RESEARCH ARTICLE

Electric Distribution Network With Multi-Microgrids Management Using Surrogate-Assisted Deep Reinforcement Learning Optimization

NIPHON KAEWDORNHAN[1], **CHITCHAI SRITHAPON**[2], **(Member, IEEE),**
AND RONGRIT CHATTHAWORN[1]

[1]Department of Electrical Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
[2]School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 11428 Stockholm, Sweden

Corresponding author: Rongrit Chatthaworn (rongch@kku.ac.th)

ABSTRACT Increased electric vehicle demand and uncertain electricity generation from solar photovoltaics lead to fluctuating and inefficient operation in the Distribution Network (DN). MicroGrids (MG) connecting to the DN with suitable energy trading is an effective way to solve this issue. However, proper energy trading considering grid constraints with a low computational burden is a significant challenge for solving the Optimal Energy Management (OEM). Therefore, this paper proposes an OEM using Deep Neural Networks developed as surrogate models to assist the Deep Reinforcement Learning Optimization for reducing the computational burden. The proposed method is deployed to a bi-level OEM for multi-MGs connected in the DN with real-time pricing consideration, represented as the proposed strategy. The power system parameters of the MG estimated by probabilistic power flow are predicted by well-trained surrogate models to mitigate the computational time for finding the optimal solution of Deep Reinforcement Learning. To validate and demonstrate the proposed method and strategy, the modified IEEE-33 bus system and a residential low-voltage distribution system defined as the DN and MG, respectively are utilized. Simulation results show that the proposed method can reduce computational burden by 89.23% compared with the Differential Evolution algorithm. Furthermore, the proposed strategy can provide the optimal purchased energy price offered to each MG. Also, it can decrease the total cost of DN between 0.01% to 0.44% compared with the cost estimated by the strategies using fixed purchased energy prices.

INDEX TERMS Deep neural network, deep reinforcement learning, electric vehicle, microgrid, optimal energy management, photovoltaics, probabilistic power flow.

I. INTRODUCTION

Many countries regard climate change and global warming affected by internal combustion vehicle usage, especially carbon emissions. Hence, Electric Vehicle (EV) usage is increasingly supported by the government because it can reduce carbon emissions in the city. However, massive EV integration in Distribution Networks (DN) and MicroGrid (MG) increase electricity demand exponentially, which can cause negative impacts on DN and MG, such as voltage violation, and reliability problem [1], [2]. In addition, the increased EV demand also impacts the generation planning of electricity utility, which has to generate more power from fossil fuel power plants, causing an increase in carbon emissions. Therefore, it does not guarantee that increased EV usage in DN and MG can reduce carbon emissions.

The associate editor coordinating the review of this manuscript and approving it for publication was Christopher H. T. Lee.

However, Renewable Energy Resources (RERs) are increasingly used for power generation, especially solar Photo-Voltaic (solar PV) and wind turbine, which can reduce power generation from fossil fuel power plants [3], [4].

Solar PV can generate clean energy, and its cost is continuously decreasing. Therefore, the solar PV system is widely used for power generation. In Thailand, the Alternative Energy Development Plan 2018 (AEDP 2018) [5] created by the ministry of energy has set the target for solar PV installed capacity in 2037 as 9,290 MW. However, the uncertainty of solar PV generation harms power quality issues such as voltage and line limit violations [6], [7]. One of the approaches that can mitigate the issues is transforming the conventional grid to a self-sufficient distribution system or MG. Generally, the small power system (low or medium voltage) that has Distributed Generators (DGs), Battery Energy Storage Systems (BESS) with the Energy Management System (EMS) and can sufficiently supply its demand is developed as the MG [8]. However, only alone MG may not alleviate the negative impacts of solar PV generation and increased EV demand. Thus, many MGs will group to become multi-MGs connected to DN for energy exchange to sufficiently supply the demand for both DN and MGs and to solve those negative impacts [9]. This situation can increase energy exchange between DN and MGs, decreasing the receiving energy from the main grid. The BESS controlled by EMS is employed for energy management in the situation. Therefore, the EMS is an important mechanism for finding Optimal Energy Management (OEM) solutions in DN with multi-MGs. Besides, the strategy for EMS design is a major challenge for solving OEM problems.

For the EMS framework, there are four types of architectures of the EMS for DN with multi-MGs: centralized, decentralized, hybrid, and nested structures [9]. Commonly, The Distribution System Operator (DSO) and MG Operators (MGOs) are for-profit entities. The DSO cannot access to control the in-depth operation of MG. MGOs transfer only necessary information to DSO, which can protect their private information. The type is called a hybrid EMS or bi-level EMS [10]. A bi-level EMS is very popular for the EMS design in DN with multi-MGs. Several researchers designed the bi- level EMS as follows: Toutounchi et al. [11] and Haseeb et al. [12] applied a bi-level EMS to the test system to increase power exchange between DN and multi-MGs and to plan day-ahead and real-time frames for scheduling energy storage devices and controllable generators in DN and all MGs. Also, DSO tries to find a fair price for purchased energy from multi-MGs. Khavari et al. [13] proposed the bi-level EMS for energy management in multi-MGs. The first level is modeled for maximizing MG profits under the Point of Common Coupling (PCC) line capacity, which is the connection points between each MG, whereas the second level is modeled to find a fair energy price for all MGs. From the above literature, energy trading is essential for motivating power exchange between DN and MG. Nevertheless, the above literature has not considered the uncertainties of RERs generation and overall demand in the OEM.

Uncertainty factors are an essential part of EMS training to guarantee OEM solutions. There are many uncertainty factors in OEM problems, such as solar PV power fluctuation and EV demand. Solar PV power is varied due to ambient temperature and solar irradiance. In contrast, EV demand is evaluated by the uncertainties of vehicle types, departure time, arrival time, and daily travel distance. There are many methods to deal with those uncertainties. In reference, [14] author applied the budget of uncertainty method based on the forecasted concept to deal with the solar PV power fluctuation. Marzband et al. [15] proposed Taguchi's orthogonal array testing method for considering the uncertainty of solar PV generation, wind turbine power, and demand. Aghdam et al. [16] and Nagpal et al. [17] adapted the chance-constrained programming approach to adjust all device constraints, which can mitigate the negative impacts of RERs and demand uncertainties. Moreover, EMS operating under operation constraints of power system using power flow calculation is essential and can guarantee the OEM solution. However, the power system constraints were not taken into account by the above works of literature.

The Deterministic Power Flow (DPF) is an efficient way to evaluate AC power parameters. For example, Liu et al. [18] and Du et al. [19] applied the path following the interior point algorithm and the linearized optimal power flow to calculate the DPF in the DN, respectively. Nevertheless, the DPF may estimate the parameters with low accuracy when the system has high uncertain parameters. A Probabilistic Power Flow (PPF) is taken into account in several EMS research works to solve this problem. The PPF is an iterative method relying on a large DPF result to create probability distributions of power system parameters. Scenario generation for DPF calculation is an essential part of the PPF. There are many methods for scenario generation. For example, in references [20], [21], and [22], authors applied the PPF using Monte Carlo Simulation (MCS) method to create the scenarios. Moreover, Srithapon et al. [23] presented Zhao's Point Estimation Method combined with Nataf transformation (PEMN). Numerical results showed that the PEMN can provide the parameters with a high confidence level which is used to guarantee OEM solutions. Also, the MCS has to generate more scenarios than the PEMN to improve PPF accuracy. Thus, the MCS consumes a high computational burden. However, the DPF result of these methods is estimated by an iterative approach such as Newton Raphson approach. If the approach is still employed to calculate the DPF, the computational burden may not be able to decrease.

To deal with the above problem, Machine Learning (ML) is adopted as a surrogate model to evaluate the power parameters instead of the iterative approach. In references, [24] and [25] authors developed a radial basis function and deep belief network, respectively, as a surrogate model to predict the power parameters. Srithapon et al. [23] proposed Deep Neural Networks (DNNs) as the surrogate model to deal with the problem that can decrease the OEM calculation time by 88.76%. However, the above research works only adapted the

surrogate model to predict the parameters of a single power system, which has not been developed for various power systems such as multi-MGs connected in DN.

There are many numerous optimization approaches for solving the OEM problem in DN with multi-MGs. Commonly, a rule-based control approach is applied to manage electrical energy within the DN or MG. For example, Bogaraj et al. [26] applied the approach to schedule energy storage devices and controllable generators, which can provide the optimal action with low computational time. Although the approach can reduce the computational burden because it controls the devices using the if-else concept, it may be trapped to the local optimum. Authors in reference [27] have proposed a Branch and Bound algorithm to minimize energy losses, and lifetime degradation of energy storage devices. In references, [28], [29], [30], and [31] authors presented a new combined control algorithm based on the mixed integer linear programming concept to solve the OEM problems. Moreover, metaheuristic algorithms were employed in the OEM task. The algorithms can solve complex nonlinear problems with a high dimension of inputs and outputs [32]. Lv et al. [33] solved the OEM problems using a hybrid approach based on a genetic algorithm to achieve sustainable DN and multi-MGs interaction. Nikmehr et al. [34] employed the Particle Swarm Optimization (PSO) to deal with the uncertainties of RERs and loads, and to solve the scheduling problem. Tan et al. [35] utilized a bi-layer solution approach, which consists of adaptive multi-objective evolutionary and Differential Evolution (DE) algorithms to solve the bi-layer OEM both DN and multi-MGs. Although many metaheuristic algorithms have the opportunity to provide the global optimum more than others, they consume a high calculational time to solve the complex OEM problem [36].

Recently, many researchers applied Reinforcement Learning (RL) to solve OEM problems. The RL provides efficient solutions for complex problems with low time consumption. The RL has a training process that allows agents to interact with the environment to find global decision control. Foruzan et al. [37] employed the RL to develop the OEM strategy without prior information. Du et al. [38] developed the DNN as the environment to predict the behavior of DN with multi-MGs, which can protect user privacy. The RL was applied to minimize the demand-side Peak-to-Average Ratio (PAR) and increase the DSO profit from selling energy. In addition, Guo et al. [39] employed a Deep RL (DRL) to manage the power exchange between the DN and multi-MGs. Also, the DNN was modeled as an agent which can interact together, learn overall solutions, and remember the optimal solution. Numerical results showed that the DRL can minimize the overall cost of DSO and MGOs less than the cost obtained by metaheuristic algorithms. However, the environment needs to have excellent and quick responses in interacting and reducing the computational burden. Furthermore, the above research works do not consider power system constraints and RERs uncertainties in the optimization process. Thus, the optimal solution may violate the DN and MGs constraints.

To deal with many limitations in the previous research works, this paper proposes a multi-agent bi-level OEM with Real-Time Pricing (RTP) framework using the DRL. The proposed framework is designed following a hybrid EMS assuming that the DSO and MGO are for-profit entities. Deep Deterministic Policy Gradient (DDPG), one of the DRL types, is employed to solve the OEM problem of a single DN and three MGs. Solar PV generation and EV usage are considered in each MG. To improve the computational burden, MG behaviors are developed as a DNN surrogate model. In addition, DSO and MGO behaviors are modeled as the DNN agents and trained simultaneously. The first level aims to minimize each MGO cost through BESS dispatching, whereas the second level aims to reduce the DSO cost by proposing the purchased energy price to each MGO. The main contributions of this paper are summarized as follows:

1) This paper introduces the architecture of the EMS framework with the PPF for solving the OEM problem of a single DN with three MGs considering the impacts of EV demand and uncertain solar PV generation, which have not been studied in previous works.
2) The DNN surrogate model is developed for estimating the power system parameters of each MG instead of using PPF calculation and utilized to assist the DDPG algorithm in discovering the optimal solution. This method can reduce the computational burden of solving the OEM problem. Also, the performance of the proposed method is compared with one of the best tools of metaheuristic algorithms named Differential Evolution which was employed to solve many problems in the past [40], [42], [43].
3) The novel bi-level OEM with real-time pricing using DNN surrogate-assisted multi-agent DDPG algorithm is proposed to improve the mechanism for handling the OEM problem of a DN with three MGs under the DN and MG operational constraints.

This paper is organized as follows: Section II presents the description modeled of EMS framework which is the hybrid EMS. Section III presents the probabilistic power flow used to evaluate the power system parameters. The problem formulation for this work is represented in Section IV whereas Section V shows the proposed methodology. The simulation results and discussion are presented in Sections VI and VII, respectively. Finally, the conclusion is presented in Section VIII.

II. EMS FRAMEWORK DESCRIPTION

In this section, the introduction architectures of the EMS framework for DN with the multi-MGs are described in the first subsection. Then, the hybrid EMS framework applied to this work is presented in the second subsection.
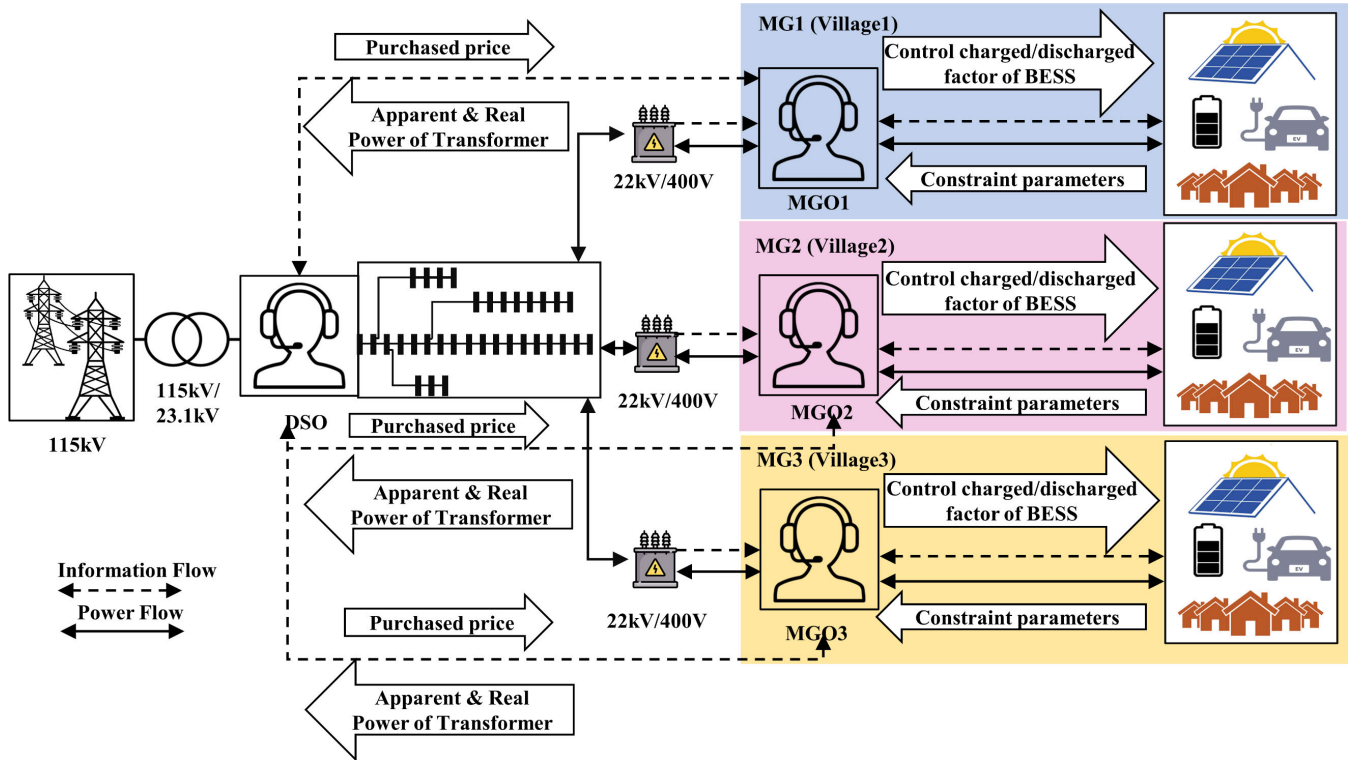
FIGURE 1. The proposed EMS framework for single-DN with three-MGs.

A. ARCHITECTURES OF THE EMS FRAMEWORK

The operation of muti-MGs in networked MGs is an efficient way to improve the reliability of the DN [44]. For the popular EMS frameworks, there are three types of the frameworks for DN with multi-MGs namely centralized, decentralized, and hybrid frameworks [9]. A centralized EMS is commonly employed for controllable networked MGs. The DSO can control the operation of storage energy devices and controllable generators of multi-MGs independently. However, the centralized EMS cannot be applied for uncontrollable networked MGs such as the DSO and multi-MGOs which are for-profit entities. If two and more MGs connected in the DN are different entities, a decentralized EMS is utilized for this strategy. Multi-MGOs can independently manage their demand and supply to obtain a high profit. Also, they can generate energy trading in the DN. Nevertheless, poor stability of the DN may occur due to the energy trading without considering DN operation constraints. To deal with the problem, the hybrid EMS is employed to optimize energy trading under the DN operation constraints. Therefore, the hybrid EMS is applied to manage the energy and create optimal energy trading for DN with multi-MGs. The detail is presented in the following subsection.

