An Applied Deep Reinforcement Learning Approach to Control Active Networked Microgrids in Smart Cities with Multi-Level Participation of Battery Energy Storage System and Electric Vehicles

Sepehrzad, Reza ; Godazi Langeroudi, Amir Saman ; Khodadadi, Amin ; Adinehpour, Sara ; Al-Durra, Ahmed ; Anvari-Moghaddam, Amjad

Published in:
Sustainable Cities and Society

DOI (link to publication from Publisher):
10.1016/j.scs.2024.105352

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Sustainable Cities and Society 107 (2024) 105352

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

An applied deep reinforcement learning approach to control active networked microgrids in smart cities with multi-level participation of battery energy storage system and electric vehicles

Reza Sepehrzad [a,*], Amir Saman Godazi Langeroudi [a], Amin Khodadadi [b,c], Sara Adinehpour [b,c], Ahmed Al-Durra [d], Amjad Anvari-Moghaddam [e]

[a] Department of Electrical Engineering, Politecnico di Milano University, Milan, Italy
[b] Department of Electrical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
[c] Department of Electrical Engineering, Arman Niroo Hormozgan Company, Bandar Abbas, Iran
[d] Advanced Power & Energy Center, EECS Department, Khalifa University, Abu Dhabi, UAE
[e] Department of Energy (AAU Energy), Aalborg University, 9220 Aalborg, Denmark

ARTICLE INFO

Keywords:
Networked microgrid
Electric vehicles
Energy management
Deep reinforcement learning algorithm

ABSTRACT

This study proposed an intelligent energy management strategy for islanded networked microgrids (NMGs) in smart cities considering the renewable energy sources uncertainties and power fluctuations. Energy management of active power and frequency control approach is based on the intelligent probabilistic wavelet fuzzy neural network-deep reinforcement learning algorithm (IPWFNN-DRLA). The control strategy is formulated with deep reinforcement learning approach based on the Markov decision process and solved by the soft actor-critic algorithm. The NMG local controller (NMGLC) provides information such as the frequency, active power, power generation data, and status of the electric vehicle's battery energy storage system to the NMG central controller (NMGCC). Then the NMGCC calculates active power and frequency support based on the IPWFNN-DRLA approach and sends the results to the NMGLC. The proposed model is developed in a continuous problem-solving space with two structures of offline training and decentralized distributed operation. For this purpose, each NMG has a control agent (NMGCA) based on the IPWFNN algorithm, and the NMGCA learning model is formulated based the online back-propagation learning algorithm. The proposed approach demonstrates a computation accuracy exceeding 98%, along with a 7.82% reduction in computational burden and a 61.1% reduction in computation time compared to alternative methods.

1. Introduction

In recent years, on the one hand, changing the framework of traditional electrical networks and on the other hand, various challenges such as environmental concerns caused by the increase in fossil fuel consumption, the availability of electrical energy, increasing the reliability of the electrical network and improving the security indicators of power systems caused researchers to present the microgrid (MG) idea (Huang et al., 2023). The MG idea has made significant progress with the expansion of the power electronics and energy storage industry (Sepehrzad et al., 2023). Because MG has the ability to operate in two grid-connected and islanded modes, it has created a suitable platform for the presence of renewable energy sources (RES) and energy storage system (ESS) in power systems (Lovering, 2023), (Hassanzadeh et al., 2023). The presence of the RES and EES in the MG structure and topology has posed the current power systems with numerous technical, economic and environmental opportunities and challenges (Parast et al., 2023). The ESSs and especially the battery energy storage system (BESS) are not only considered as one of the most important supplements of RES but BESS units are considered one of the most important components of electric vehicles (EVs) (Abid et al., 2024), (Ajagekar, Decardi-Nelson & You, 2024).

