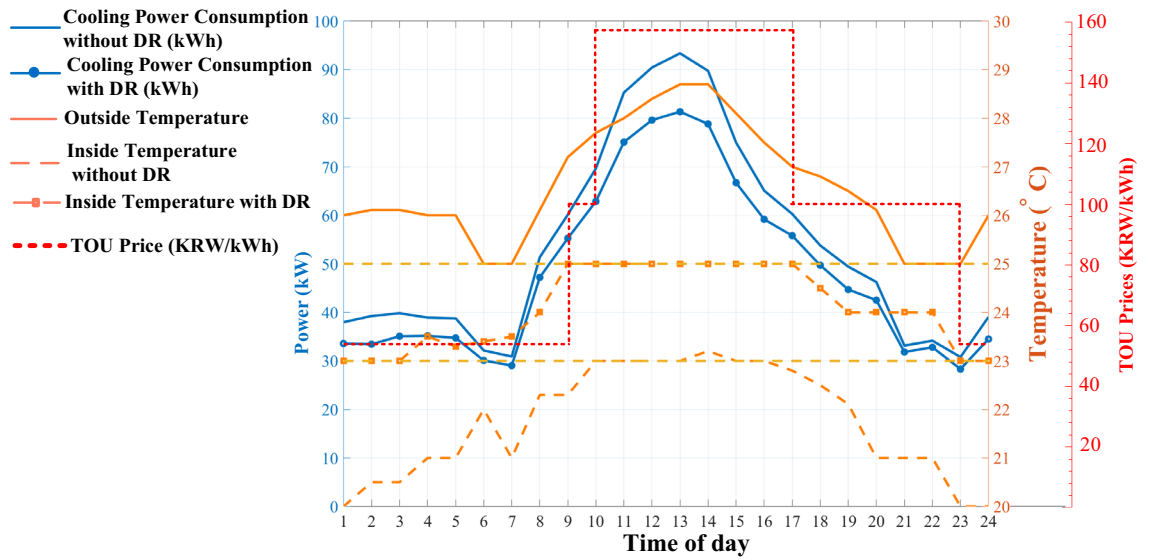**Figure 9.** Washing machine power consumption with and without DR.

Figure 10. Cooling system power consumption with and without DR.
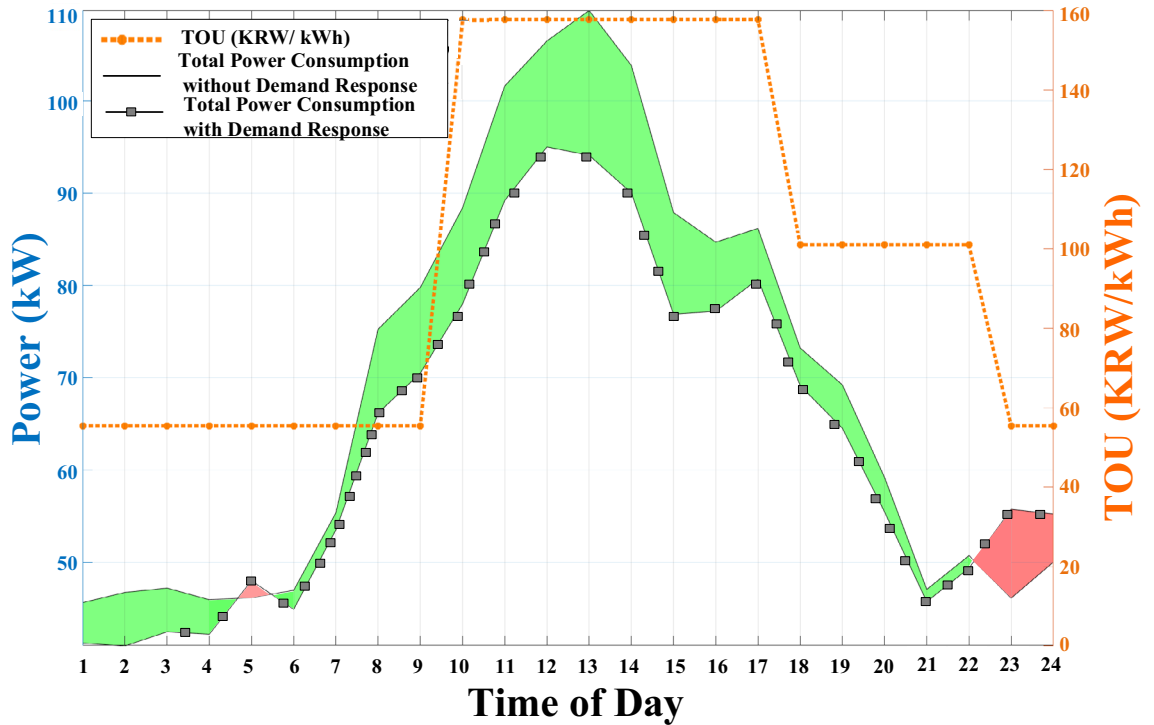
Figure 11. Total residential power consumption with and without DR.