B. A HYBRID EMS FRAMEWORK

In this work, a hybrid EMS is applied to define the OEM problem with the RTP. The proposed EMS framework is shown in Fig. 1. Distribution Network (DN), supervised by

a DSO, has three MGs. A bi-level EMS is developed using the hybrid EMS framework. Each MG is controlled by the MGO under solar PV and EV uncertainties. The MGO manages electric energy within the MG through BESS control. Moreover, the MGO must coordinate with the DSO to prevent violated situations that may occur within DN. For the role of DSO, it controls the power flow within the DN and exchanges power with MGOs in real-time. The DSO and multi-MGOs can send/receive general information through the interaction between DSO and MGO under the EMS framework to analyze and provide the optimal decision. Even though the energy exchange of the MG cannot be controlled directly by the DSO, it can be motivated by proposed energy prices using the RTP, which can increase the attention of MGOs to exchange their energy.

III. PROBABILISTIC POWER FLOW

The DN and MG operation constraints are taken into account in the EMS training to guarantee the OEM solution without any constraint violations. A Deterministic Power Flow (DPF) is an essential step of a Probabilistic Power Flow (PPF) because the PPF is an iteration method relying on a large DPF result to create the probability distribution of desired power parameters. Since solar PV generation and EV usage are integrated into each MG, which have high uncertainties, thus the PPF is utilized to evaluate the power parameters of MG. In contrast, the DPF is employed to evaluate the power parameters of DN. This section explains the stochastic model

TABLE 1. The expected EV charging probability following the time-of-use rate [48], [49].

Interval (h)	TOU rate	Probability
$1 \leq t < 2$	Off-peak	15%
$2 \leq t < 9$	Off-peak	1%
$9 \leq t < 19$	peak	1%
$19 \leq t < 22$	peak	6%
$22 \leq t \leq 24$	Off-peak	15%

of the uncertainty variables for the PPF defined in the first subsection, whereas the PPF procedure is described in the second subsection.

A. STOCHASTIC MODELS

Uncertainty variables are an essential input to generate the scenarios. Usually, the distribution of each uncertainty variable is modeled as a Probability Density Function (PDF). In this research work, overall demands such as home baseload and EV demand are denoted as uncertainty variables. Also, power generation of solar PV integrated into each MG is uncertain. Firstly, the stochastic models of home baseload and solar PV generation are presented. Then, the stochastic model of EV demand is defined in this subsection.

1) HOME BASELOAD AND SOLAR PV GENERATION

Since the residential low-voltage distribution network is modeled as the MG. Thus, the home baseload within the MG is considered in this work. The baseload behavior is commonly modeled as a normal PDF [45] which can be written as the following equation:

$$f(x : \mu_n, \sigma_n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu_n)^2}{2\sigma_n^2}\right) \quad (1)$$

where $f(x : \mu_n, \sigma_n)$ is the normal PDF of the uncertainty variable x whereas μ_n and σ_n are the mean and standard deviation of the PDF, respectively.

An ambient temperature and solar irradiance are uncertainty variables for evaluating the output power of solar PV. The ambient temperature is modeled as a normal PDF similar to the home baseload whereas the solar irradiance is modeled as a beta PDF [45], [46]. The beta PDF can be formulated by the following equation:

$$f(r : \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) + \Gamma(\beta)} r^{\alpha-1} (1 - r)\beta \quad (2)$$

where $f(r : \alpha, \beta)$ is the beta PDF of the uncertainty variable r which is the solar irradiance (W/m^2). Γ is the gamma function. Then, α and β are the exponents of random variable and control deviation, respectively.

The solar PV output power has an inherent uncertainty that depends on the ambient temperature and solar irradiance [23],

which is expressed as follows:

$$R_t = \begin{cases} r_t^2 / (r_c r_{std}) & r_t < r_c \\ r_t / r_{std} & r_c \leq r_t \leq r_{std} \\ 1 & r_t > r_{std} \end{cases} \quad (3)$$

$$T_{C,t} = T_{a,t} + R_t (T_n - 20) \quad (4)$$

$$P_{PV,t} = \eta_g P_{PV,r} R_t \left[1 + \alpha_P \left(T_{C,t} - T_{C,STC}\right)\right] \quad (5)$$

where $P_{PV,t}$ is the solar PV output power (kW) at time t. η_g and $P_{PV,r}$ are the overall efficiency of the solar PV generation system and the rated solar PV output power (kW), respectively. α_P is the coefficient of the output power due to the temperature (kW/°C) whereas $T_{C,t}$ and $T_{C,STC}$ are the temperature of the PV cell (°C) at time t and Standard Test Conditions (STC) temperature (°C) which is set as 25 °C, respectively. $T_{a,t}$ is the ambient temperature (°C) at time t. T_n denotes the nominal operating cell temperature (°C) which is usually set as 41 °C. r_t and r_{std} denote the solar irradiance (W/m^2) at time t and the STC solar irradiance which is commonly assigned as 1,000 W/m^2, respectively. Then, r_c is a certain radiation point which is usually defined as 150 W/m^2.

2) EV DEMAND MODEL

The EV charging demand is considered in the residential networks and evaluated based on many uncertainty variables. The stochastic model of EV demand is constructed based on the type of vehicle, departure time, arrival time, and daily travel distance [18]. Based on the data from the National Household Travel Survey (NHTS) [47], the uncertainties of the arrival and departure times are modeled as normal PDFs. In contrast, the daily travel distance is modeled as a lognormal PDF that can be written by the following equation:

$$f(d : \mu_l, \sigma_l) = \frac{1}{d\sigma_l\sqrt{2\pi}} \exp\left(-\frac{(\ln d - \mu_l)^2}{2\sigma_l^2}\right) \quad (6)$$

where $f(d : \mu_l, \sigma_l)$ is the lognormal PDF of the uncertainty variable d that is the daily travel distance (km) whereas μ_l and σ_l are the mean and standard deviation of the PDF, respectively.

However, predictive EV demand is a massive challenge for the OEM task because EV demand is high and uncertain. Hence, governments of many countries apply Time-Of-Use (TOU) rate to decrease the uncertainty of EV charging demand. The TOU can create several EV charging percentages [48], [49] shown in Table 1. The stochastic model of EV demand can be generated in the following steps:

Step 1: From Table 1, the number of EVs ($n_{EV,t}$) that start charging at time t for 24 hours is defined below:

$$n_{EV,t} = P_t \cdot N_{EV} \quad t = 0, 1, 2, \ldots, 23 \quad (7)$$

where P_t and N_{EV} are the expected EVs probability at time t and the total number of EVs in the test system, respectively.

Step 2: Random EV types to label the ith EV for 24 hours.

Step 3: Random the daily travel distance to evaluate the required charging duration ($T_{CHd,i}$) of the ith EV for 24 hours which can be calculated by the following equation [23]:

$$T_{CHd,i} = \frac{\varepsilon_i d_i}{\eta_{ev,i} P_{r,i}} \qquad (8)$$

where ε_i and d_i are the consumption rate (kWh/km) and daily travel distance (km) of the ith EV, respectively. $\eta_{ev,i}$ denotes the charging efficiency whereas $P_{r,i}$ is the charger power rating (kW) of the ith EV.

Step 4: In this work, EVs are simultaneously charged at the rated power. From step 3, $T_{CHd,i}$ of each EV at time t may require more than an hour. For 24 hours, the number of EVs ($N_{EV,t}$) charged at time t can be evaluated as follows:

$$N_{EV,t} = n_{EV,t} + nr_{EV,t-1} + nr_{EV,t-2} + \ldots \qquad (9)$$

where $n_{EV,t}$ denote the number of EVs from step 1 whereas $nr_{EV,t-1}$ is the number of EVs that is not fully charged at time $t-1$.

Step 5: The Monte Carlo Simulation (MCS) is applied to generate EV charging scenarios by repeating 10,000 times for steps $1-4$. Therefore, the stochastic model of EV demand is presented through probability values which is a ratio between $N_{EV,t}$ and N_{EV}.

B. PPF PROCEDURE
The PPF calculation is an efficient way to estimate the power system parameters of the power system with high uncertainty. In this subsection, the Nataf Transformation (NT) combined with the Point Estimation Method (PEM) is employed to generate DPF scenarios and evaluate the desired power parameters. Thus, DPF scenario generation is described in the first part whereas the PPF evaluation is represented in the last part.

1) DPF SCENARIO GENERATION
Random variables are necessary inputs for scenario generation. The dependence on those variables is essential for the sampling process. Joint Probability Distribution (JPD) of all variables is one of the models to illustrate their dependence. Nevertheless, the JPD is generated using complete statistical information which is not easy. Marginal Probability Distribution (MPD) is an efficient way to construct the JPD of those variables with limited information. The NT is one of the methods to construct the JPD of random variables using the MPD concept [50], whereas the sampling points are determined by Zhao's PEM [51]. Hence, the above method is applied to create the DPF scenarios in this work, shown as follows:

Step 1: Evaluate the Correlation Coefficients (CC) in the Original Distribution Space (ODS) of the n input random variables using Spearman's Rank-order Correlation Method [52]. The CC can be show as the following equation:

$$R_X = \begin{bmatrix} \rho_{X,11} & \rho_{X,12} & \cdots & \rho_{X,1n} \\ \rho_{X,21} & \rho_{X,22} & \cdots & \rho_{X,2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{X,n1} & \rho_{X,n2} & \cdots & \rho_{X,nn} \end{bmatrix} \qquad (10)$$

where R_X denotes the CC in the ODS whereas $\rho_{X,ij}$ is the correlation coefficient of the ODS input variables X_i and X_j, respectively.

Step 2: Transfer R_X in the ODS to the Correlated Normal Space (CNS). More explanation of the transfer method can be found in [23]. The tranfer method satisfies the following equation:

$$R_Z = \begin{bmatrix} \rho_{Z,11} & \rho_{Z,12} & \cdots & \rho_{Z,1n} \\ \rho_{Z,21} & \rho_{Z,22} & \cdots & \rho_{Z,2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{Z,n1} & \rho_{Z,n2} & \cdots & \rho_{Z,nn} \end{bmatrix} \qquad (11)$$

where R_Z represents the CC in the CNS. $\rho_{Z,ij}$ is the correlation coefficient of CNS input variables Z_i and Z_j.

Step 3: Evaluate the Lower triangular matrix (L) of the. using the Cholesky Decomposition as follows:

$$R_Z = LL^T \qquad (12)$$

Step 4: Given m points of each input variable are sampled in the Standard Gaussian Space (SGS) using the Gaussian Points to estimate the mean and standard deviation of desired output parameters. Thus, the points can be shown as follows:

$$U_{m \times n} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m2} & \cdots & u_{mn} \end{bmatrix} \qquad (13)$$

where $U_{m \times n}$ is the SGS sample matrix which has $m \times n$ dimensions, whereas u_{ki} is the SGS sample at the kth point of the ith variable.

Step 5: Transfer $U_{m \times n}$ in the SGS to the CNS using the Lower triangular matrix (L) from step 3 which can be shown in the below equation:

$$Z_{m \times n} = LU = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{bmatrix} \qquad (14)$$

where $Z_{m \times n}$ is the CNS sample matrix which has $m \times n$ dimensions whereas z_{ki} is the CNS sample at the kth point of the ith variable.

Step 6: Transfer $Z_{m \times n}$ in the CNS to the ODS which can be represented in the below equations [23]:

$$X_i = F^{-1}[f(Z_i)] = [x_{1i}, x_{2i}, x_{3i}, \ldots, x_{mi}]^T \qquad (15)$$

$$X_{m \times n} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \qquad (16)$$

where X_i is the ODS sample vector of the ith variable mapped from Z_i by using the Cumulative Distribution Function (CDF) $f(Z_i)$ and the inverse CDF F^{-1} of Z_i. $X_{m \times n}$ denotes the ODS sample matrix which has $m \times n$ dimensions whereas x_{ki} is the ODS sample at the kth point of the ith variable.

Step 7: Generate the DPF scenarios based on the principle of Zhao's PEM which can be shown below:

$$S = \begin{bmatrix} S_1 \\ S_o \\ \vdots \\ S_{(m-1)n+1} \end{bmatrix} = \begin{bmatrix} x_{k1} & \mu_{x1} & \cdots & \mu_{xn} \\ \mu_{x1} & x_{ki} & \cdots & \mu_{xn} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{x1} & \mu_{x2} & \cdots & x_{kn} \end{bmatrix} \quad (17)$$

where

$$o = 1, 2, 3, \ldots, (m-1)n+1$$
$$k = 1, 2, 3, \ldots, m$$
$$i = 1, 2, 3, \ldots, n \quad (18)$$

where S is the DPF scenario matrix. S_o denotes the oth scenario. x_{ki} and μ_i are the kth sampling point of ith input variable and the mean of ith input variable, respectively. From the NT with Zhao's PEM concept, the number of scenarios S equal to $(m-1)n+1$ [53].

2) PPF EVALUATION

From the previous part, the S is generated by the NT with Zhao's PEM. Then, the evaluation of desired power parameters is represented in this part.

G is defined as a structural response function of desired power parameters in the system such as bus voltage, line current, and real/reactive power of the transformer. Therefore, G is represented as a function of random input variables as follows:

$$Y = G\left(\left[S_1, S_2, \ldots, S_{(m-1)n+1}\right]^T\right) \quad (19)$$

where Y is the output vector from $(m-1)n+1$ scenarios elevated in the G.

μ_y and σ_y are defined the mean and standard deviation values of the output variable, respectively. These parameters can be evaluated using the following equations [53]:

$$\mu_y = \int G(S)f(S)\, dS \quad (20)$$

$$\sigma_y = \sqrt{\int \left(G(S) - \mu_y\right)^2 f(S)\, dS} \quad (21)$$

where $f(S)$ is the JPD function of random input variables. However, if μ_y and σ_y are calculated by using (20) and (21), respectively, it leads to the computational burden [53]. Since S is created by the NT with Zhao's PEM, it can solve the problem, decrease the number of scenarios, and provide a high accuracy comparable with (20) and (21). Thus, the μ_y and σ_y are evaluated as follows [23]:

$$\mu_y = G(S_\mu) + \sum_{i=1}^{n} \left[\mu_i - G(S_\mu)\right] \quad (22)$$

$$\sigma_y = \sqrt{\sum_{i=1}^{n} \sigma_i^2} \quad (23)$$

$$\mu_i = \sum_{k=1}^{m} w_k G(S_{ki}) \quad (24)$$

$$\sigma_i^2 = \sum_{k=1}^{m} w_k \left[G(S_{ki}) - \mu_i\right]^2 \quad (25)$$

where $G(S_\mu)$ is the DPF result when all inputs are set to the mean value S_μ. μ_i and σ_i denote the mean and standard deviation values of all DPF results from scenarios created using sampling points of ith variable, respectively. The S_{ki} is the scenario created from the kth sampling point associated with i^{th} variable whereas other variables are set as the mean value. $G(S_{ki})$ represents the DPF result of the scenario S_{ki} in k^{th} scenario. Then, w_k is the Gaussian coefficient weight.

IV. PROBLEM FORMULATION

In this work, the bi-level Optimal Energy Management (OEM) for a single DN with three MGs is proposed to be adopted into the EMS. The bi-level OEM based on the hybrid EMS framework is developed to minimize the overall cost of DSO by offering the purchased energy price to the MGOs. In contrast, the MGOs try to control the BESS to reduce its overall cost at the same time. Moreover, AC power parameters using both PPF and DPF calculation are subjected in the OEM task. Therefore, this section presents objective functions and constraints of both DSO and MGO. Also, the bi-level OEM strategy is formulated in this section.

A. OBJECTIVE FUNCTIONS

Objective functions of DSO and MGO significantly impact determining the direction of optimal solutions. This work has three objective functions: exchanged energy cost, carbon emission cost, and BESS degradation cost.

1) EXCHANGED ENERGY COST

The MGO tries to push an energy exchange between MG and DN to minimize the net exchanged energy cost of the MGO, whereas the DSO tries to reduce the purchased energy cost from the main grid and multi-MGs. The MGO responds to the purchased energy price proposed by the DSO to increase the opportunity for its energy selling. In contrast, they manage energy to reduce the received energy cost following the Time-Of-Use (TOU) rate from the DSO. The TOU of Thailand [54], [55] is applied in this work. The exchanged energy cost of DSO and MGO can be formulated as the following equations [39]:

$$F_{1,t} = C_{maingrid,t}^{TOU} \cdot P_t^{DN} + \sum_{i=1}^{3} C_{DSO,t}^{pursh} \cdot P_{i,t}^{MGs} \quad (26)$$

$$f_{1,i,t} = C_{DSO,t}^{TOU} \cdot P_{i,t}^{MGb} - C_{DSO,t}^{pursh} \cdot P_{i,t}^{MGs} \quad (27)$$

where $F_{1,t}$ is the purchased energy cost ($) from the main grid and MGs of the DSO at time t. $C_{maingrid,t}^{TOU}$ and P_t^{DN} denote the TOU rate ($/kWh) of the main grid and the received

power (kW) from the main grid at time t, respectively. $f_{1,i,t}$ is the net exchanged energy cost ($) of ith MGO at time t. $C_{DSO,t}^{TOU}$ and $C_{DSO,t}^{pursh}$ are the TOU rate ($/kWh) of DN and purchased energy price ($/kWh) proposed by DSO at time t, respectively. Then, $P_{i,t}^{MGb}$ is the power (kW) that MGO buys from DSO whereas $P_{i,t}^{MGs}$ represents the power (kW) that MGO sells to DSO.