The increased use of EVs together with the random behavior of EV owners, has also posed serious technical challenges to power system operators, especially in terms of frequency regulation (Sepehrzad et al., 2024). Although the presence of EV has created serious challenges for power systems, this challenge can be changed into an opportunity to

* Corresponding author.
E-mail address: sepehrzad.reza@gmail.com (R. Sepehrzad).

https://doi.org/10.1016/j.scs.2024.105352
Received 23 November 2023; Received in revised form 11 February 2024; Accepted 16 March 2024
Available online 18 March 2024

Nomenclature

Variables and parameters

$PDF(P_{solar}), PDF(f_{EV}^{AT}), PDF(f_{EV}^{DT}), PDF(f_D)$ probability distribution function of different parameters
$\sigma_{solar}, \mu_{solar}$, $\sigma_{EV}^{AT}, \mu_{EV}^{AT}$, $\sigma_{EV}^{DT}, \mu_{EV}^{DT}$ mean and standard deviation of various parameters
P_{solar}, P_{PV} solar power, PV power (kW)
$P_{BESS-EV}(t)$ the EV battery power (kW)
$P_{BESS-EV}(t), P_{BESS-EV}^{\min}, P_{BESS-EV}^{\max}$ the EV battery power, The minimum EV battery power, The maximum power of EV battery (kW)
$P_{BESS}^{ref}, P_{BESS,t}^{i,ch}, P_{BESS,t}^{i,dis}$ the BESS reference power, The charging and discharging power of the BESS unit in each time slot (kW)
$\Delta P_{SG}^{i,out}, \Delta P_{gov}^{i}, \Delta P_{PV}^{i}, \Delta P_{tie}^{i}, \Delta P_{Load}^{i}, \Delta P_{BESS}^{i}, \Delta P_{CS}^{i}$ deviation of synchronous generator output power, Deviation of the governor position, Deviation of the PV power, The tie-line power changes, Load power changes, The BESS power changes, The control signals power to be obtained by the controller (kW)
P_{ref}, P_{OP} reference active power, operation power (kW)
$P_{\text{Pre}}^{ref}, P_{\text{Pre}}^{OP}$ predicted the reference active power and operation power (kW)
$E_{BESS,t}^{i,rate}, E_{BESS,t}^{i}$ the BESS rated energy in each time slot, BESS energy in each time slot (kWh)
$SOC(P_{BESS-EV})$ the EV's battery state of charge (%)
$SOH_{BESS,t}$ the BESS state of health in each time slot (%)
$\Delta SOH_{BESS,t}, \Delta SOC_{BESS,t}$ the BESS state of charge and state of health changes (%)
f_{ref}, f_{OP} reference frequency, the operation frequency (Hz)
$\Delta\omega_{MG,i,j}$ the $MG_{i,j}$ frequency deviation (Rad/s)
N the PV panel modules
d the variance of the distribution (km)
c_1, c_2 shape and scale factors
k_v, k_i voltage and current gain
k_{EV}^{AT}, k_{EV}^{DT} standard deviation
$T_{PV}, T_{Amb}, T_{norm}$ the PV panel surface temperature, Ambient temperature, Normal temperature (C^o)
V_{MPP}, V_{OC} maximum power point voltage and PV open circuit voltage (V)
$V_{MG,i,j}$ the $MG_{i,j}$ main bus voltage (V)
I_{MPP}, I_{SC} maximum power point current and PV short circuit current (A)
$i_{q-P}^{OP}, i_{d-Q}^{OP}$ converters operation signals (A)
$D_{BESS-EV}^{\max}, D_{BESS-EV}$ maximum distance traveled by EV, The distance traveled by EV (km)
$b_{BESS-EV}^{ch}, b_{BESS-EV}^{dis}$ binary variables of the EV battery charge and discharge
t_{ch} charging time (h)
$T_{SG,t}^{i}, T_{gov}^{i}, T_{MG,ij}$ governor and turbine time constant, synchronous coefficient between $MG_{i,j}$
Δt control action time-interval (s)
η_{ch}, η_{dis} the BESS charging and discharging efficiency (%)
$D.F_{BESS,t}$ the BESS degradation factor (kWh)
$d_{BESS,1,2}$ the BESS degradation deviation (kWh)
H_i, D_i the SG inertia, The SG damping coefficient
SD_i speed drop
$\delta_{MG,i,j}^{0}$ the $MG_{i,j}$ main bus voltage angle
$X_{MG,ij}$ tie-line reactance (Ω)
ρ the BESS power change index for optimizing (kW)
χ the BESS performance uniformity index
$\beta_{1,2}$ the BESS degradation coefficients (kWh)
W^{ω}, W^{DG} weight coefficients of DRL's reward
S_d^{ω}, S_d^{DG} frequency and DG state
$e_1(k), e_2(k)$ error of reference values from operation values in each iteration k, The Prediction error from actual values in each iteration k
I_i^L, O_i^L input / output layers
L layer symbol
$NN_i^L(k)$ neural network of each layer
f_i^L function of each layer
M_j^L mean index of asymmetric Gaussian function
t_N, t_D demand power and PV unit power in different parts of each MG
$D_{r,l-j}^{L}$ asymmetric Gaussian function Right / left deviation
$\lambda_q(k)$ output summation of Wavelet function
M_{iq}^L, D_{iq}^L mean and deviation of the Wavelet function
w^L normalized weight of the membership functions
E_L racking error in each layer
σ^L each layer's error index based propagated process
μ each layer's learning rate
Δw^L each layer weight deviation
w_Γ DRL entropy weight
θ^t DRL discount factor
φ critic network parameters
$\mu_{DM_{CP}}, \sigma_{DM_{CP}}$ decision-making mean and covariance guarantee index of the IPWFNN algorithm
ζ_t random variables of normal probability distribution function
$J_Z(\varphi)$ critic network function
a^i, b^i, c^i the SG generation cost coefficients ($/kWh)