work, we will develop our study with utilizing CPP and EDRP. Furthermore, we will hire LSTM based predictor to estimate RESs output prediction.
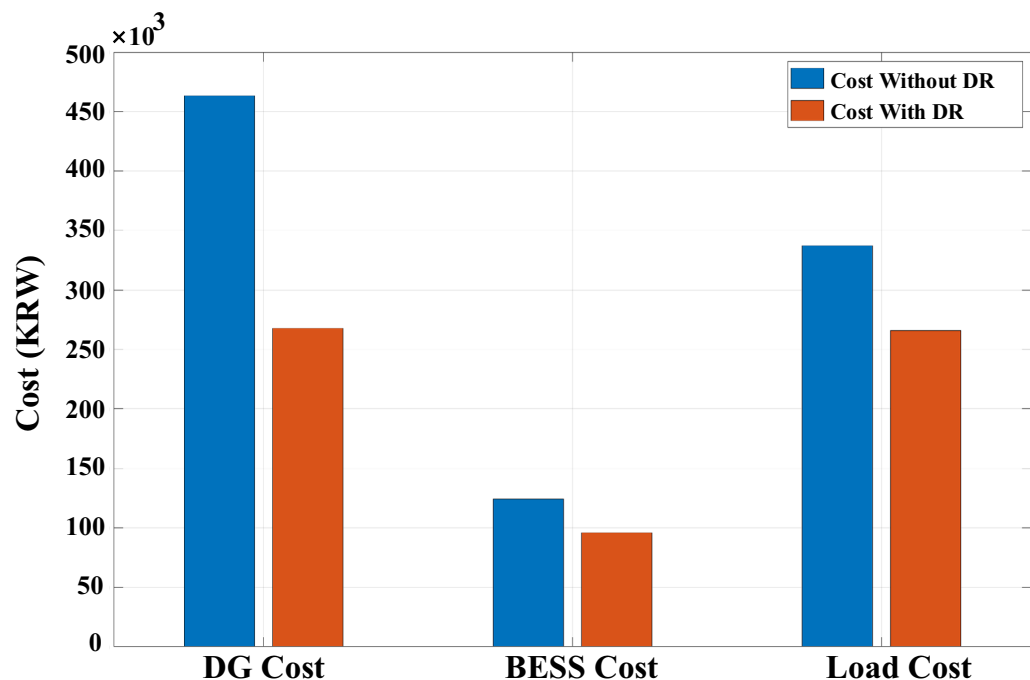
Figure 12. Comparison of DG, BESS, and load consumption cost for understudy day with and without DR.