2) CARBON EMISSION COST

Carbon emission cost is defined as the cost of capturing carbon dioxide due to received power from the external grid. To reduce carbon emissions, the MGO should reduce received power from the DN whereas the DSO should decrease received power from the main grid. In this work, the carbon emission rate (kg/kWh) and capturing carbon dioxide rate ($/kg) are evaluated as the emission rates according to [56], [57]. The carbon emission cost is formulated as a function of the power which can be defined as follows [23]:

$$F_{2,t} = W_{carbon,t} \cdot C_{capture} \cdot P_t^{DN} \tag{28}$$

$$f_{2,i,t} = W_{carbon,t} \cdot C_{capture} \cdot P_{i,t}^{MGb} \tag{29}$$

where $F_{2,t}$ and $f_{2,i,t}$ denote the carbon emission cost ($) of DSO and i^{th} MGO at time t, respectively. $W_{carbon,t}$ and $C_{capture}$ are the carbon emission rate (kg/kWh) and capturing carbon dioxide rate ($/kg), respectively.

3) BESS DEGRADATION COST

The BESS is an essential element with a high investment cost in the MG. They may have an increased end-of-life rate with excessive use. A degradation cost can be modeled to convey a level of end-of-life rate. The cost can be determined according to [58]:

$$f_{3,i,t} = \begin{cases} C_E(SoC_t) - C_E(SoC_{t-1}) & P_{B,t} > 0 \\ 0 & P_{B,t} \le 0 \end{cases} \tag{30}$$

$$C_E(SoC_t) = \frac{C_{cap}}{N_{cycle,t}} \tag{31}$$

$$N_{cycle,t} = \beta_0 \times (1 - SoC_t)^{-\beta_1} \times e^{\beta_2 \times SoC_t} \tag{32}$$

where $f_{3,i,t}$ is the BESS degradation cost ($) at time t of the ith MG. $P_{B,t}$ is the BESS power (kW) at time t. If $P_{B,t}$ is more than 0, BESS is discharging. Otherwise, BESS is charging. SoC_t represents the State-of-Charge (SoC) of BESS at time t. $C_E(SoC_t)$ denotes the one-cycle degradation cost ($) at state-of-charge SoC_t, whereas the BESS capital cost ($) is denoted as C_{cap}. $N_{cycle,t}$ is the number of cycle life at state-of-charge SoC_t. Then, β_0, β_1 and β_2 are curve-fitting coefficients.

B. CONSTRAINTS

To protect against the failure of the power system operation, the DN and MG constraints are considered to guarantee optimal solutions. The power system operation, transformer, and BESS constraints are presented in this subsection.

1) POWER SYSTEM OPERATION CONSTRAINTS

There are two types of constraints for power flow calculation composing of equality and inequality constraints. In the equality constraints, real and reactive power balance equations in the power system can be determined as follows [23]:

$$P_{G,i} - P_{D,i} = \sum_{j=1}^{Nbus} V_i V_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \tag{33}$$

$$Q_{G,i} - Q_{D,i} = \sum_{j=1}^{Nbus} V_i V_j \left(G_{ij} \sin \theta_{ij} + B_{ij} \cos \theta_{ij} \right) \tag{34}$$

where $P_{G,i}$ and $Q_{G,i}$ denote the real and reactive power (p.u.) generation whereas $P_{D,i}$ and $Q_{D,i}$ are the real and reactive power (p.u.) demand at the ith bus, respectively. V_i and V_j are the bus voltage (p.u.) at the ith and j^{th} buses, respectively. G_{ij} and B_{ij} denote the conductance and susceptance (p.u.) of the line between the ith and j^{th} buses, respectively. θ_{ij} denotes as the difference of the angle between V_i and V_j. Then, $Nbus$ is the total number of buses within the system.

For the inequality constraints, it is determined as the network operation constraints as follows:

$$V_{min} \le V_{i,t} \le V_{max} \tag{35}$$

$$\left| I_{l,t} \right| \le I_{max} \tag{36}$$

where $V_{i,t}$ is the bus voltage (V) of the ith bus at time t. V_{min} and V_{max} are minimum and maximum bus voltage (V), respectively. $I_{l,t}$ denotes the line current (kA) of the lth line at time t whereas maximum line current (kA) is represented as I_{max}.

2) TRANSFORMER CONSTRAINTS

Overloading transformers may lead to a temperature rise. The situation may cause the failure of insulators which can reduce the transformer aging. Hottest-Spot Temperature (HST) is essential factor for predictive aging of the transformer. Transformer loading and ambient temperature are the main factors to evaluate the HST. Then, the HST is used to estimate the aging acceleration factor (F_t^{AA}) of the transformer. In this work, the HST and transformer loading in kVA unit are subjected to the transformer constraints. Moreover, the F_t^{AA} also is evaluated to show the performance of the proposed method. If the F_t^{AA} is more than 1, transformer loss of life will increase exponentially [59]. The above constraints can be shown as follows:

$$\theta_{x,t}^{HS} \le \theta_{max}^{HS} \tag{37}$$

$$S_{x,t} \le S_{x,rated} \tag{38}$$

$$F_{x,t}^{AA} \le 1 \tag{39}$$

where $\theta_{x,t}^{HS}$ and θ_{max}^{HS} are the HST at time t (°C) and maximum HST (°C) of the xth transformer, respectively. $S_{x,t}$ and $S_{x,rated}$ denote the loading at time t and rated loading (kVA) of the xth transformer, respectively. Then, $F_{x,t}^{AA}$ is the aging acceleration factor (h) of the xth transformer at time t.

F_t^{AA} and θ_t^{HS} of each transformer can be formulated according to [59] as follows:

$$F_t^{AA} = \exp\left(\frac{15000}{110 + 273} - \frac{15000}{\theta_t^{HT} + 273}\right) \quad (40)$$

$$\theta_t^{HS} = \theta_t^{A} + \Delta\theta_t^{HT} + \Delta\theta_t^{TA} \quad (41)$$

where θ_t^{A} is the ambient temperature (°C). $\Delta\theta_t^{HT}$ and $\Delta\theta_t^{TA}$ are winding hottest-spot temperature rise over the top-oil temperature (°C), and the top-oil temperature rise over the ambient temperature (°C), respectively. Moreover, $\Delta\theta_t^{HT}$ and $\Delta\theta_t^{TA}$ can be estimated as follows:

$$\Delta\theta_t^{HT} = \left(\Delta\theta_t^{HT,f} - \Delta\theta_{t-1}^{HT}\right)(1 - \exp(-\Delta t/\tau_w)) + \Delta\theta_{t-1}^{HT} \quad (42)$$

$$\Delta\theta_t^{TA} = \left(\Delta\theta_t^{TA,f} - \Delta\theta_{t-1}^{TA}\right)(1 - \exp(-\Delta t/\tau_o)) + \Delta\theta_{t-1}^{TA} \quad (43)$$

where $\Delta\theta_t^{HT,f}$ and $\Delta\theta_{t-1}^{HT}$ are the final and initial winding hottest-spot temperature rise over the top-oil temperature at time t and $t - 1$, respectively. $\Delta\theta_t^{TA,f}$ and $\Delta\theta_{t-1}^{TA}$ denote the final and initial top-oil temperature rise over the ambient temperature at time t and $t - 1$, respectively. τ_w and τ_o are the winding and oil time constant, respectively. Then, Δt is time interval (h).

The $\Delta\theta_t^{HT,f}$ and $\Delta\theta_t^{TA,f}$ can be calculated as the following equations:

$$\Delta\theta_t^{HT,f} = \Delta\theta_r^{HT} k_t^{2m} \quad (44)$$

$$\Delta\theta_t^{TA,f} = \Delta\theta_r^{TA}\left[\frac{k_t^2 R + 1}{R + 1}\right]^n \quad (45)$$

where $\Delta\theta_r^{HT}$ and $\Delta\theta_r^{TA}$ are the winding hottest-spot temperature rise over the top-oil temperature and the top-oil temperature rise over the ambient temperature at rated loading, respectively. k_t is the ratio of transformer between loading at time t and its rated loading. The ratio between the power loss at full load and no load is represented as R. Then, m and n are empirically cooling exponent parameters of winding and oil which depended on transformer cooling mode, respectively.

3) BESS CONSTRAINTS

To extend the service life of BESS within the MG, charging/discharging power and State-of-Charge (SoC) of the BESS should not exceed its limits. These constraints can be shown as follows [23]:

$$\left|P_{B,t}\right| \leq P_{B,r} \quad (46)$$

$$SoC_{\min} \leq SoC_t \leq SoC_{\max} \quad (47)$$

where

$$SoC_t = (1 - \gamma)\, SoC_{t-1} - \frac{E_t}{E_{cap}} \quad (48)$$

$$E_t = \begin{cases} P_{B,t} \cdot \eta_{ch} & \text{charging} \\ P_{B,t}/\eta_{dis} & \text{discharging} \end{cases} \quad (49)$$

where $P_{B,t}$ is the charging(-) or discharging(+) power (kW) of BESS at time t whereas $P_{B,r}$ denotes its power rating (kW). SoC_t is the state-of-charge at time t. $SoC_{\min}$ and $SoC_{\max}$ are the minimum and maximum limits of SoC, respectively. γ is a self-discharge coefficient. E_t and E_{cap} are BESS energy (kWh) at time t and BESS energy (kWh) capacity, respectively. Then, the charging and discharging efficiencies are denoted as η_{ch} and η_{dis}, respectively.

C. BI-LEVEL OEM STRATEGY

In this work, the bi-level OEM strategy is proposed to achieve the objectives under the constraints in the previous subsection. Since the Deep Deterministic Policy Gradient (DDPG) algorithm is employed to solve the OEM problem, thus, the problem is mapped to the Markov Decision Process (MDP) [39]. The DDPG algorithm tries to find the optimal policy in MDP domain. Then, the MDP format for OEM problem of MGO utilized in the first level optimization is described in the second part. The last part, the MDP format of DSO utilized in the second level optimization is presented.

1) DDPG WITH MARKOV DECISION PROCESS

The DDPG is a model-free machine learning that allows an agent to interact with the environment with continuous action for finding best decision [39]. The training process of DDPG makes decisions automatically to achieve the Maximum Cumulative Long-term Reward (MCLR), which follows the optimal policy.

In the dynamic MDP for OEM problem, the agent will interact with the environment in a finite task. A Daily task (24 hours) is considered in this work. Important factors taken into account in DDPG with MDP are Current State (S_t), Continuous Action (A_t), Reward (R_t), and Next State (S_{t+1}) at the step t, and then (S_t, A_t, R_t, S_{t+1}) is stored in the memory buffer and utilized to update the DNN weight that is embedded in an agent.

There are two processes for agent training, namely exploitation, and exploration. The processes can help the agent to discover an optimal policy for mapping from states to the best actions. In the exploitation, it can be described using the following equation [39]:

$$V(S_t \mid \pi) = E[R_t + \gamma V(S_{t+1} \mid \pi)] \quad (50)$$

where $V(S_t \mid \pi)$ denotes the state value function used to estimate the MCLR at the state S_t followed the policy π. E is labeled to show the expected value estimation. Then, γ is the discount factor.

However, the policy π is not changed if the exploration process does not apply to the training. The exploration result is represented in the state-action value which is called Q-value. From the agent training in the ith episode at state S_t, they will not take action following the optimal policy (π_{op}) at the moment but they take new action to provide maximum Q-value and to increase the opportunity to find the new policy which leads to discovering the best policy. Hence, random gaussian noise $N(0, \sigma_t^2)$ is added to the action (A_t)

for the exploration. The exploration rate is reduced by changing the standard deviation (σ_t) using the decay rate (φ) as follows [39]:

$$A_t = A_t + N(0, \sigma_t^2) \tag{51}$$

$$\sigma_t = e^{-\varphi \times t} \tag{52}$$

Moreover, the Q-value can be formulated by Bellman Equation as follows [39]:

$$Q(S_t, A_t) = R(S_t, A_t) + \gamma V(S_{t+1} | \pi_{op}) \tag{53}$$

where $Q(S_t, A_t)$ and $R(S_t, A_t)$ is the Q-value and the reward of the state-action pair (S_t, A_t), respectively. $V(S_{t+1} | \pi_{op})$ denotes the state value function at the state S_{t+1} that followed the optimal policy π_{op}.

One factor that determines the number of Q-values is the number of actions. Continuous action task has a lot of actions. Thus, the number of Q-values will increase exponentially. Since the DDPG is applied to solve the problem, hence, the continuous action and Q-value are predicted by using two Deep Neural Networks (DNNs), which are called actor and critic networks, respectively, as follows:

$$A_t \approx Actor(S_t | \theta^\mu) \tag{54}$$

$$Q(S_t, A_t) \approx Critic(S_t, A_t | \theta^Q) \tag{55}$$

where $Actor(S_t | \theta^\mu)$ is the actor network for mapping from state to optimal action using a weight set θ^μ. Then, $Critic(S_t, A_t | \theta^Q)$ determines as the critic network for mapping from the state-action pair to the maximum Q-value using the weight set θ^Q.

In the training process, the weights of actor and critic networks are updated by applying the sampled policy gradient and minimizing the mean square error, respectively, to find the well-trained weight set, showing more detail in [39]. Furthermore, the target networks of the actor and critic are created and applied to accelerate convergence and improve the stability of DDPG learning. Their networks are generated using the same as the actor and critic structures. Soft updating is employed to change the weight set of the target networks. The soft factor (τ) is set $\tau \ll 1$. Then, the weight set of target actor (θ^μ_{tar}) and target critic (θ^Q_{tar}) are updated using below equations:

$$\theta^\mu_{tar} \leftarrow \tau\theta^\mu + (1-\tau)\theta^\mu_{tar} \tag{56}$$

$$\theta^Q_{tar} \leftarrow \tau\theta^Q + (1-\tau)\theta^Q_{tar} \tag{57}$$

2) FIRST LEVEL

From the objective functions and constraints of the ith MGO, it can be mapped to state ($S^{MGO}_{i,t}$), action ($A^{MGO}_{i,t}$), and reward ($R^{MGO}_{i,t}$) in DDPG with MDP format at time t as follows:

$$S^{MGO}_{i,t} = \left[SoC_{i,t}, P^{PV}_{i,t}, P^{Tr}_{i,t}, S^{Tr}_{i,t}, C^{TOU}_{DSO,t}, C^{pursh}_{DSO,t}, \right.$$
$$\left. W_{carbon,t} \right] \tag{58}$$

$$A^{MGO}_{i,t} = \left[x^B_{i,t} \right] \tag{59}$$

$$R^{MGO}_{i,t} = -\left(f_{1,i,t} + f_{2,i,t} + f_{3,i,t} + PN^{MGO}_{i,t} \right) \tag{60}$$

$$PN^{MGO}_{i,t} = \omega \left(PN^{Current}_{i,t} + PN^{Volt}_{i,t} + PN^{HST}_{i,t} + PN^{TrLoading}_{i,t} \right) \tag{61}$$

$$PN^y_{i,t} = \ln \frac{\left| y_{i,t} - y^{max}_{i,t} \right| + \left| y_{i,t} - y^{min}_{i,t} \right|}{\left(y^{max}_{i,t} - y^{min}_{i,t} \right)} \tag{62}$$

where $P^{Tr}_{i,t}$ and $S^{Tr}_{i,t}$ are the real (kW) and apparent (kVA) power of the transformer in the ith MG at time t, respectively. $P^{PV}_{i,t}$ denotes the total solar PV power (kW) in the ith MG at time t. $PN^{MGO}_{i,t}$ is the constraint penalty value of the ith MGO. Line current ($PN^{Current}_{i,t}$), bus voltage ($PN^{Volt}_{i,t}$), the HST of the transformer ($PN^{HST}_{i,t}$), and transformer loading ($PN^{TrLoading}_{i,t}$) are considered as the ith MGO penalty values. ω denotes as the penalty coefficient. $PN^y_{i,t}$ is the penalty value of the y parameter calculated by using (62). Then, $y_{i,t}$ is the parameter value with the boundary [$y^{min}_{i,t}$, $y^{max}_{i,t}$]. Also, each parameter of the state fed to DNN of MGO agent is normalized within the range of [−1, 1].

Moreover, $x^B_{i,t}$ is charging/discharging factor of BESS which is determined as an action variable of the ith MGO agent at time t. The purpose of installing the BESS is to store the remaining power from solar PV generation and help reduce the transformer burden appropriately. Therefore, charging or discharging power ($P^B_{i,t}$) of BESS in the ith MG at time t depend on the factor and transformer loading ($P^{Tr}_{i,t}$) which can be defined as follows:

$$P^B_{i,t} = x^B_{i,t} \cdot P^{Tr}_{i,t} \tag{63}$$

where $x^B_{i,t}$ has the boundary as [$0, a$] when a is a positive integer.