Abbreviations

AGF asymmetric Gaussian function
ANN artificial neural network
BESS battery energy storage system
DRL deep reinforcement learning
EV electric vehicle
IPWFNN-DRLA intelligent probabilistic wavelet fuzzy neural network-deep reinforcement learning algorithm
MASAC multi-agent soft actor-critic
MDP Markov decision process
NMGLC NMG local controller
NMGCC NMG central controller
NMGCA NMG control agent
NMG networked microgrid
RES renewable energy sources
OBPL online back-propagation learning
SAC soft actor-critic
SOC state of charge
SOH state of health
SG synchronous generator

control the vital indicators of MGs such as frequency (Mobarakeh et al., 2023). Implementation, optimal operation and intelligent development of MGs are the most important challenges of this idea. Therefore, networked microgrids (NMG) are considered a desirable framework to satisfy the objectives and challenges of power systems (Zhou et al., 2023). Control hierarchy in MGs are generally divided into three categories, which are primary, secondary and tertiary. Each of these levels are presented and developed in a specific time frame (Kumar & Karthikeyan, 2024). For example, primary control actions are defined in the time frame of milliseconds and seconds, and secondary actions are defined in the time frame of minutes and hours, while tertiary control can be defined in daily, weekly and even yearly time frames (Bustos et al., 2023), (Reza Sepehrzad et al., 2022). The most important primary goals include controlling vital MGs indicators such as voltage and frequency (Zhao, Wang & Guo, 2023), (Rosero, Díaz & Trujillo, 2021). The frequency index of the power system changes and fluctuates according to the changes in active power. As mentioned, RES units are widely used in the MGs structure. Because the performance of the RES depends on weather conditions, the power generated by these units has a completely probability pattern (Wang et al., 2022), (Eid, Mohammed & El-Kishky, 2022). Changes in the power generated by RES influence the demand-supply balance and the MG may experience frequency fluctuations (Yang, Cui & Wang, 2023). Therefore, MG frequency control and support in the presence of RES is an important challenge.

In recent years, many research studies have been done in order to control the MG frequency in the presence of RES units. The presented works are generally divided into two categories: conventional control models (Mukhopadhyay & Das, 2021) and advanced models (Sepehrzad et al., 2021b). The most important conventional control models are the droop control method (Ferahtia et al., 2022), virtual impedance control method (Babayomi, Li & Zhang, 2022) and the most important advanced models are meta-heuristic algorithms (Sharma et al., 2022), model predictive control (Yamashita et al., 2022), fuzzy algorithm (Sepehrzad et al., 2021a) and artificial neural networks (Almaleck et al., 2024). The most important challenge of advanced and conventional algorithms is their lack of efficiency when faced with diverse data, as well as the exponential increase in operational processing. In the last decade, intelligent algorithms based on artificial intelligence and advanced techniques have not only provided accurate models, but also improved the performance in different operating conditions. Control systems based on deep reinforcement learning (DRL) techniques are one of the successful models in providing accurate and fast control platforms in the MG structure (Onile et al., 2023). Intelligent algorithms based on machine learning, relying on a data driven structure, significantly increase the accuracy and speed of calculations. Evaluation and analysis of problems based on data diversity, such as energy distribution systems, significantly depends on data mining, data analysis and extraction of their features. Algorithms based on machine learning using a layered and data driven structure are a good answer to the challenge of data diversity in energy distribution systems with wide data diversity. Intelligent algorithms based on the DRL benefit from two important features. The most important feature of these algorithms is the structure and probability framework based on random variables. Therefore, the modeling of various uncertainties in power networks is easily done. On the other hand, DRL algorithms have significant efficiency in dealing with big data, as well as providing the best solutions for data mining and data feature extraction (Wu et al., 2020).