Optimization solution	Operational cost without DR (KRW)	Operational cost with DR (KRW)
CPLEX solver	832.88×10^3	707.948×10^3
ADMM	1091.441×10^3	709.44×10^3
SAC-NF-ADMM	924.95×10^3	629.84×10^3

Table 3. Comparison of proposed optimization method cost performance (KRW) with centralized-ADMM and CPLEX solver.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Huh, T. Analyzing the configuration of knowledge transfer of the green island projects in S. Korea. *Korean J. Policy Stud.* **32**(1), 1–26 (2017).
2. Tightiz, L., Yang, H. & Piran, M. A survey on enhanced smart micro-grid management system with modern wireless technology contribution. *Energies* **13**(9), 2258 (2020).
3. Ganjei, N., Zishan, F., Alayi, R., Samadi, H., Jahangiri, M., Kumar, R. & Mohammadian, A. Designing and sensitivity analysis of an off-grid hybrid wind-solar power plant with diesel generator and battery backup for the rural area in Iran. *J. Eng.* **2022** (2022).
4. Kang, T. South Korea's experience with smart infrastructure services: Smart grids. Inter-American Development Bank 1–36 (2020).
5. Zhang, X., Wei, Q. S. & Oh, B. S. Cost analysis of off-grid renewable hybrid power generation system on Ui Island, South Korea. *Int. J. Hydrog. Energy* **47**(27), 13199–13212 (2022).
6. Hwang, W. Microgrids for electricity generation in the Republic of Korea. NAPSNet Special Reports (2020).
7. Rangu, S. K., Lolla, P. R., Dhenuvakonda, K. R. & Singh, A. R. Recent trends in power management strategies for optimal operation of distributed energy resources in microgrids: A comprehensive review. *Int. J. Energy Res.* **44**(13), 9889–9911 (2020).
8. Ullah, Z. *et al.* Advanced energy management strategy for microgrid using real-time monitoring interface. *J. Energy Storage* **52**, 104814 (2022).
9. Alarcón, M. A., Alarcón, R. G., González, A. H. & Ferramosca, A. Economic model predictive control for energy management of a microgrid connected to the main electrical grid. *J. Process Control* **117**, 40–51 (2022).
10. Raghav, L. P., Kumar, R. S., Raju, D. K. & Singh, A. R. Analytic hierarchy process (AHP)-swarm intelligence based flexible demand response management of grid-connected microgrid. *Appl. Energy* **306**, 118058 (2022).
11. Raghav, L. P., Kumar, R. S., Raju, D. K. & Singh, A. R. Optimal energy management of microgrids using quantum teaching learning based algorithm. *IEEE Trans. Smart Grid* **12**(6), 4834–4842 (2021).
12. Hu, M., Xiao, F. & Wang, S. Neighborhood-level coordination and negotiation techniques for managing demand-side flexibility in residential microgrids. *Renew. Sustain. Energy Rev.* **135**, 110248 (2021).