3) SECOND LEVEL

From the objective functions and constraints of the DSO, it can be mapped to state (S^{DSO}_t), action (A^{DSO}_t), and reward (R^{DSO}_t) in DDPG with MDP format at time t as follows:

$$S^{DSO}_t = \left[P^{TrDN}_t, P^{Tr}_{i,t}, S^{Tr}_{i,t}, C^{TOU}_{maingrid,t}, W_{carbon,t} \right] \tag{64}$$

$$A^{DSO}_t = \left[C^{pursh}_{DSO,t} \right] \tag{65}$$

$$R^{DSO}_t = -\left(F_{1,t} + F_{2,t} + PN^{DSO}_t \right) \tag{66}$$

$$PN^{DSO}_t = \alpha \left(PN^{Current}_t + PN^{Volt}_t \right) \tag{67}$$

where P^{TrDN}_t denotes the DN transformer loading (MW) at time t. PN^{DSO}_t is the constraint penalty value of the DSO at time t. Line current ($PN^{Current}_t$) and bus voltage (PN^{Volt}_t) are considered as the DSO penalty values only because the DN transformer rating is large enough to supply the total demand of the DN with three MGs. Moreover, $PN^{Current}_t$ and PN^{Volt}_t can be calculated by using (62). Then, α is the penalty coefficient. Also, each parameter of the state fed to DNN of DSO agent is normalized within the range of [−1,1].

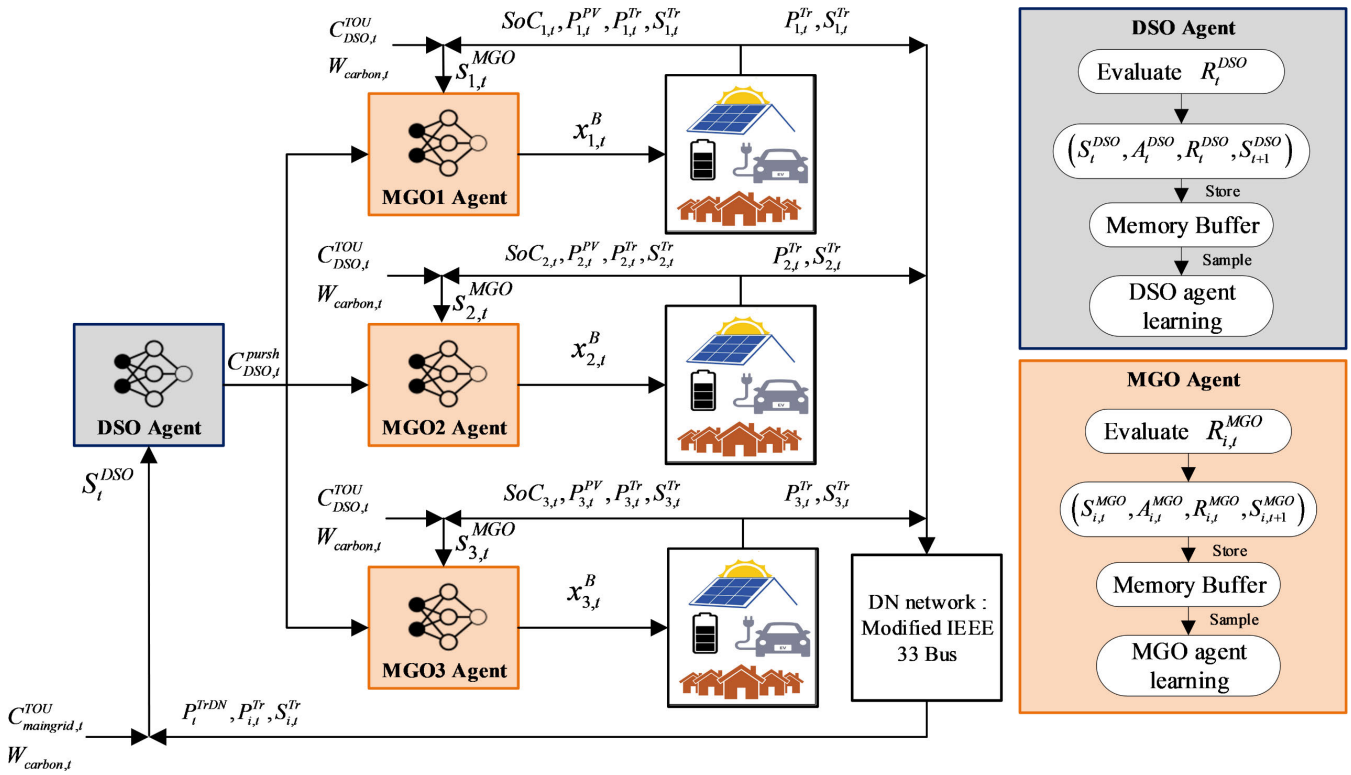For all the above two levels, the DDPG with MDP framework can be shown in Fig. 2.

FIGURE 2. The DDPG with MDP framework for the bi-level OEM of the single DSO with three MGOs.

V. METHODOLOGY

The OEM problem is more complex when considering multi-level optimization. Commonly, the OEM problem is improved by using numerical methods, leading to the computational burden, especially the bi-level OEM problem. In addition, to guarantee the global optimum, it is necessary to consider the uncertainties of solar PV generation and overall demand using PPF. This situation leads to the computational burden exponentially. Opting for a robust numerical method and increased PPF calculation speed can guarantee the global optimum and decrease the computational burden. Thus, the DNN surrogate-assisted multi-agent DDPG algorithm is employed to solve the OEM problem in this work.

A. DNN SURROGATE MODEL

Deep Neural Network (DNN) trained through deep learning has an impressive capability for learning complex system behavior [60], especially the power system with high uncertainty. Moreover, the performance of MG behavior learning using the DNN surrogate model is guaranteed by Xiao et al. [61] that it has good adaptability to predict the MG behavior required high-dimension inputs and provides high accuracy for the prediction. To mitigate the computational burden of solving the OEM problem, DNN surrogate models are developed and used to substitute the PPF for estimating the power system parameters of three MGs.

1) DNN ARCHITECTURE

The DNN architecture includes an input layer, a multi-hidden layer, and an output layer which can be shown in Fig. 3. The multi-hidden layer is essential for deep learning, which can remember a complex relationship of the problem. Also, the number of neurons in each layer can be determined. In each neuron, all inputs are multiplied by the weight set and brought together with the bias. Subsequently, the outputs are determined by the activation function. In deep learning, the weight set of all neurons are updated to provide the correct output.

2) MG BEHAVIOR GENERATION

Since deep learning is one of the supervised learning, thus data generation for the learning is essential. The purpose of developing the DNN surrogate model is to predict power system parameters of each MG instead of the PPF calculation. Hence, the data is generated by using the PPF calculation.

Pandapower tool is an open-source power systems analysis based on Python programming language. The data structure of the power system can construct and import to the tool easily. In addition, power flow results can be evaluated and exported simply [62]. The unbalance DPF can be calculated using the tool perfect for the PPF calculation in the low voltage distribution system such as the MG. Also, the performance of the tool is guaranteed by Thurner et al. [62], which can improve the computational burden and accuracy
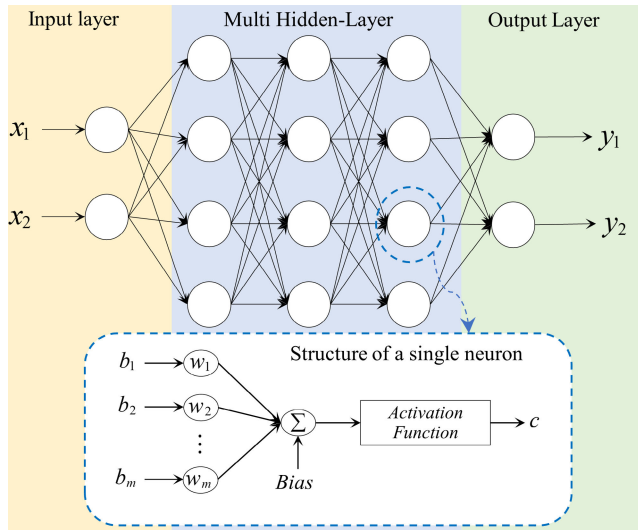
FIGURE 3. The structure of a deep neural network (DNN) [23].

FIGURE 4. MG behavior generation using LHS and NT with PEM.

compared with MATPOWER and others. Hence, the MG is constructed as an Actual Engineering Model (AEM) for DPF calculation in Pandapower.

Latin Hypercube Sampling (LHS) method and the NT with Zhao's PEM are applied for the MG behavior generation, which can be shown in Fig. 4. It starts with the MG construction in the Pandapower determined as the AEM. In the second step, the charging/discharging factor of BESS is sampled in the feasible space to generate 1,000 scenarios per hour using the LHS. After that, the input variables of the PPF are imported to calculate the PPF using NT with Zhao's PEM for each scenario. The fourth step evaluates the mean of desired power parameters such as maximum/minimum bus voltage (V), maximum line current (kA), and transformer loading in kW and kVA units at time t. Finally, the parameters are saved to CSV files.

3) MG BEHAVIOR LEARNING

Keras Library based on Python language is the most constructed deep learning framework, which has many modules available for choices, such as layer modules, optimization modules, and activation function modules [63]. In addition, those modules can be implemented simply. Therefore, the DNN developed as the MG behavior is created and learned in the Keras library. The data set of the MG behavior is divided into two parts: training 80% and testing 20%, respectively. The DNN configuration is set according to [23]. Root Mean Square Error (RMSE) and R-squared are used to guarantee the performance of the DNN leaning.

B. DNN SURROGATE-ASSISTED MUTI-AGENT DDPG ALGORITHM

From the previous subsection, the DNN surrogate model is developed to predict MG behavior completely. Then, in this subsection, multi-agent DDPG algorithm is assisted by three

DNN surrogates to improve the computational burden for solving the OEM problem.

1) THE TRAINING PROCESS

To find the best DNN weight set of the DSO and MGO agents, the training process is essential to achieve the purpose. Four agents, such as a single DSO and three MGOs are trained simultaneously. The agents try to interact with their environment to learn and remember the best control solution.

The training of the DNN surrogate-assisted multi-agent DDPG algorithm is shown in Fig. 5. In the first step, DDPG and DNN-agent parameters are set. Then, the purchased energy price randomized by the DSO agent in each hour is proposed to three MGOs. Subsequently, three MGO agents randomly select the charging/discharging factor to evaluate the BESS power. Next, the DNN surrogate models predict the power system parameters by feeding the BESS power obtained by the previous step into the models. Also, the objective functions of each MGO are evaluated in this step. After that, the objective and penalty values are mapped to the MGO reward. Next, each MGO transmits the transformer loading states to the DSO to evaluate power system parameters of the DN using the DPF calculation in Pandapower. The objective functions and penalties of the DN are estimated and mapped to the DSO reward in the next step. Finally, the parameters associated with the DDPG with MDP of each agent are transmitted into its memory which is used to update the weight set of the agent. The above steps are repeated 24 times (24 hours) in each episode to estimate 24 rewards. After that, the rewards of each agent are summed as the score. The average score of each agent is calculated by using previous scores of N_{ep}. Subsequently, the mean of all average scores of N_{agent} are
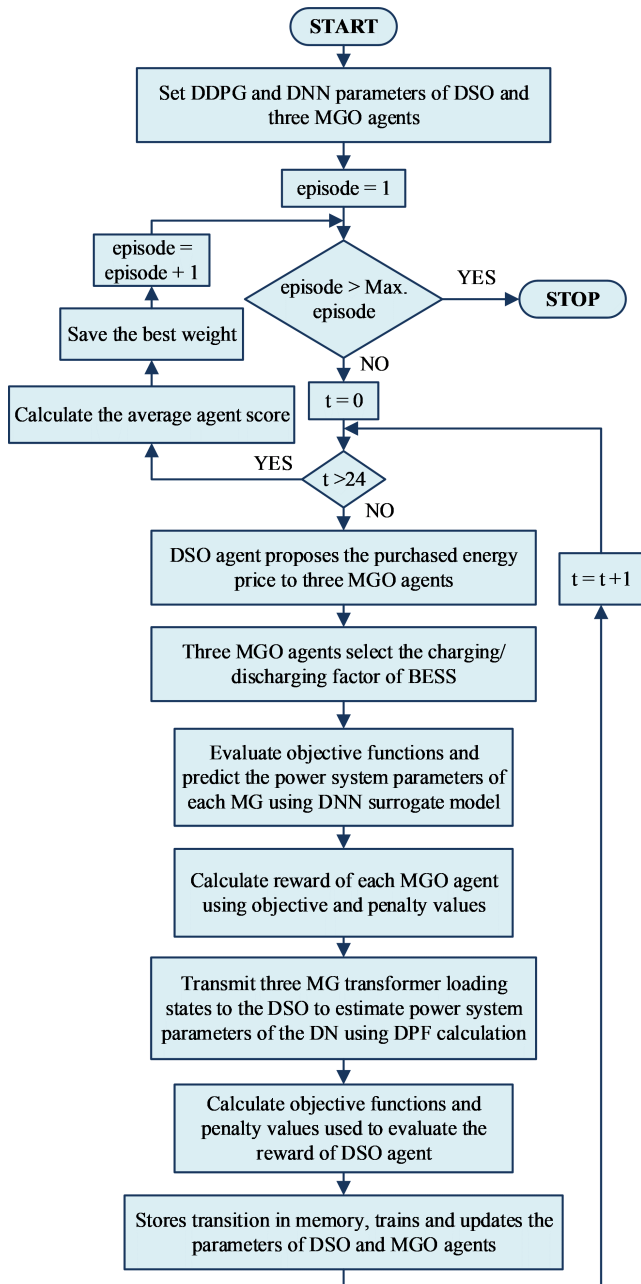
FIGURE 5. The training process for DNN surrogate-assisted multi-agent DDPG algorithm.
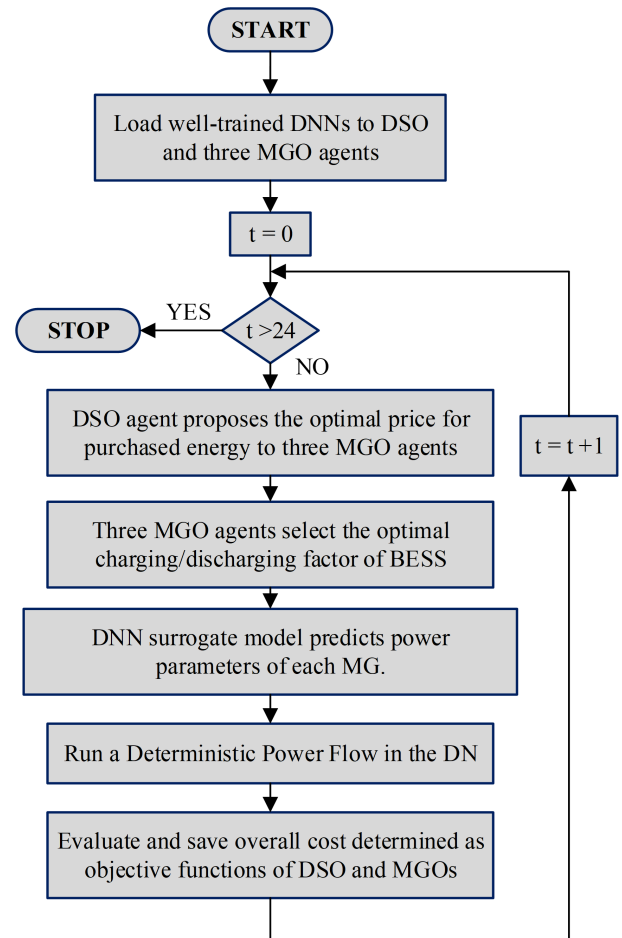

FIGURE 6. The testing process for DNN surrogate-assisted multi-agent DDPG algorithm.
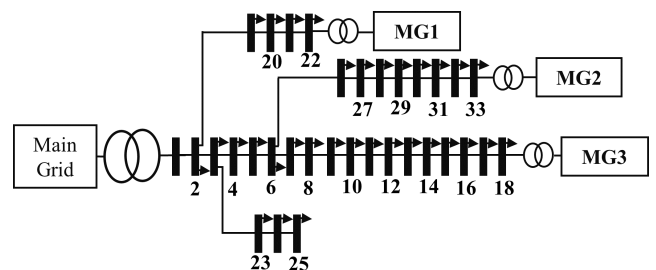
FIGURE 7. Single-line diagram of the modified IEEE-33 test system.

estimated as the average agent score. After that, if the average agent score at the ep^{th} is more than the average agent score obtained by the previous episode, the weight set of all agents is saved. Then, the episode is updated to the next episode. The average scores of all and each agent can be formulated as the following equations:

$$S_{ep}^{avag} = \frac{\sum_{j=1}^{N_{agent}} S_{j,ep}^{av}}{N_{agent}}, \quad S_{j,ep}^{av} = \frac{\sum_{m=ep-N_{ep}}^{N_{ep}} S_{j,m}}{N_{ep}} \quad (68)$$

where S_{ep}^{avag} is the average agent score of N_{agent} at the ep^{th}. Also, $S_{j,ep}^{av}$ denotes the average score of previous N_{ep} of the

j^{th} agent whereas $S_{j,m}$ is the score at the mth episode of the jth agent.