1.1. Literature review

To control the frequency index and improve transient stability in RES-based power systems, load frequency control (LFC) based on a data-driven model and considering the uncertainties of RES is presented in Yan and Xu (2018). In this study, the proposed algorithm based on the DRL method has been developed in the continuous action domain. The proposed strategy has provided a nonlinear strategy to control the frequency deviation and improve the dynamic response of the MG. The proposed algorithm includes offline optimization of LFC and DRL and online control in the policy network structure with the aim of removing accumulated noise. Although the results of this strategy and approach have significant efficiency in controlling the frequency index, the lack of probability loads modeling such as EVs is considered the most important challenge of this research. To develop the DRL model, in Yan and Xu (2020) the LFC model is presented considering the data-driven model in a multi-area power system and based on the multi-agent deep reinforcement learning (MA-DRL) algorithm in the continuous action domain. A nonlinear control strategy based on centralized learning and decentralized execution is developed. In this research, multi-agent deep deterministic policy gradient (DDPG) has been used to adjust the parameters of the control agent by considering the nonlinear behavior of power generation sources. The proposed strategy has been modeled in the New-England 39-bus network and the results of this strategy and approach show the effectiveness of the proposed strategy in facing the RES uncertainties and controlling load changes.

In Liu, Yao and Hu (2019), the intelligent coordinated automatic generation control (DIC-AGC) model is proposed for LFC and creating dynamic balance in a multi-area integrated energy system (IES). In order to realize LFC, dynamic balance and improve the performance of DIC-AGC, the evolutionary imitation curriculum multi-agent deep deterministic policy gradient (EIC-MADDPG) algorithm has been used. The proposed strategy is presented in a four-area structure based on considering the uncertainties of RES and controlling the probability loads, as well as central learning and decentralized operation. Although the proposed strategy has a significant performance in reducing the frequency deviation of multi-area networks, the lack of random loads modeling such as EVs have challenged the proposed strategy. In Cui and Zhang (2021) frequency optimal control model based on the DRL approach and search in nonlinear control policy space is presented. The proposed strategy is developed based on the Lyapunov function approach and the use of neural networks algorithm. In this study, increasing the flexibility of power systems by considering the RES uncertainty and improving the transient stability of power systems has been analyzed.

To manage the participation of EV and RES units in controlling the technical and economic indicators of power systems in Zhang et al. (2022), the strategy of minimizing the operation cost and satisfying the EV stations charging standards in the context of RES-based microgrids is presented. The optimization problem model is programmed by the DRL algorithm based on delayed deep deterministic policy gradient (TD3). The energy management and optimization problem are formulated according to the MDP. The results of this strategy and approach show the flexibility of the proposed strategy in reducing the operating costs of RES-based MGs and EV charging stations. In Dorokhova et al. (2021), the EV units charging control model in the MG platform based on the RES and considering demand response programs is presented. The mathematical model of the control strategy has been developed in two discrete and continuous analysis spaces based on the DRL algorithm. The proposed strategy performance compared to other methods such as rule-based control mode, deterministic optimization algorithms and model predictive control has better performance and efficiency.

Although advanced models are being developed and expanded, the most important challenge of these algorithms is the learning structure and improving the computation time of DRL algorithms. So far, many attempts have been made in providing different learning algorithms. For example, in Reza Sepehrzad et al. (2022), the DRL approach based on integrated Monte Carlo tree search and a deep neural network (DNN) is presented with the aim of improving the learning structure of the DRL algorithm and energy management of BESS units. The multi-level strategy of preventive maintenance activities has been developed by considering the optimization index of operating costs, and capacities and increasing the reliability index. The proposed learning model in the state and action space shows an increase in the learning rate and a

Table 1
Comparison of reviewed literature.