13. Chen, G. & Yang, Q. An ADMM-based distributed algorithm for economic dispatch in islanded microgrids. *IEEE Trans. Ind. Inform.* **14**(9), 3892–3903 (2018).
14. Lyu, C., Jia, Y. & Xu, Z. A novel communication-less approach to economic dispatch for microgrids. *IEEE Trans. Smart Grid* **12**(1), 901–904 (2021).
15. Bhattacharjee, V. & Khan, I. A non-linear convex cost model for economic dispatch in microgrids. *Appl. Energy* **222**, 637–648 (2018).
16. Zhou, C. & Zheng, C. *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)* 839–844 (2020).
17. Wu, K. *et al.* Fast distributed Lagrange dual method based on accelerated gradients for economic dispatch of microgrids. *Energy Rep.* **6**, 640–648 (2020).
18. Chang, X. *et al.* Accelerated distributed hybrid stochastic/robust energy management of smart grids. *IEEE Trans. Ind. Inform.* **17**(8), 5335–5347 (2021).
19. Ullah, M. H. & Park, J. Peer-to-peer energy trading in transactive markets considering physical network constraints. *IEEE Trans. Smart Grid* **12**(4), 3390–3403 (2021).
20. Li, Z., Li, P., Yuan, Z., Xia, J. & Tian, D. Optimized utilization of distributed renewable energies for island microgrid clusters considering solar-wind correlation. *Electric Power Syst. Res.* **206**, 107822 (2022).
21. Biagioni, D. *et al.* Learning-accelerated ADMM for distributed dc optimal power flow. *IEEE Control Syst. Lett.* **6**, 1–6 (2020).
22. Ichnowski, J. *et al.* Accelerating quadratic optimization with reinforcement learning. *Adv. Neural Inf. Process. Syst.* **34**, 21043–21055 (2021).
23. Zeng, S., Kody, A., Kim, Y., Kim, K. & Molzahn, D. K. A reinforcement learning approach to parameter selection for distributed optimization in power systems. arXiv preprint arXiv:2110.11991 (2021).
24. Haarnoja, T., Zhou, A., Abbeel, P. & Levine, S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor (2018). In *Paper Presented at the 35th International Conference on Machine Learning, Stockholm, Sweden* (2018).
25. Tightiz, L. & Yang, H. Resilience microgrid as power system integrity protection scheme element with reinforcement learning based management. *IEEE Access* **9**, 83963–83975 (2021).
26. Shin, C. *et al.* The ENERTALK dataset, 15 Hz electricity consumption data from 22 houses in Korea. *Sci. Data* **6**(1), 1–13 (2019).
27. Seo, Y. & Hong, W. Constructing electricity load profile and formulating load pattern for urban apartment in Korea. *Energy Build.* **78**, 222–230 (2014).
28. Lee, D. & Choi, S. Reinforcement learning-based energy management of smart home with rooftop solar photovoltaic system, energy storage system, and home appliances. *Sensors* **19**(18), 3937 (2019).
29. Shi, Y. *et al.* Distributed model predictive control for joint coordination of demand response and optimal power flow with renewables in smart grid. *Appl. Energy* **290**, 116701 (2021).
30. Korea electric power corporation. Electric rates table in South Korea. http://cyber.kepco.co.kr/ckepco/front/jsp/CY/E/E/CYEEHP00201.jsp. Accessed 30 May 2022.
31. Luo, F., Kong, W., Ranzi, G. & Dong, Z. Y. Optimal home energy management system with demand charge tariff and appliance operational dependencies. *IEEE Trans. Smart Grid* **11**(1), 4–14 (2019).
32. Ghasemi, S., Mohammadi, M. & Moshtagh, J. A new look-ahead restoration of critical loads in the distribution networks during blackout with considering load curve of critical loads. *Electric Power Syst. Res.* **191**, 106873 (2021).
33. Mincu, P. R. & Boboc, T. A. *2017 International Conference on Energy and Environment (CIEM)* 480–484 (IEEE, 2017).
34. Mhanna, S., Verbič, G. & Chapman, A. C. Adaptive ADMM for distributed AC optimal power flow. *IEEE Trans. Power Syst.* **34**(3), 2025–2035 (2018).
35. Ward, P. N., Smofsky, A. & Bose, A. J. Improving exploration in soft-actor-critic with normalizing flows policies. arXiv preprint arXiv:1906.02771 (2019).
36. Kobyzev, I., Prince, S. J. & Brubaker, M. A. Normalizing flows: An introduction and review of current methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**(11), 3964–3979 (2020).
37. Mazoure, B., Doan, T., Durand, A., Pineau, J. & Hjelm, R. D. Leveraging exploration in off-policy algorithms via normalizing flows (2020). In *Paper Presented at 3rd Conference on Robot Learning Osaka, Japan* (2020).
38. Open energy system databases renewables.ninja. http://www.renewables.ninja. Accessed 1 June 2022.
39. The weather year round anywhere. https://weatherspark.com/h/d/142033/2017/8/15/. Accessed 1 June 2022.
40. Diamond, S. & Boyd, S. CVXPY: A python-embedded modeling language for convex optimization. *J. Mach. Learn. Res.* **17**(1), 1–5 (2016).
41. Stellato, B., Banjac, G., Goulart, P., Bemporad, A. & Boyd, S. OSQP: An operator splitting solver for quadratic programs. *Math. Program. Comput.* **12**(4), 637–672 (2020).

Acknowledgements

Author contributions

Conception: L.T. Acquisition of data: L.T. Analysis of data: all authors. Writing: L.T. Critical revision: J.Y. Final approval: all authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.