2) THE TESTING PROCESS

From the training, the single DSO and three MGOs behaviors are trained to remember the OEM solution. The well-trained weights saved from the train are loaded to the agents to control the test system appropriately which is called the testing process. The process is shown in Fig. 6.

In Fig. 6, the process starts with the well-trained DNN loading of all agents. Then, the optimal price for purchased energy in each hour is proposed by the DSO agent. After that,
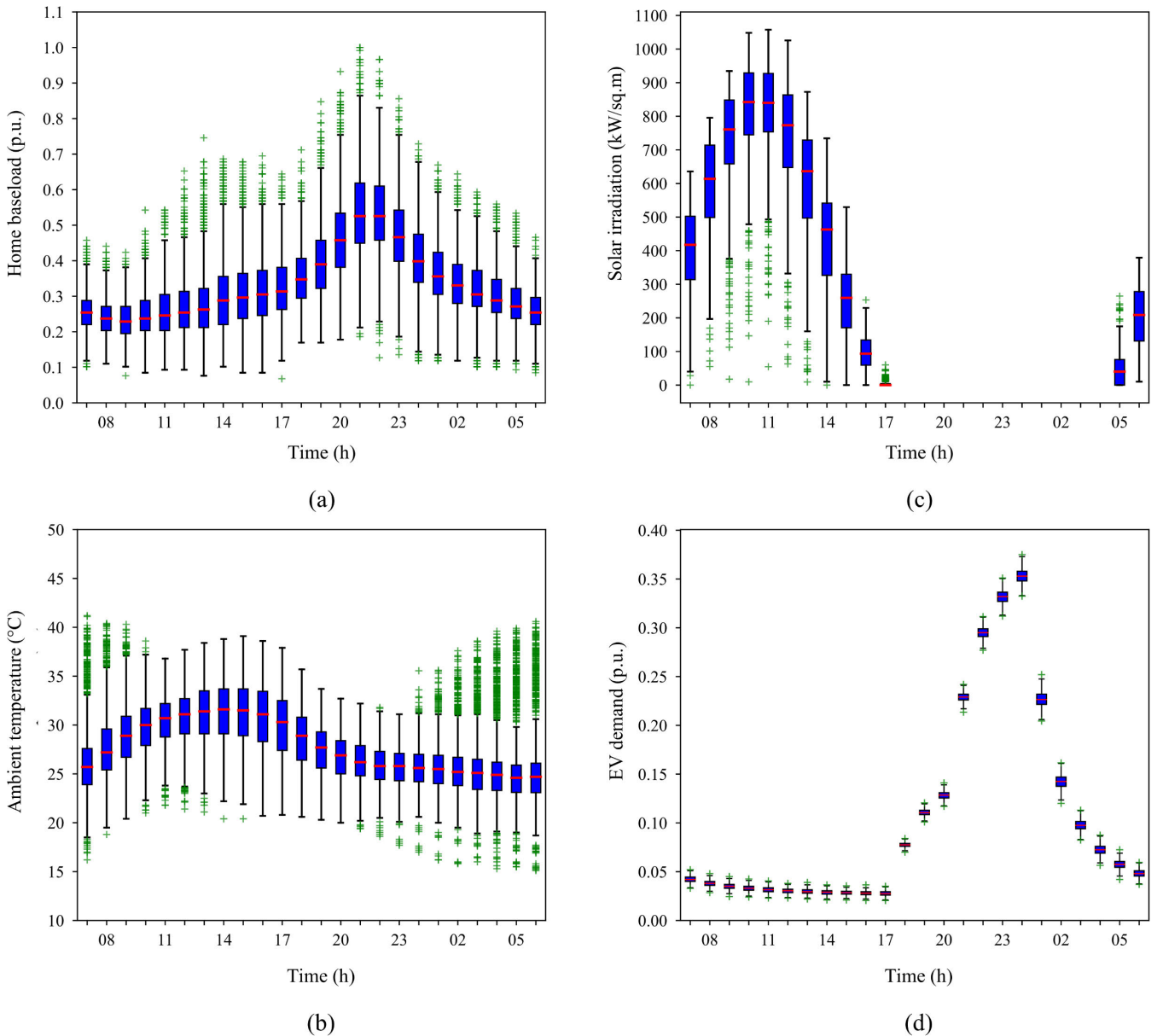
FIGURE 8. The distribution of four random inputs of the PPF in MG; (a) Home baseload, (b) Ambient temperature, (c) Solar irradiation and (d) EV demand [23].

each MGO agent will select the optimal charging/discharging factor of its BESS to respond to the price of the DSO. Subsequently, three DNN surrogate models predict the power system parameters of three MGs and send the MG transformer loading status to the DSO for the DPF calculation. Then, the overall cost of DSO and MGO are evaluated and saved. The above steps are repeated 24 times (24 hours).

VI. SIMULATION RESULTS
A. ASSUMPTIONS
1) DN AND MG DEFINITION
A low voltage distribution network in reference [23] is modeled as the MG. The MG is a village that has 36 buses

and 27 households. Each household is assumed to have solar PV, a home baseload, and two EVs. Moreover, the BESS is installed near the transformer of the MG. There are three MGs in this work. Maximum loads of three MG are scaled by 100%, 50%, and 25% of the maximum load according to [23], which are called MG1, MG2, and MG3, respectively. Three MGs are connected to the DN through the 100kVA, 22kV/400V distribution transformers located at the 22nd, 33rd, and 18th buses, which is shown in Fig. 7. The transformer specification is determined according to [23]. The IEEE 33-bus radial distribution system is modified and used as a distribution network in this work. The electrical elements of the DN are modified to conform to a Provincial Electricity Authority (PEA) medium voltage distribution
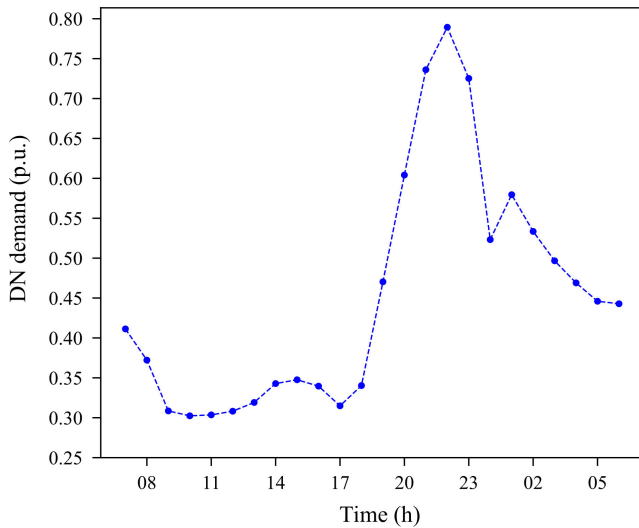
FIGURE 9. DN load profile [64].

TABLE 2. The VRFB specification.

Parameter	Value	Parameter	Value
$P_{B,r}$ (kW)	80	C_{cap} ($)	79,080
E_{cap} (kWh)	450	β_0	4,901
γ	0.001	β_1	1.98
η_{ch}	0.85	β_2	0.016
η_{dis}	0.80	$\%SoC_{\min}, \%SoC_{\max}$	10, 100

network standard, such as transformer specification and line impedance. The network is connected to the main grid through the 50MVA, 115kV/23.1kV power transformer. All DN peak loads are scaled by 2.7 times the original load of the IEEE 33-buses.

In Thailand, the power system load typically peaks in summer over four months (March - June). This situation leads to difficulty in energy management and challenges for the EMS design. Hence, the residential load less than and more than 150 kWh of summer in Udon Thani, Thailand, are modeled as load profiles of MG and DN, respectively. The load data are provided by Provincial Electricity Authority (PEA) [64] from 2017 – 2020. Moreover, The hourly ambient temperature and solar irradiation data in summer over three years (2015 – 2017) used to estimate solar PV power are obtained from the Thai Meteorological Department [65] and the Department of Alternative Energy Development and Efficiency in Thailand [66], respectively. The rated power of solar PV generation is determined as 5 kW for rooftop solar PV of each home. Furthermore, the stochastic EV demand is generated using four EV types and random daily travel distance according to [23] and [47], respectively. The home-rated charging power of all EV types is defined as 3.3 kW whereas the charging power factor is determined as 0.95. In this work, all EVs are assumed to be simultaneously charged by the rated power. To generate the stochastic EV demand, the MSC with 10,000 iterations is applied to estimate the hourly charging probabilistic of EVs. All the above data used in the MG can be represented through the boxplot in Fig. 8, whereas the DN load profile is shown in Fig. 9.

Vanadium Redox Flow Battery (VRFB) is a popular battery employed to store/supply electrical energy in the power system, especially the system with RERs generation, which is called the commercial BESS. The VRFB can be designed suitable for each power system with limited power and capacity, whereas it has a low maintenance cost and a long

operational life [67]. Therefore, the VRFB is modeled as the BESS installed in each MG. The VRFB specification, cost parameters, and curve-fitting coefficients are determined according to [23], [68], and [58], respectively, which are presented in Table 2.

Moreover, the hourly carbon emission rate (kg/kWh) is defined by the International Energy Agency (IEA) [56]. In the electricity sector, the capturing carbon dioxide rate ($/kg) is evaluated using EUR 60 per tonne of carbon price benchmark defined by the Organization for Economic Co-operation and Development (OECD) [57].

2) STRATEGY
To demonstrate the performance of the proposed bi-level OEM using DNN surrogate-assisted multi-agent DDPG algorithm, there are four strategies in this work:

- Rule-based control strategy (the first strategy): All BESS operations are controlled by rule-based control. In the strategy, the role of the BESS is defined as storing the remaining power from solar PV generation and reducing the transformer burden. The power rating of the transformer is determined as 80 kVA, expressed as a rule-based condition. If transformer loading is more than 80 kVA, BESS will discharge to alleviate the transformer burden. Otherwise, BESS will charge or not operate. Also, the original TOU and FIT are used as the energy selling and purchase prices of the DSO to propose three to MGOs.
- Fixed-purchased energy price strategy (the second strategy): All BESSs are controlled by three MGO agents to minimize their operation costs. In contrast, the original TOU and FIT are used as the energy selling and purchase prices of the DSO to propose to three MGOs.
- Maximum-purchased energy price strategy (the third strategy): All BESSs are controlled by three MGO agents to minimize their operation costs. The original TOU is defined as the selling energy price of the DSO. In contrast, the TOU of the main grid is utilized as the maximum price of possible purchased energy of the DSO to motivate the attention of three MGOs for selling their energy to the DN instead of purchasing energy from the main grid. Then, the above selling and purchase prices of the DSO are proposed to three MGOs.
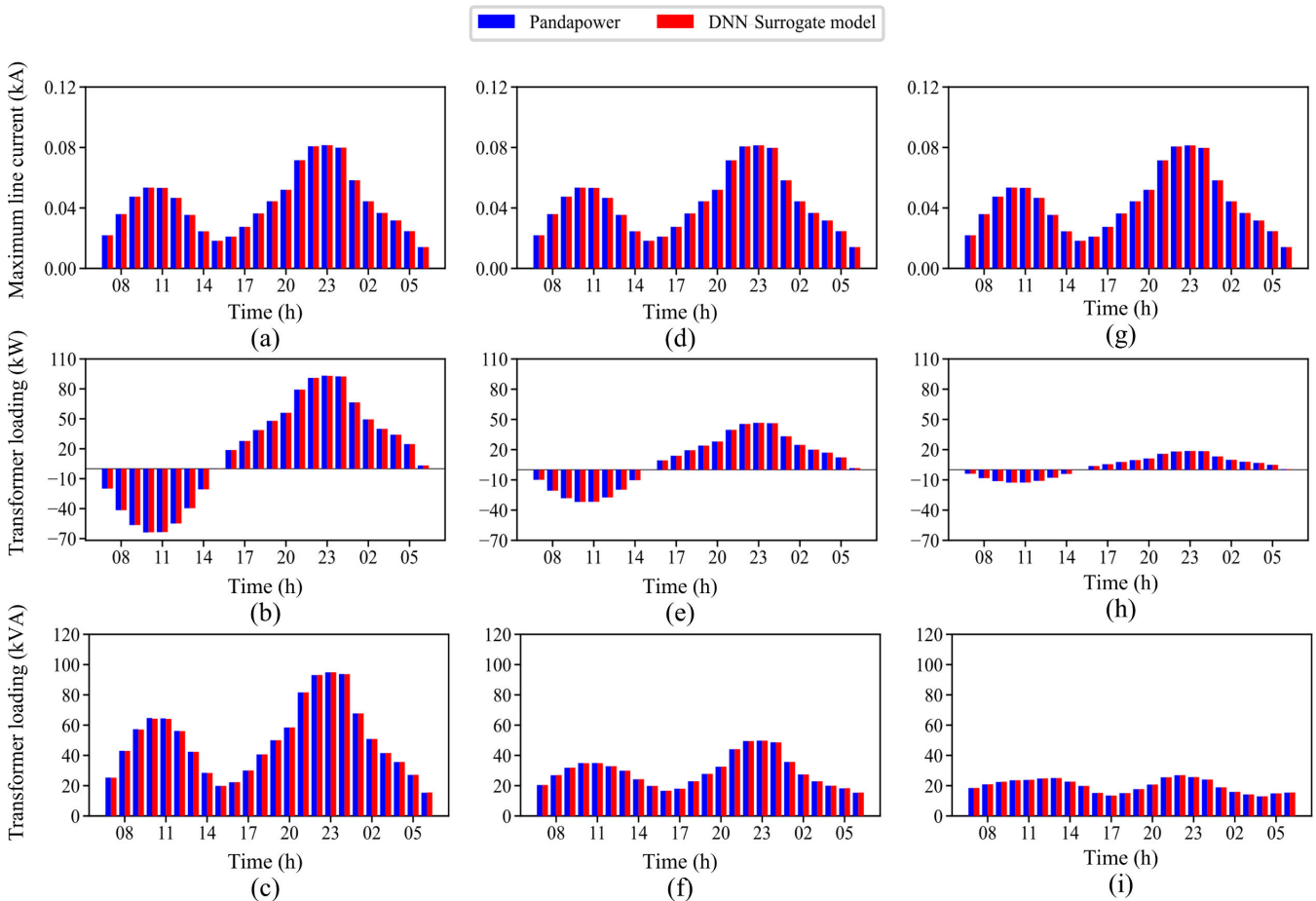
FIGURE 10. The mean hourly maximum line current, transformer loading in kW and kVA units using the DNN surrogate model and the PPF in AEM using Pandapower; (a) – (c) BESS factor is 0.0, (d) – (f) BESS factor is 0.5 and (g) – (i) BESS factor is 0.8.

- Optimal-purchased energy price strategy (the proposed strategy): All BESSs are controlled by three MGO agents to minimize their operation costs. The original TOU and the real-time FIT submitted by the DSO agent are utilized as the energy selling and purchase prices of the DSO which are proposed to three MGOs.

B. PREDICTIVE PERFORMANCE OF THE DNN SURROGATE MODEL

To verify the accuracy performance of DNN surrogate models used to predict the power system parameters of MG, the performance of the models is compared with the PPF calculation using NT with Zhao's PEM in the AEM. Spyder program based on Python language is utilized to develop the models, whereas the PPF is calculated according to scenarios generated by NT with Zhao's PEM using unbalanced power flow in the Pandapower library. A personal computer is used to run the above process. The computer specification consists of Intel(R), Core (TM) i7-8700, CPU 3.20GHz, and 16.0GB RAM.

The mean hourly power parameters of MG2 are used to verify the accuracy performance, which is tested using three

different BESS factors. Some parameters, such as hourly maximum line current (kA) and hourly transformer loading in kW and kVA units, are shown in Fig. 10. The MG2 parameter estimation results are represented in Table 3. The results show that the maximum absolute errors that occur in 24 hours of all parameters using the DNN surrogate model are less than 1% compared with the PPF in AEM using the Pandapower. Moreover, the computational time of the PPF in AEM varies from about 63 – 66 seconds, whereas the DNN surrogate model is only 4 seconds, reducing the computational time by about 93%.

C. DNN SURROGATE-ASSISTED DDPG VERIFICATION

From the previous subsection, the DNN surrogate models can provide high accuracy and exponentially decrease the computational time for evaluating the power system parameters. To deal with the computational burden of the optimization task, the DNN surrogate model has assisted the optimization algorithm. In this work, the DDPG algorithm with the DNN surrogate model is applied to solve the OEM problem, which is represented as the proposed method. Moreover, the Differential Evolution (DE) algorithm, guaranteed as one of

TABLE 3. The comparison of maximum absolute errors and computational times between the DNN surrogate model and the PPF using Pandapower.