Reference	Non-RES	RES	ESS	Probability analysis	Stability analysis	Technical analysis	Computation time analysis	Computation burden analysis	Diverse data analysis	Probability loads (EV)	DRL algorithm
Ref (Yan & Xu, 2018)	*	*	*	–	–	*	–	–	*	–	*
Ref (Yan & Xu, 2020)	*	*	*	–	–	*	–	–	*	*	*
Ref (Liu et al., 2019)	*	*	*	*	–	*	*	*	–	–	*
Ref (Cui & Zhang, 2021)	*	*	*	*	*	*	–	–	*	*	*
Ref (Zhang et al., 2022)	*	*	*	*	–	*	*	–	–	*	*
Ref (Dorokhova et al., 2021)	*	*	*	*	–	*	*	*	*	–	*
Ref (Reza Sepehrzad et al., 2022)	*	*	*	*	–	*	*	*	–	–	*
Ref (Ali et al., 2020)	*	*	*	–	–	*	*	*	*	–	*
Ref (Haarnoja et al., 2018)	*	*	*	–	–	*	*	*	*	–	*
Ref (Ma et al., 2023)	*	*	*	*	–	*	*	*	*	–	*
Ref (Cui et al., 2022)	*	*	*	–	*	*	–	–	–	–	*
Ref (Li et al., 2022)	*	*	*	–	*	*	*	–	*	–	*
Ref (Khalid et al., 2022)	*	*	*	–	*	*	*	*	–	–	*
Ref (Yan et al., 2022)	*	*	*	–	*	*	*	*	–	–	*
Proposed Method	*	*	*	*	*	*	*	*	*	*	*

decrease in the loss function due to the deviation of the estimated values from the reference values. In Ali et al. (2020), a low-inertia MG secondary controller model based on RES is presented to control the voltage and frequency index based on the DRL algorithm. The proposed algorithm model is developed and modeled based on deep deterministic policy gradient (DDPG) structure. The proposed strategy has reduced the deviation of voltage and frequency from the reference values by using the probability algorithms and the DRL approach of the developed learning structure as well as the intelligent control model between RES and ESs units. The results of the proposed strategy have been compared with other methods such as the droop control model in addition to validation in different scenarios. The results of this strategy and approach show the effectiveness of the proposed strategy in facing various uncertainties.

In Haarnoja et al. (2018), the economic analysis and optimal control of BESS units to support the frequency of power systems and consider indicators such as the aging cost of charging and discharging cycles of BESS units, generation cost, and also unscheduled interchange price are presented. The proposed strategy is based on the DRL method and data-driven approach. The results of this strategy and approach show the effectiveness of the proposed strategy in the optimal control of operating costs of BESS units. The proposed strategy is presented without considering the uncertainty of the price index in different time frames. Also, in this research, the random behavior of RES units and their effect on the performance of the proposed strategy has not been investigated. In Ma, Hu and Hu (2023), the primary frequency control model based on the RL method and the developed RNN model are presented in the structure of grids based on RESs and power electronic converters. The proposed strategy provides network frequency support with a local controller approach. The results of this study show the effectiveness of the proposed strategy in restoring the frequency of networks based on RESs and power electronics. The results of this strategy and approach have only mentioned the recovery of the frequency in the primary. While it is necessary to compensate for the deviation of the operating frequency from the reference value by the secondary controller at the secondary level. Also, another challenge of this research is the lack of analysis of the random loads such as EVs, and its effect on the control of the power grid frequency.

The stability of electric power distribution systems connected to converters based on RESs is presented in Cui, Jiang and Zhang (2022). In this study, the frequency control and stability approach based on the RL method has been developed. The proposed structure has been validated in different operation scenarios. The results of this research have described the high stability rate of the proposed approach in different operation scenarios. The primary challenge in this research lies in the absence of a comprehensive assessment of computation time and burden, along with a lack of evaluation in the context of probability analysis. In Li, Yu and Zhang (2022), an evolutionary imitation curriculum multi-agent deep deterministic policy gradient (EIC-MADDPG) algorithm is presented as a deep reinforcement learning algorithm for the dynamic balance of multiple energy fluctuations and frequency regulation in a multi-region integrated energy system (IES). The results of this research are tested and validated in real networks. The most obvious challenge of this research is the lack of assessment of probability loads and probability assessment.