BESS factor	Maximum absolute error occurred in 24 hours					Computational time (s)		
	Maximum Voltage (%V)	Minimum Voltage (%V)	Maximum Current (%kA)	Transformer loading (%kW)	Transformer loading (%kVA)	PPF using Pandapower	DNN Surrogate model	%Different
0.0	0.120	0.060	0.095	0.903	0.873	66	4	93.94%
0.5	0.080	0.064	0.065	0.979	0.471	64	4	93.75%
0.8	0.084	0.067	0.072	0.970	0.684	63	4	93.65%

TABLE 4. The TOU rates for MG and DN [54], [55].

Interval (h)	TOU rate for MG ($/kWh)	TOU rate for DN ($/kWh)
$9 \leq t < 22$	0.1143	0.0989
$22 \leq t < 24$	0.0711	0.0638
$1 \leq t < 9$	0.0711	0.0638

TABLE 5. The performance of DE and DDPG.

Algorithm	Median fitness value ($)	Training time (s)
DE	60.880	63,150
DDPG	61.724	6,804
% Difference	1.39%	89.23%

FIGURE 11. The distribution of fitness values using DDPG and DE.

the best tools of metaheuristic algorithms and employed to solve many problems according to [40], [41], [42], and [43], is applied with the DNN surrogate model to verify the performance of proposed method.

The MG2 is modeled as the test system for the optimization task and is only one MG connected to the DN. The TOU and FIT rates are fixed as shown in Table 4 [54], [55] and 0.0014 $/kWh [69] for every hour, respectively. Exchanged energy cost, carbon emission cost, and BESS degradation cost are modeled as the objective functions. For the DDPG parameters, the actor and critic networks have two hidden layers and 400 neurons in each layer. The actor and critic learning rates are 0.005 and 0.02, respectively. The soft factor is defined as 0.005, whereas 0.9 is assigned as the discount factor value. Also, the decay rate is 0.002. For the DE parameters, A population size is set to 10. Scaling factor and crossover rate are defined as 0.5. Furthermore, the fitness value is defined as the sum of all objective and penalty values in 24 hours. The fitness values are evaluated 10 times using 100 episodes per time. From the above condition, the performance results can be shown in Fig. 11 and Table 5, respectively.

In Fig. 11, the results show that the fitness value using the DE varies from about 58 – 65, whereas the fitness value us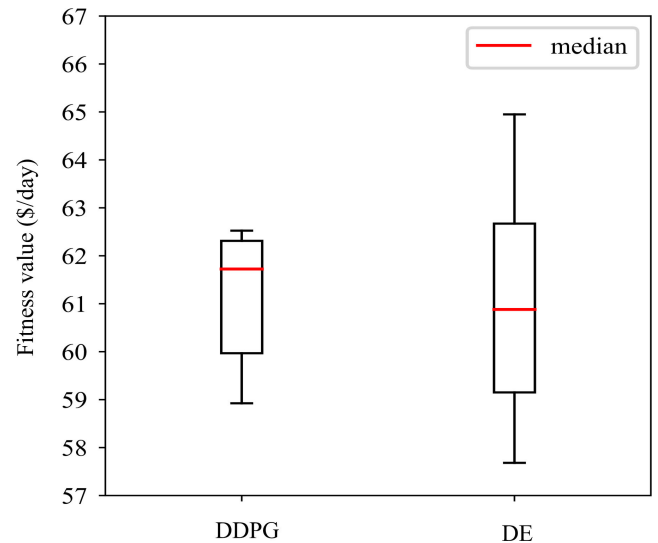ing the DDPG varies in a narrow range from about 59 – 63 compared with the DE. In Table 5, the median of all fitness values obtained by the DDPG is more than the DE but does not exceed 2%. However, the DDPG can decrease computational time by about 89% compared with the DE. Therefore, the above result can guarantee the accuracy of the DDPG and significantly reduce the computational burden.

D. THE BI-LEVEL OEM USING DNN SURROGATE-ASSISTED MULTI-AGENT DDPG ALGORITHM

From the previous subsections, the DNN surrogate-assisted DDPG optimization can improve accuracy and computational burden for solving the OEM problem of a single MG. Hereafter, the proposed method is applied to improve the performance of the multi-agent optimization for solving the bi-level OEM problem within a single DN with three MGs. Therefore, DNN surrogate-assisted multi-agent DDPG algorithm is proposed to reduce the computational burden. Also, it is applied in the last three strategies to solve the problem and to discover the best strategy for this work. There are three DNN surrogate models that are developed to predict power system parameters of three MGs in multi-agent DDPG optimization. Furthermore, the performance of OEM with RTP presented as the proposed strategy is compared with the first three strategies shown in beginning of this subsection.
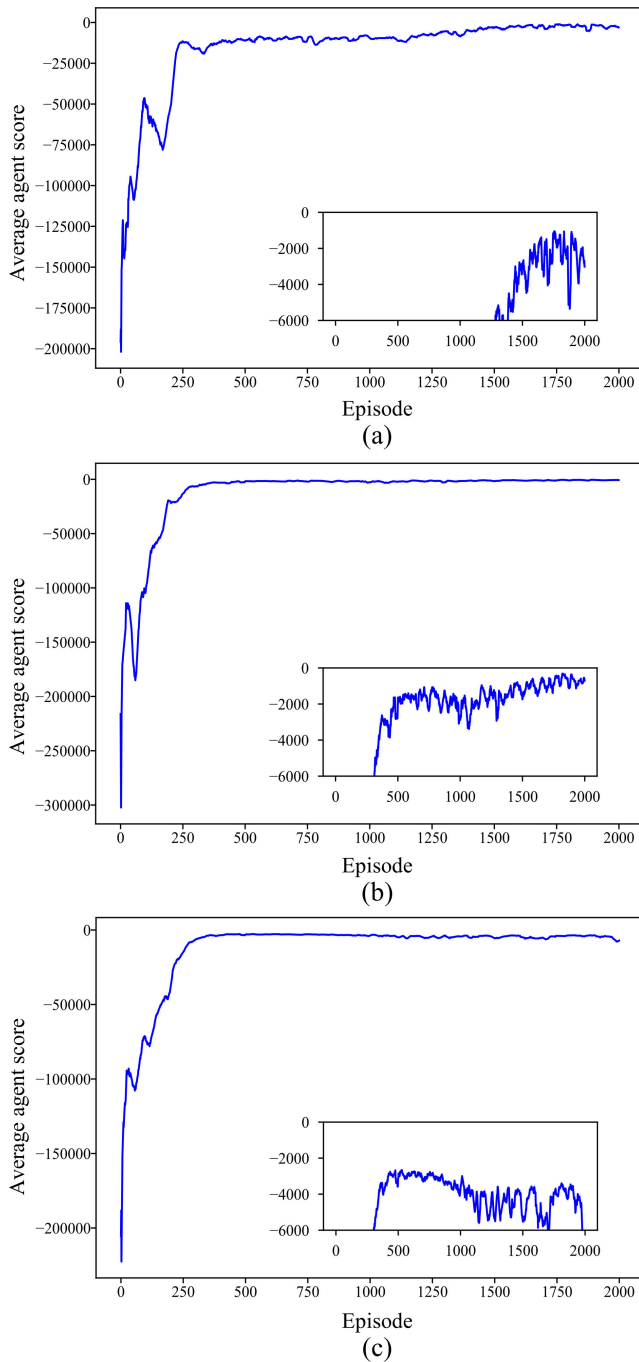
FIGURE 12. The average agent score of the three strategies, (a) Second strategy, (b) Third strategy and (c) Proposed strategy.

In the first strategy, all BESS are controlled using the rule-based control strategy whereas they are controlled by three MGO agents designed to follow the DDPG concept in the last three strategies. In the second and third strategies, only three MGO agents handle three BESS, whereas DSO is responsible for ensuring the operation constraints of the DN. However, the DSO agent can generate and propose the purchased energy price to three MGO agents in the fourth strategy (the proposed strategy). The DDPG parameters of

each agent are set the same as in the previous subsection, but the maximum number of episodes is determined as 2,000. The average agent score of the last three strategies can be shown in Fig. 12. Numerical results show that the average agent score will converge to the best value approximately in the 300th episode for all the above strategies. However, the training will continue until it discovers the well-trained weight set of all agents. In the last three strategies, the well-trained weight set is saved about the 1700th, 1900th, and 500th episodes, respectively. The best score of the third strategy is more than the score obtained by the second strategy. Thus, the overall cost of the MGO can reduce more when applying the purchased energy price as the TOU of the main grid. Furthermore, the results of the second and third strategies cannot compare with the fourth strategy because the average agent score provided by the proposed strategy is calculated from four agents (three MGOs and a DSO). In contrast, it is estimated from three agents (three MGOs) in the second and third strategies. However, the comparison results of all strategies are shown in the testing result.

In the testing process, well-trained weight sets are loaded into all agent models to control the system in the last three strategies, not including the first strategy. For the testing results, the power state and state of charge of BESS, and parameters of the transformers can be presented in Fig. 13 and Fig. 14, respectively. In Fig. 13, it shows that surplus power (negative red color) stored in the BESS of MG3 is not discharged to reduce the transformer burden when applying the first strategy because the normal baseload of the transformer does not exceed 80 kVA. Also, the BESS of MG2 only discharges the power to decrease the transformer burden sometimes. Thus, the surplus power of the above MGs is not utilized as expected, whereas it is completely discharged to reduce the transformer burden in MG1. Nevertherless, the surplus power is not adequate to decrease the transformer burden of MG1 when applying the first strategy, which causes the transformer limit violation such as transformer overloading, HST violation and increase of aging acceleration factor during certain hours like Fig. 14. For the last three strategies, the well-trained weight sets are loaded into all agents to schedule the charging/discharging power of the BESS in all MGs appropriately. Also, it can prevent violation problems in the transformer of MG1, as shown in Fig. 14. Nevertheless, the above strategies have different BESS dispatch in some MGs. Still, the BESS dispatch in MG1 has a similar operation when employing the above three strategies. The total amount of surplus power within MG1, which is stored in the BESS, is not enough for supplying to assist the transformer loading reduction, as noticed in Fig. 13(b). Hence, in the last three strategies, the power is drawn into MG1 through the transformer during the surplus power period (07.00 – 14.00) to charge the BESS with the surplus power. This situation causes the charged energies of BESS to be more than the energies due to applying the first strategy. Moreover, applying the last three strategies can protect the violation scenarios of the MG1 transformer limitation, as represented in Fig. 14(d)
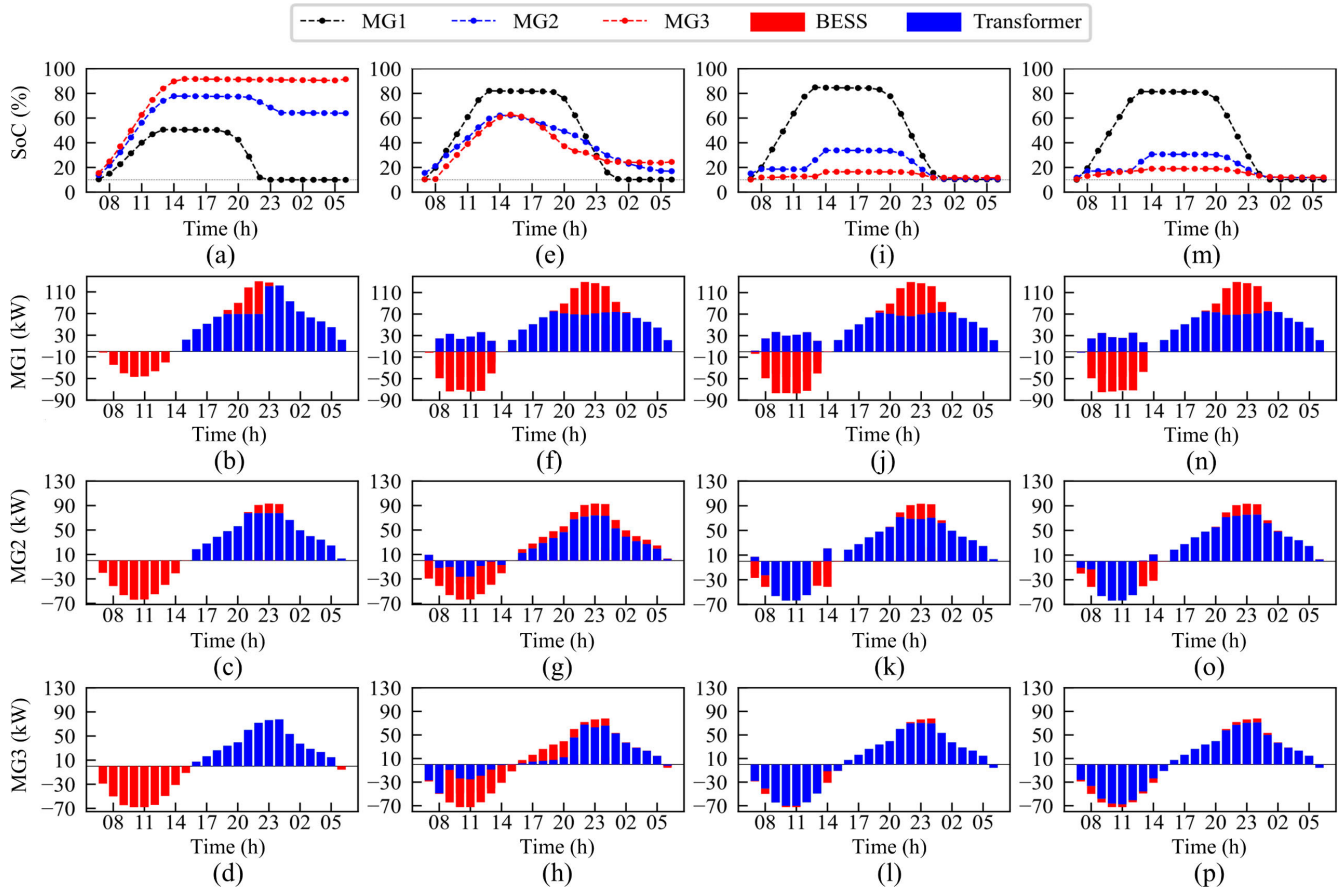
FIGURE 13. The power state and state of charge of BESS in each MG; (a) – (d) First strategy, (e) – (h) Second strategy, (i) – (l) Third strategy and (m) – (p) Proposed strategy.

to Fig. 14(l). For the difference in BESS dispatch, the charged energies in BESSs of MG2 and MG3 in the second strategy are more than the energies due to applying the third and proposed strategies, which can be noticed from an increased SoC on the charging period (07.00 - 15.00) in Fig. 13(e). Additionally, the charged energies of BESSs in MG2 and MG3 are not completely discharged which causes a daily unbalanced charging and discharging energy situation noticed from the end SoC. However, the problem is improved when the DSO offers the purchased energy price using the TOU of the main grid (third strategy) and the price presented by the proposed strategy, as noticed from the end SoC of MG2&MG3 BESS in Fig. 13(i) and Fig. 13(m).

In addition, the BESS is almost dispatched every hour when applying the second strategy, as shown in Fig. 13(g) and Fig. 13(h). In contrast, it is used at certain hours when employing the third and proposed strategies observed from MG2 & MG3 power states plotted in Fig. 13(k), Fig. 13(l), Fig. 13(o), and Fig. 13(p). Thus, the BESS degradation cost in the second strategy is more than the third and proposed strategies, as shown in Fig. 15.

The overall cost of each MG is shown in Fig. 15. In the last three strategies, the overall costs of MG1 have

a similar value. This situation is a consequence of the limited surplus power and increased demand within MG1. The BESS can not be dispatched for energy trading, but it is controlled to decrease the transformer burden only. Hence, the injected energy cost is equal to 0$ per day. In contrast, MGO2 and MGO3 agents try to control the BESS for energy trading to minimize the overall cost. Employing the third and proposed strategies can motivate the MGO2 and MGO3 to inject an amount of energy into the DN more than the second strategy, presented in Fig. 16. For MG2 and MG3, although the drawn energies and the drawn energy costs in the third and proposed strategies are more than the second strategy, the net energy cost of those strategies is still less than the second strategy.

The overall costs of DSO are presented in Table 6. The results show that the carbon emission costs of all strategies have approximately the same value which is 83$ per day. The purchased energy costs from the main grid in the third and proposed strategies are roughly the same. Also, they are less than the first two strategies. For the first strategy, the purchased energy cost from the main grid is high because the surplus energy within MG2, and MG3 does not inject into the DN observed from the SoC in the first strategy plotted in Fig. 13(a). Thus, the DSO greatly receives power from the
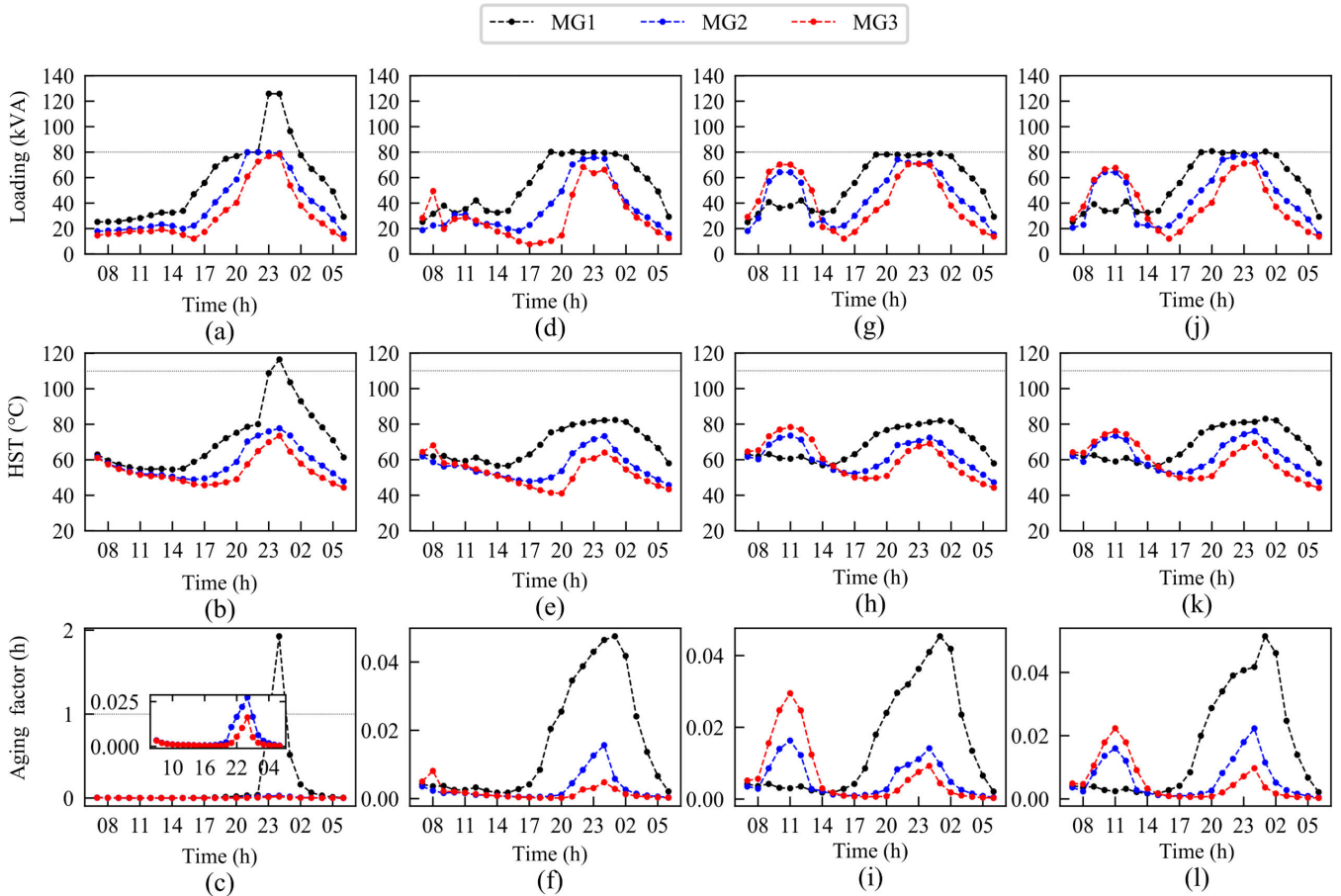
FIGURE 14. The parameters of each MG transformer; (a) – (c) First strategy, (d) – (f) Second strategy, (g) – (i) Third strategy and (j) – (l) Proposed strategy.

main grid. Moreover, the purchased energy costs from three MGs are equal to 0$ per day in this strategy, whereas it is increased to 3.50$ per day in the second strategy. Nevertheless, the purchased energy costs from the main grid in the first and second strategies are still high. In contrast, it can be reduced to about 9,011$ per day when employing the third and proposed strategies. The purchased energy cost from three MGs is still increased when applying the third strategy because the purchased price submitted to three MGOs is exponentially higher than the original FIT, which increases the total cost of DSO to about 9,158$ per day.

Furthermore, the purchased energy price presented by the DSO agent in the proposed strategy is lower than the third strategy, as shown in Fig. 17, where it can attract the attention of MGO2 and MGO3 to inject their energies into the DN approximately equal to the third strategy noticed from injected energy in Fig. 16. Therefore, the total cost of DSO when employing the proposed strategy decreases from about 9,158$ to 9,117$ per day compared with the third strategy, which is the lowest cost in this work. Also, the proposed strategy can decrease the total cost between 0.01% to 0.44% compared with the cost provided by the first three strategies.

VII. DISCUSSION

The heavy computational burden and performance of the EMS for solving the OEM problem within single DN with multi-MGs, especially dealing with the uncertainty of RERs generation and DN&MG system constraints, is a significant problem that should be improved. Thus, MG behavior developed as a DNN surrogate model is applied to mitigate the computational time due to the PPF calculation. According to the verified results in Table 3, the DNN surrogate model can guarantee the high accuracy of the predicted MG parameters. The results show that the maximum absolute errors that occur in 24 hours of all parameters using the DNN surrogate model are less than 1% compared with the PPF in AEM. Moreover, the performance of the DNN surrogate-assisted DDPG algorithm (proposed method) is improved and compared with the DNN surrogate-assisted metaheuristic algorithm, as shown in Fig. 11. The proposed method can approximately provide optimal solutions and reduce the computational burden to 89.23%, as shown in Table 5. Therefore, the proposed method is guaranteed from the above result that can improve performance and decrease the computational burden of the OEM task.
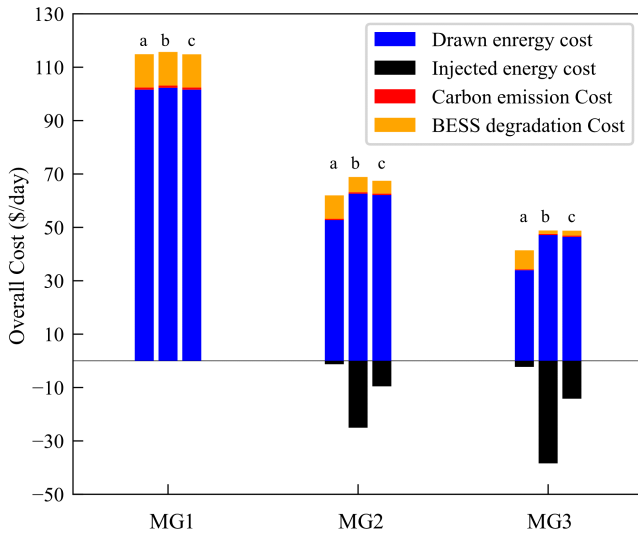
FIGURE 15. Overall energy cost per day of each MG: (a) Second strategy, (b) Third strategy and (c) Proposed strategy.
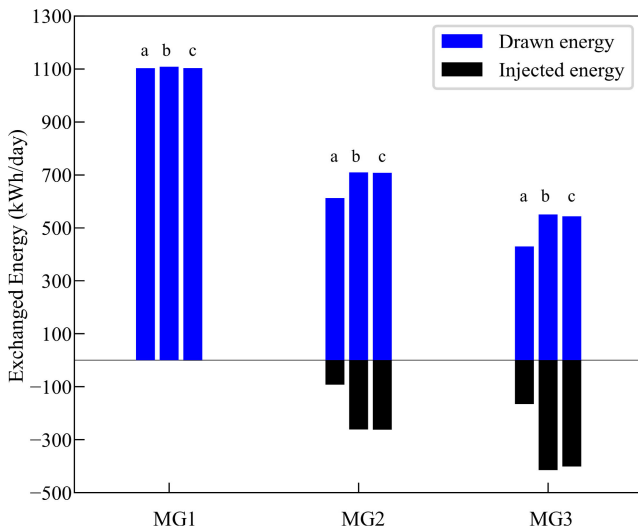

FIGURE 16. Drawn and injected energy per day of each MG; (a) Second strategy, (b) Third strategy and (c) Proposed strategy.

TABLE 6. The comparison of DSO costs per day in four strategies.

Strategies	Purchased Energy Cost ($)		Carbon Emission Cost ($)	Total Cost ($)	Decreased (%)
	Main Grid	Three MGs			
First	9,068.03	0.00	83.59	9,151.61	0.37
Second	9,032.32	3.50	83.23	9,119.05	0.01
Third	9,011.71	63.47	83.11	9,158.29	0.44
Proposed	9,011.07	23.77	83.12	9,117.96	-

The DNN surrogate-assisted rule-based control is considered in the first strategy. In addition, the DNN surrogate-assisted multi-agent DDPG algorithm is applied to solve
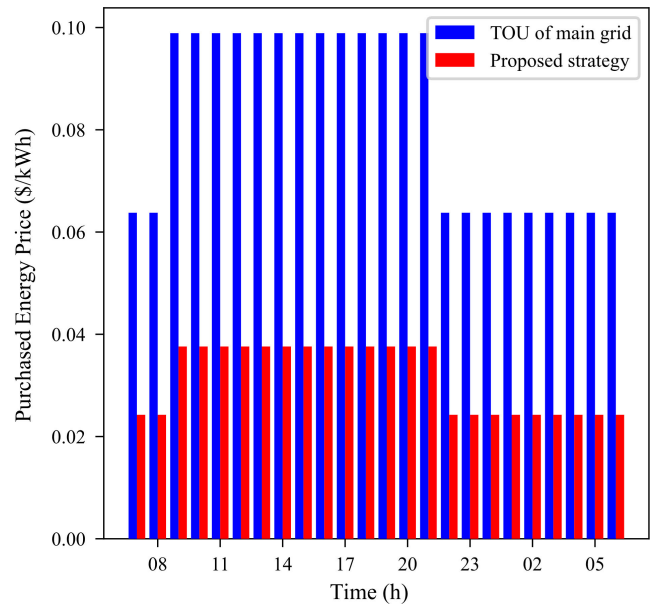
FIGURE 17. The purchased energy price for three MGOs.

the OEM problem in the last three strategies. Results in Table 6 demonstrate that the proposed strategy can reduce the overall cost of the DSO to about 9,117$ per day. The proposed strategy can motivate the MGO agent to inject power into the DN more than or equal to the other strategies under DN and MG operation constraints. Also, it can provide a minimum the total cost of the DSO compared with other strategies. In Fig. 14, three MGO agents under the proposed strategy can control the BESS to prevent transformer limit violation, over-limit of HST, and increased leaps of aging acceleration factor during certain hours. Furthermore, the BESS controlled by the MGO agent in the proposed strategy is managed and used correctly, as represented in Fig. 13, which causes a decrease in the BESS degradation cost. Therefore, the above results can confirm the proposed strategy as an efficient way for DSO to submit the purchased energy price to multi-MGOs, which can minimize the purchased energy and carbon emission costs of DSO.

VIII. CONCLUSION

This work presents the EMS frameworks for the DN with the multi-MGs framework to minimize the overall cost of the DSO. Two objective functions for DSO are minimizing the purchased energy and carbon emission costs. In contrast, the exchanged energy cost between MG and DN, carbon emission cost, and BESS degradation cost are modeled as MGO objective functions. The DNN surrogate models are developed to estimate the power system parameters of multi-MGs that reduces the computational burden considering the probabilistic power flow. The DDPG algorithm is guaranteed as a robust optimization approach and employed to improve the performance of OEM tasks within a DN with three MGs, considering the uncertainties of the solar PV and increased

EV demand integrated into each MG. The purchased energy price based on real-time pricing is proposed using DNN surrogate-assisted multi-agent DDPG algorithm to minimize the overall cost of the DSO. The OEM problem is mapped to the DDPG with the Markov decision process to create the environment for agent interaction. In addition, the DNN is developed as the DSO and MGO agents to interact with the environment, which leads to discovering the optimal solution for OEM solving. The BESS installed in each MG is properly controlled by the MGO agent to deal with solar PV and EV uncertainties and to minimize the overall cost of MGO. Moreover, three DNN surrogate models predict the mean of three MG power parameters to guarantee the decision of MGO agents instead of applying the probabilistic power flow using NT with Zhao's PEM in the actual system. Also, the DN operation constraints are confirmed using the deterministic power flow. The modified IEEE 33 bus distribution network is modeled as a distribution network, and a residential low-voltage distribution system at Udon Thani in Thailand is modeled as three MGs.

The numerical results show that the DNN surrogate model can provide an error for MG power parameters prediction lower than 1% compared with the results using the probabilistic power flow in the actual system. Moreover, the DNN surrogate-assisted DDPG algorithm can provide the optimal solution close to the solution obtained by the DE and decrease the computational burden by 89.23% compared with the DE. The DNN surrogate-assisted multi-agent DDPG algorithm proposed in this work can generate the optimal purchased energy price, control the BESS appropriately, and prevent transformer limit violation. Also, the proposed strategy can obtain the lowest cost of DSO by about 9,117$ per day and track MGO behavior to inject more energy into the DN and close to the energy provided using a higher purchase price. From the above results, the DNN surrogate-assisted multi-agent DDPG algorithm can decrease the computational convergence time for solving the optimal energy management task in the DN with multi-MGs.

Besides, wireless communication technologies and electric elements are continuously developed. This scenario leads to applying the real-time pricing strategy for managing energy within the power system, especially the DN with multi-MGs. Also, the OEM with the real-time pricing strategy needs to rely on the optimization approach that can provide the decision accurately and quickly. DNN surrogate-assisted multi-agent DDPG algorithm is one of the optimization approaches that have a good capability for improving the computational burden in OEM solving with a real-time pricing strategy. Furthermore, the well-trained weight set of all agents can be easily updated by adding modified power systems data to the environment, which they are learned more about through the DDPG training process. In contrast, the DNN surrogate model can still easily learn more to correct the changed behavior of MG through deep learning. Therefore, the DNN surrogate-assisted multi-agent DDPG algorithm can apply to future studies on optimal energy management tasks in a large and complex energy system.

REFERENCES

[1] S. Deb, K. Tammi, K. Kalita, and P. Mahanta, "Impact of electric vehicle charging station load on distribution network," *Energies*, vol. 11, no. 1, pp. 1–25, 2018, doi: 10.3390/en11010178.

[2] O. Beaude, S. Lasaulce, M. Hennebel, and J. Daafouz, "Minimizing the impact of EV charging on the electricity distribution network," in *Proc. Eur. Control Conf. (ECC)*, vol. 7, no. 6, Jul. 2015, pp. 648–653, doi: 10.1109/ECC.2015.7330615.

[3] R. Li, W. Wang, and M. Xia, "Cooperative planning of active distribution system with renewable energy sources and energy storage systems," *IEEE Access*, vol. 6, pp. 5916–5926, 2018, doi: 10.1109/ACCESS.2017.2785263.

[4] S. Hu, Y. Xiang, J. Liu, C. Gu, X. Zhang, Y. Tian, Z. Liu, and J. Xiong, "Agent-based coordinated operation strategy for active distribution network with distributed energy resources," *IEEE Trans. Ind. Appl.*, vol. 55, no. 4, pp. 3310–3320, Jul./Aug. 2019, doi: 10.1109/TIA.2019.2902110.

[5] *Department of Alternative Energy Development and Efficiency, Alternative Energy Development Plan*, document AEDP2018, Ministry Energy, Bangkok, Thailand, 2021. http://www.eppo.go.th/index.php/th/conservation/aedp

[6] Provincial Electricity Authority (PEA). (2020). *Provincial Electricity Authority Power Quality Standards B.E.2563 (2020)*. [Online]. Available: https://www.pea.co.th/Portals/0/Document/WorkStandard/standardmix_2564_1.pdf

[7] A. Woyte, V. Van Thong, R. Belmans, and J. Nijs, "Voltage fluctuations on distribution level introduced by photovoltaic systems," *IEEE Trans. Energy Convers.*, vol. 21, no. 1, pp. 202–209, Mar. 2006, doi: 10.1109/TEC.2005.845454.

[8] A. Hirsch, Y. Parag, and J. Guerrero, "Microgrids: A review of technologies, key drivers, and outstanding issues," *Renew. Sustain. Energy Rev.*, vol. 90, pp. 402–411, Jul. 2018, doi: 10.1016/j.rser.2018.03.040.

[9] B. Zhou, J. Zou, C. Yung Chung, H. Wang, N. Liu, N. Voropai, and D. Xu, "Multi-microgrid energy management systems: Architecture, communication, and scheduling strategies," *J. Modern Power Syst. Clean Energy*, vol. 9, no. 3, pp. 463–476, 2021, doi: 10.35833/MPCE.2019.000237.

[10] V. Bui, A. Hussain, and H.-M. Kim, "A multiagent-based hierarchical energy management strategy for multi-microgrids considering adjustable power and demand response," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1323–1333, Mar. 2018, doi: 10.1109/TSG.2016.2585671.

[11] A. N. Toutounchi, S. Seyedshenava, J. Contreras, and A. Akbarimajd, "A stochastic bilevel model to manage active distribution networks with multi-microgrids," *IEEE Syst. J.*, vol. 13, no. 4, pp. 4190–4199, Dec. 2019, doi: 10.1109/JSYST.2018.2890062.

[12] M. Haseeb, S. A. A. Kazmi, M. M. Malik, S. Ali, S. B. A. Bukhari, and D. R. Shin, "Multi objective based framework for energy management of smart micro-grid," *IEEE Access*, vol. 8, pp. 220302–220319, 2020, doi: 10.1109/ACCESS.2020.3041473.