In Khalid et al. (2022), a deep reinforcement learning approach is presented to adjust the parameters of the proportional-integral-derivative (PID) controller with the aim of controlling frequency fluctuations and improving stability in electrical networks based on the RESs. The proposed approach shows high efficiency for handling non-linear decision variables. The results of this study are validated in different operation scenarios. The most obvious challenge of this research is the lack of evaluation of various data and the probability analysis. In Yan et al. (2022), secondary frequency control in islanded microgrids is presented based on quantum DRL method in IEEE13-Bus network in the presence of 4 synchronous generators where the approach has shown high stability in the face of load fluctuations. Table 1 illustrates a comparative analysis of the proposed approach in this study with previous research studies, highlighting various aspects.

1.2. Problem statement, challenges and highlights

In this research, the intelligent control of NMG frequency index in smart cities is presented in the context of primary control considering the uncertainty of power generation of PV units and loads behaviors (i.e., consumption profiles of EVs). In this regard, the control model based on intelligent probabilistic wavelet fuzzy neural network-deep reinforcement learning algorithm (IPWFNN-DRLA) is developed and presented.

Traditional and non-intelligent control methods not only do not have enough accuracy in controlling power grid indicators but also have a low speed in responding to power grid changes. In traditional networks, the structure of frequency control is simpler than MGs, which have the ability to operate in two modes, islanded and grid-connected. The noteworthy point in MG operation is that each operation mode has its own operational requirements. For example, when the MG is connected to the main grid, two hypotheses are stated to implement power flow equations. The first hypothesis is when the MG receives power from the main grid. Therefore, in this case, the MG connection bus to the main grid is considered a P-Q bus. The second hypothesis is when the MG has the ability to inject power into the main grid. In this case, the MG is considered a P-V bus in the power flow equations. In addition, it should be mentioned that according to the two stated hypotheses and the operating mode of the grid-connected, the power of the MG will be dependent. In this operating mode, the task of controlling the power and frequency of the MG is performed by the slack bus of the main grid. But when the MG is operated in islanded mode, the internal power generation units must meet the voltage and frequency security constraints. Therefore, islanded and grid-connected modes have a significant impact on MG voltage and frequency, and each operation mode requires the implementation of a specific control algorithm.

Another noteworthy point is that in recent years, the use of RES and energy storage have grown significantly. Extracted power from the RES depends heavily on meteorological conditions thus exhibiting stochastic behaviors which in turn increases the uncertainty levels in operation management of MGs. The increase of uncertainty levels not only causes disturbances in the voltage and frequency index, but also causes the complexity of different MG controllers. On the other hand, the use of EVs has grown significantly in recent years. The battery of EVs provides the possibility for the subscribers that the subscribers based on the demand response programs are also introduced as one of the actors of the optimal power distribution system.

The power consumption pattern and demand of EVs are probabilistic and based on the behavior of subscribers. Therefore, the uncertainty caused by the power of EVs affects the operation of MGs. Another noteworthy and important point is that when the battery is in charging mode, it receives power from the MG. In this case, EV batteries are considered as loads in power flow equations. But when the batteries have the ability to inject power into the MG, then the battery unit is considered a power supplier in the power flow equations. Therefore, the use of EVs complicates the problem of power distribution and also the control of vital MG indicators such as frequency.

Different equipment with linear, non-linear, random, and probability features are used in the MG structure. Different equipment with different characteristics causes the continuous change of the MG voltage and frequency indicators. Therefore, providing a robust, reliable, flexible, and fast control model is considered a requirement of MGs. Therefore, in this study, the effect and the role of BESS units in frequency control and support based on random behavior of the RES and load is presented with a deep reinforcement learning approach in order to accurately and quickly control the MG frequency.

Therefore, it can be mentioned that uncertainties cause power changes in the structure of power grids, and therefore the voltage and frequency indicators of the power grid experience many changes and cause the power and energy distribution networks instability. Therefore, the provision of intelligent controllers based on uncertainties is considered one of the most important pillars of MG operation.