[13] F. Khavari, A. Badri, and A. Zangeneh, "Energy management in multi-microgrids considering point of common coupling constraint," *Int. J. Electr. Power Energy Syst.*, vol. 115, Feb. 2020, Art. no. 105465, doi: 10.1016/j.ijepes.2019.105465.

[14] B. Zhang, Q. Li, L. Wang, and W. Feng, "Robust optimization for energy transactions in multi-microgrids under uncertainty," *Appl. Energy*, vol. 217, pp. 346–360, May 2018, doi: 10.1016/j.apenergy.2018.02.121.

[15] M. Marzband, M. Javadi, S. A. Pourmousavi, and G. Lightbody, "An advanced retail electricity market for active distribution systems and home microgrid interoperability based on game theory," *Electr. Power Syst. Res.*, vol. 157, pp. 187–199, Apr. 2018, doi: 10.1016/j.epsr.2017.12.024.

[16] F. H. Aghdam, N. T. Kalantari, and B. Mohammadi-Ivatloo, "A chance-constrained energy management in multi-microgrid systems considering degradation cost of energy storage elements," *J. Energy Storage*, vol. 29, Jun. 2020, Art. no. 101416, doi: 10.1016/j.est.2020.101416.

[17] H. Nagpal, I.-I. Avramidis, F. Capitanescu, and P. Heiselberg, "Optimal energy management in smart sustainable buildings—A chance-constrained model predictive control approach," *Energy Buildings*, vol. 248, Oct. 2021, Art. no. 111163, doi: 10.1016/j.enbuild.2021.111163.

[18] Y. Liu, L. Guo, and C. Wang, "A robust operation-based scheduling optimization for smart distribution networks with multi-microgrids," *Appl. Energy*, vol. 228, pp. 130–140, Oct. 2018, doi: 10.1016/j.apenergy.2018.04.087.

[19] Y. Du, Z. Wang, G. Liu, X. Chen, H. Yuan, Y. Wei, and F. Li, "A cooperative game approach for coordinating multi-microgrid operation within distribution systems," *Appl. Energy*, vol. 222, pp. 383–395, Jul. 2018, doi: 10.1016/j.apenergy.2018.03.086.

[20] M. J. Morshed, J. B. Hmida, and A. Fekih, "A probabilistic multi-objective approach for power flow optimization in hybrid wind-PV-PEV systems," *Appl. Energy*, vol. 211, pp. 1136–1149, Feb. 2018, doi: 10.1016/j.apenergy.2017.11.101.

[21] M. Roustaee and A. Kazemi, "Multi-objective stochastic operation of multi-microgrids constrained to system reliability and clean energy based on energy management system," *Electr. Power Syst. Res.*, vol. 194, May 2021, Art. no. 106970, doi: 10.1016/j.epsr.2020.106970.

[22] Z. Ullah, S. Wang, J. Radosavljević, and J. Lai, "A solution to the optimal power flow problem considering WT and PV generation," *IEEE Access*, vol. 7, pp. 46763–46772, 2019, doi: 10.1109/ACCESS.2019.2909561.

[23] C. Srithapon, P. Fuangfoo, P. K. Ghosh, A. Siritaratiwat, and R. Chatthaworn, "Surrogate-assisted multi-objective probabilistic optimal power flow for distribution network with photovoltaic generation and electric vehicles," *IEEE Access*, vol. 9, pp. 34395–34414, 2021, doi: 10.1109/ACCESS.2021.3061471.

[24] C. Gai, Y. Chang, T. Xu, and C. Li, "Optimization of emergency load-shedding based on surrogate-assisted differential evolution," in *Proc. IEEE/IAS Ind. Commercial Power Syst. Asia (I&CPS Asia)*, Jul. 2021, pp. 868–873, doi: 10.1109/ICPSAsia52756.2021.9621625.

[25] T. Su, Y. Liu, J. Zhao, and J. Liu, "Deep belief network enabled surrogate modeling for fast preventive control of power system transient stability," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 315–326, Jan. 2022.

[26] T. Bogaraj and J. Kanakaraj, "Intelligent energy management control for independent microgrid," *Sādhanā*, vol. 41, no. 7, pp. 755–769, Jul. 2016, doi: 10.1007/s12046-016-0515-6.

[27] D. Morin, Y. Stevenin, C. Grolleau, and P. Brault, "Evaluation of performance improvement by model predictive control in a renewable energy system with hydrogen storage," *Int. J. Hydrogen Energy*, vol. 43, no. 45, pp. 21017–21029, Nov. 2018, doi: 10.1016/j.ijhydene.2018.09.118.

[28] T. M. Kneiske, M. Braun, and D. I. Hidalgo-Rodriguez, "A new combined control algorithm for PV-CHP hybrid systems," *Appl. Energy*, vol. 210, pp. 964–973, Jan. 2018, doi: 10.1016/j.apenergy.2017.06.047.

[29] H. Hu, S. S. Yu, J. Zhao, T. Chau, F. Ding, T. Fernando, and H. Trinh, "MPC-based double-layer real-time conditional cSelf-restoration for interconnected microgrids," *Int. J. Electr. Power Energy Syst.*, vol. 129, Jul. 2021, Art. no. 106745, doi: 10.1016/j.ijepes.2020.106745.

[30] A. Mansour-Saatloo, Y. Pezhmani, M. A. Mirzaei, B. Mohammadi-Ivatloo, K. Zare, M. Marzband, and A. Anvari-Moghaddam, "Robust decentralized optimization of multi-microgrids integrated with power-to-X technologies," *Appl. Energy*, vol. 304, Dec. 2021, Art. no. 117635, doi: 10.1016/j.apenergy.2021.117635.

[31] W. Jiang, K. Yang, J. Yang, N. Xue, and Z. Zhuo, "Energy management strategy for maximization of renewable energy consumption in multi-microgrids," in *Proc. 6th Int. Conf. Syst. Informat. (ICSAI)*, Nov. 2019, pp. 325–329, doi: 10.1109/ICSAI48974.2019.9010441.

[32] V. I. Herrera, A. Milo, H. Gaztañaga, I. Etxeberria-Otadui, I. Villarreal, and H. Camblong, "Adaptive energy management strategy and optimal sizing applied on a battery-supercapacitor based tramway," *Appl. Energy*, vol. 169, pp. 831–845, May 2016, doi: 10.1016/j.apenergy.2016.02.079.

[33] T. Lv and Q. Ai, "Interactive energy management of networked microgrids-based active distribution system considering large-scale integration of renewable energy resources," *Appl. Energy*, vol. 163, pp. 408–422, Feb. 2016, doi: 10.1016/j.apenergy.2015.10.179.

[34] N. Nikmehr, S. Najafi-Ravadanegh, and A. Khodaei, "Probabilistic optimal scheduling of networked microgrids considering time-based demand response programs under uncertainty," *Appl. Energy*, vol. 198, pp. 267–279, Jul. 2017, doi: 10.1016/j.apenergy.2017.04.071.

[35] B. Tan and H. Chen, "Multi-objective energy management of multiple microgrids under random electric vehicle charging," *Energy*, vol. 208, Oct. 2020, Art. no. 118360, doi: 10.1016/j.energy.2020.118360.

[36] L. Yin and S. Li, "Hybrid Metaheuristic multi-layer reinforcement learning approach for two-level energy management strategy framework of multi-microgrid systems," *Eng. Appl. Artif. Intell.*, vol. 104, Sep. 2021, Art. no. 104326, doi: 10.1016/j.engappai.2021.104326.

[37] E. Foruzan, L.-K. Soh, and S. Asgarpoor, "Reinforcement learning approach for optimal distributed energy management in a microgrid," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5749–5758, Sep. 2018, doi: 10.1109/TPWRS.2018.2823641.

[38] Y. Du and F. Li, "Intelligent multi-microgrid energy management based on deep neural network and model-free reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1066–1076, Mar. 2020, doi: 10.1109/TSG.2019.2930299.

[39] C. Guo, X. Wang, Y. Zheng, and F. Zhang, "Optimal energy management of multi-microgrids connected to distribution system based on deep reinforcement learning," *Int. J. Electr. Power Energy Syst.*, vol. 131, Oct. 2021, Art. no. 107048, doi: 10.1016/j.ijepes.2021.107048.

[40] D. Zhang and B. Wei, "Comparison between differential evolution and particle swarm optimization algorithms," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Aug. 2014, pp. 239–244, doi: 10.1109/ICMA.2014.6885702.

[41] A. Al-Dujaili, M. R. Tanweer, and S. Suresh, "DE vs. PSO: A performance assessment for expensive problems," in *Proc. IEEE Symp. Ser. Comput. Intell.*, Dec. 2015, pp. 1711–1718, doi: 10.1109/SSCI.2015.240.

[42] A. Deb, B. Gupta, and J. S. Roy, "Performance comparison of differential evolution, genetic algorithm and particle swarm optimization in impedance matching of aperture coupled microstrip antennas," in *Proc. 11th Medit. Microw. Symp. (MMS)*, Sep. 2011, pp. 17–20, doi: 10.1109/MMS.2011.6068519.

[43] M. O. Akinsolu, B. Liu, P. I. Lazaridis, K. K. Mistry, M. E. Mognaschi, P. D. Barba, and Z. D. Zaharis, "Efficient design optimization of high-performance MEMS based on a surrogate-assisted self-adaptive differential evolution," *IEEE Access*, vol. 8, pp. 80256–80268, 2020, doi: 10.1109/ACCESS.2020.2990455.

[44] X. Hu and T. Liu, "Co-optimisation for distribution networks with multi-microgrids based on a two-stage optimisation model with dynamic electricity pricing," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 9, pp. 2251–2259, Jun. 2017, doi: 10.1049/iet-gtd.2016.1602.

[45] H. R. Baghaee, M. Mirsalim, G. B. Gharehpetian, and H. A. Talebi, "Application of RBF neural networks and unscented transformation in probabilistic power-flow of microgrids including correlated wind/PV units and plug-in hybrid electric vehicles," *Simul. Model. Pract. Theory*, vol. 72, pp. 51–68, Mar. 2017, doi: 10.1016/j.simpat.2016.12.006.

[46] B. R. Prusty and D. Jena, "Combined cumulant and Gaussian mixture approximation for correlated probabilistic load flow studies: A new approach," *CSEE J. Power Energy Syst.*, vol. 2, no. 2, pp. 71–78, Jun. 2016, doi: 10.17775/cseejpes.2016.00024.

[47] *National Household Travel Survey (NHTS)*, USA Dept. Transp., Washington, DC, USA, 2009.

[48] K. Qian, C. Zhou, M. Allan, and Y. Yuan, "Modeling of load demand due to EV battery charging in distribution systems," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 802–810, May 2011, doi: 10.1109/TPWRS.2010.2057456.

[49] C. Wu, F. Wen, Y. Y. Lou, and F. Xin, "Probabilistic load flow analysis of photovoltaic generation system with plug-in electric vehicles," *Int. J. Elect. Power Energy Syst.*, vol. 64, pp. 1221–1228, Jan. 2015, doi: 10.1016/j.ijepes.2014.09.014.

[50] C. Yan, W. Jinyu, and C. Shijie, "Probabilistic load flow calculation method based on polynomial normal transformation and Latin hypercube sampling," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 294–301, May 2013.

[51] Y. G. Zhao and T. Ono, "New point estimates for probability moments," *J. Eng. Mech.*, vol. 126, no. 4, pp. 1998–2001, 2000.

[52] C. Wang and C.-Y. Hsu, "Rankings correlation study," in *Proc. 5th IEEE Int. Conf. Big Data Anal.*, 2020, pp. 6–10, doi: 10.1109/ICBDA49040.2020.9101286.

[53] H. Li, Z. Lü, and X. Yuan, "Nataf transformation based point estimate method," *Sci. Bull.*, vol. 53, no. 17, pp. 2586–2592, Sep. 2008, doi: 10.1007/s11434-008-0351-0.

[54] Provincial Electricity Authority (PEA). (2018). *Electricity Tariffs for User*. Accessed: Mar. 25, 2022. [Online]. Available: https://www.pea.co.th

[55] Electricity Generating Authority of Thailand (EGAT). (2020). *Electricity Tariffs for PEA and MEA*. Accessed: Aug. 25, 2022. [Online]. Available: https://www.egat.co.th

[56] International Energy Agency (IEA). (2020). *Average CO₂ Emissions Intensity of Hourly Electricity Supply in the European Union, 2018 and 2040 by Scenario and Average Electricity Demand in 2018*. Accessed: Mar. 25, 2022. [Online]. Available: https://www.iea.org

[57] *Effective Carbon Rates 2021*, Effective Carbon Rates 2021, Organisation Econ. Co-Operation Develop. (OECD), Paris, France, 2021.

[58] J.-O. Lee and Y.-S. Kim, "Novel battery degradation cost formulation for optimal scheduling of battery energy storage systems," *Int. J. Electr. Power Energy Syst.*, vol. 137, May 2022, Art. no. 107795, doi: 10.1016/j.ijepes.2021.107795.

[59] *IEEE, IEEE Guide for Loading Mineral-Oil-Immersed Transformers and Step-Voltage Regulators Redline*, Standard C57.91-2011, 2011.

[60] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[61] H. Xiao, W. Pei, W. Deng, L. Kong, H. Sun, and C. Tang, "A comparative study of deep neural network and meta-model techniques in behavior learning of microgrids," *IEEE Access*, vol. 8, pp. 30104–30118, 2020, doi: 10.1109/ACCESS.2020.2972569.

[62] L. Thurner, A. Scheidler, F. Schafer, J.-H. Menke, J. Dollichon, F. Meier, S. Meinecke, and M. Braun, "Pandapower—An open-source Python tool for convenient modeling, analysis, and optimization of electric power systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6510–6521, Nov. 2018, doi: 10.1109/TPWRS.2018.2829021.

[63] Z. Jiang and G. Shen, "Prediction of house price based on the back propagation neural network in the keras deep learning framework," in *Proc. 6th Int. Conf. Syst. Informat. (ICSAI)*, Nov. 2019, pp. 1408–1412, doi: 10.1109/ICSAI48974.2019.9010071.

[64] Provincial Electricity Authority (PEA). *Load Research of PEA*. Accessed: Sep. 1, 2022. [Online]. Available: http://peaoc.pea.co.th/loadprofile/en/

[65] Thai Meteorological Department. *Hourly Temperature for Udon Thani, Thailand (2015–2017)*. Accessed: Sep. 1, 2022. [Online]. Available: https://www.tmd.go.th/en/

[66] Department of Alternative Energy Development and Efficiency. *Hourly Solar Irradiation for Udon Thani, Thailand (2015–2017)*. Accessed: Sep. 1, 2022. [Online]. Available: https://weben.dede.go.th/webmax/

[67] K. Lourenssen, J. Williams, F. Ahmadpour, R. Clemmer, and S. Tasnim, "Vanadium redox flow batteries: A comprehensive review," *J. Energy Storage*, vol. 25, Oct. 2019, Art. no. 100844, doi: 10.1016/j.est.2019.100844.

[68] M. R. Jannesar, A. Sedighi, M. Savaghebi, and J. M. Guerrero, "Optimal placement, sizing, and daily charge/discharge of battery energy storage in low voltage distribution network with high photovoltaic penetration," *Appl. Energy*, vol. 226, pp. 957–966, Jun. 2018, doi: 10.1016/j.apenergy.2018.06.036.

[69] Energy Policy and Planning Office (EPPO). (2022). *Review of the Purchase of Electricity From Incremental Renewable Energy Under Natural Gas Management (2022)*. Accessed: Apr. 28, 2022. [Online]. Available: http://www.eppo.go.th/index.php/en/

NIPHON KAEWDORNHAN received the B.Eng. degree (Hons.) in electrical engineering from Khon Kaen University, Khon Kaen, Thailand, in 2021, where he is currently pursuing the M.Eng. degree with the Department of Electrical Engineering. His research interests include power system analysis, microgrid, energy management systems, renewable energy resources, and machine learning for energy management optimization and prediction.

CHITCHAI SRITHAPON (Member, IEEE) received the B.Eng. and Ph.D. degrees in electrical engineering from Khon Kaen University, Khon Kaen, Thailand, in 2002 and 2021, respectively. From 2008 to 2021, he was an Electrical Engineer with the Provincial Electricity Authority (PEA), Thailand. He is currently a Postdoctoral Researcher with the Department of Electrical Engineering, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include smart microgrids, electric vehicles grid integration, energy management systems, optimization algorithm, and machine learning.

RONGRIT CHATTHAWORN received the B.Eng. degree (Hons.) in electrical engineering from Khon Kaen University, Thailand, in 2009, and the M.Eng. and Ph.D. degrees in electrical engineering from Chulalongkorn University, Thailand, in 2011 and 2015, respectively. From 2015 to 2017, he was a Researcher with the Energy Regulatory Commission (ERC), Thailand. Since 2018, he has been an Assistant Professor with the Department of Electrical Engineering, Khon Kaen University. His research interests include power system planning, power system reliability, smart grid, energy management systems, and renewable energy resources.

● ● ●