The **most important highlights** of the IPWFNN algorithm model in this research are:

1. The use of probability structure in the classification and feature extraction of data based on probability distribution functions. Various decision-making variables such as power generation and consumption indicators based on RESs and electric vehicles (EVs) are modeled, categorized, and feature extracted using probability distribution functions. Therefore, the classified results are provided to the proposed algorithm, and the proposed algorithm no longer needs to spend time classifying and extracting data features. Then the time and volume of calculations due to the probability data will bring a significant reduction.
2. Because the control and analysis of the frequency index in power networks is done in milliseconds and seconds, therefore the computation time and burden are of great importance. Therefore, to increase flexibility, reduce the computation time and burden, various arrangements have been considered, which include:
   - The control algorithm execution at both local and central controller levels.
   - The MG information is updated in 5 ms interval.
   - Presenting the DRL approach which is formulated based on the Markov decision process (MDP) and solved by the soft actor-critic (SAC) algorithm. Therefore, the proposed model is developed in the continuous solution space with central training and distributed operation structure. The proposed structure not only reduces the computation time but also the computation burden. This condition is also due to the fact that the implementation of repetitive and unnecessary algorithms has been prevented in the structure of central training and distributed operation.
   - Using the online back-propagation learning (OBPL) algorithm increases the proposed model's flexibility.

1.3. Innovation and main contribution

In this study, the intelligent frequency control and support model in the NMG and smart cities structure, in presence of RES and BESS units (electric vehicles) based on IPWFNN-DRLA, is presented. The proposed strategy is modeled based on the behavior and probability pattern of PV and EVs. The probability pattern of PV panel units and EVs are modeled by the Beta and Weibull probability distribution functions. Considering that the MG topology is based on the NMG structure, the connection between different MGs is modeled by the NMGCC and the NMGLC in the context of the communication system. The NMGLC sends information such as voltage, frequency, active power, reactive power, information on power generation sources and the status of BESS units in specific time periods (5 ms) to the NMGCC. The NMGCC implements NMG frequency support and control algorithms based on the IPWFNN-DRLA approach and sends the results to the NMGLC. The NMGCC model based on DRL and MDP approach is formulated and solved by the SAC algorithm.

The proposed strategy is developed in two structures of centralized training and decentralized operation. For this purpose, each MG has an NMGCA based on the IPWFNN algorithm. The learning model of the IPWFNN algorithm is based on the OBPL algorithm. The proposed learning algorithm is not only used in the deep learning algorithm of artificial neural networks with more than one hidden layer to calculate the weight gradient more accurately, but this method often provides the decreasing gradient calculation of the objective function by optimizing the learning algorithm and stabilizing the weight of the neurons. In addition, the proposed algorithm is also used for feed-forward neural

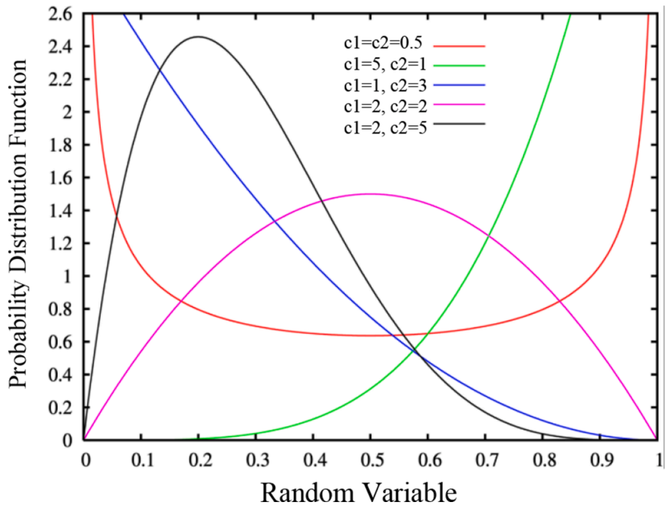

Fig. 1. The Beta probability distribution function based on $0.5 < c_1, c_2 < 5$ changes.

networks that require supervised learning. In the proposed algorithm, reducing the computation time and improving the dynamic response of NMG along with MG's frequency support in the presence of PV panel units and BESS units have been developed.

The **main contribution** of the study can be summarized as follows:

1. Non-linear, multi-level and intelligent frequency control in NMG structure with developed central learning approach and decentralized operation based on IPWFNN-DRLA method.
2. Improving the dynamic response of the proposed controller based on the OBPL learning algorithm with the aim of reducing the gradient calculations of the objective function and also reducing the loss function of the learning algorithm in the DRL structure with a supervision-based learning approach.
3. Improving the frequency response and support of NMG by considering the RES uncertainties and the EV units participation (BESS) in power and management control and reducing the frequency deviation from the reference values as well as increasing the MG's subscribers participation in the NMG power management and control structure.

The rest of the article is: In Section 2, the modeling of NMG structure and the probability model of RES and EV units are presented. In Section 3, the control algorithm is presented along with the structure of the central and local controller. In Section 4, implementation requirements, simulation framework and modeling results are presented. Finally, in Section 5, the conclusion along with the challenges of this research are stated.

2. NMG modeling and structure

2.1. RES uncertainty model

Basically, the behavior of EV and RES units depend on the random consumption pattern of subscribers and weather conditions, respectively. The consumption pattern of subscribers and weather conditions are also subject to probability and random variables. In order to describe the behavior of EVs and RESs, it is necessary to use probability models.

Because the description of the EV and RES unit behavior is based on probabilistic data analysis, therefore, data modeling based on probability distribution functions is highly capable. The use of probability distribution functions depends on the type of data, data recording periods, data volume and data sampling space (continuous or discrete space). In this research, the Weibull and Beta probability distribution function model and generalized models of the expressed functions have been used in modeling the behavior of EVs and RESs units. The proposed strategy does not show the output power of PV units and EV power, but only provides behavior analysis in order to estimate the power of these two units. To model the power generation of PVs, the Beta probability distribution function model is used (Mobarakeh et al., 2023).

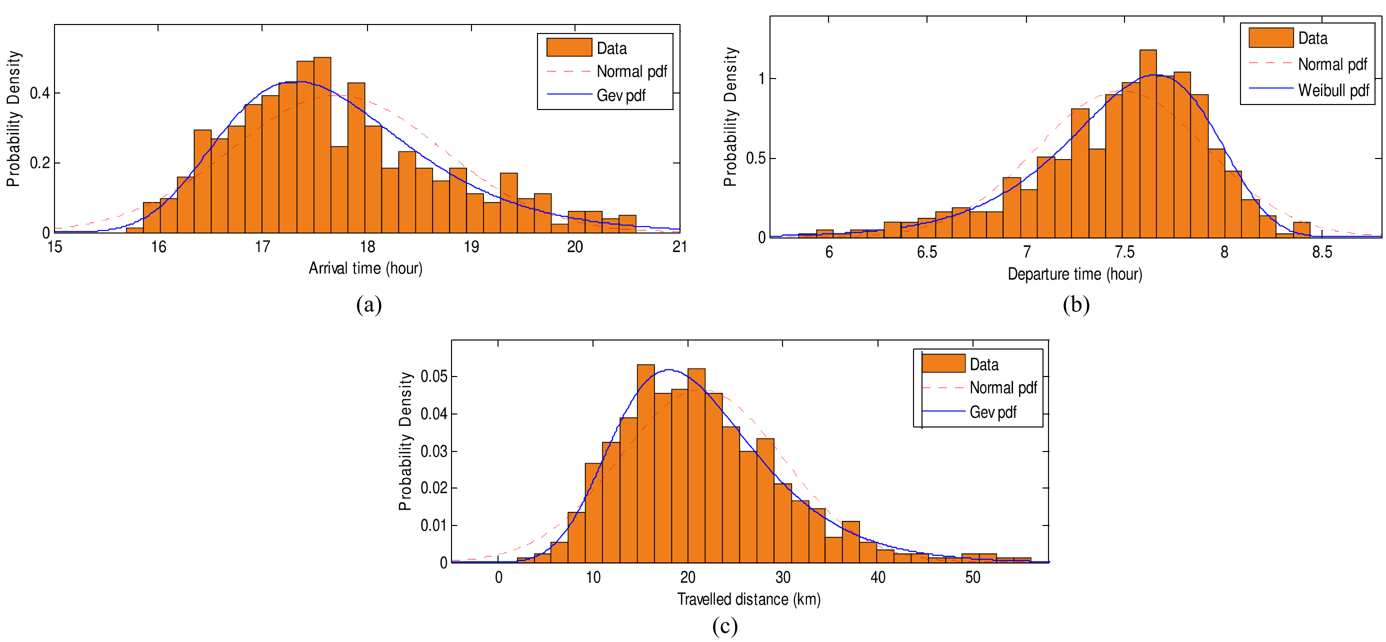

Fig. 2. Probability behavior model of EV units in different categories.

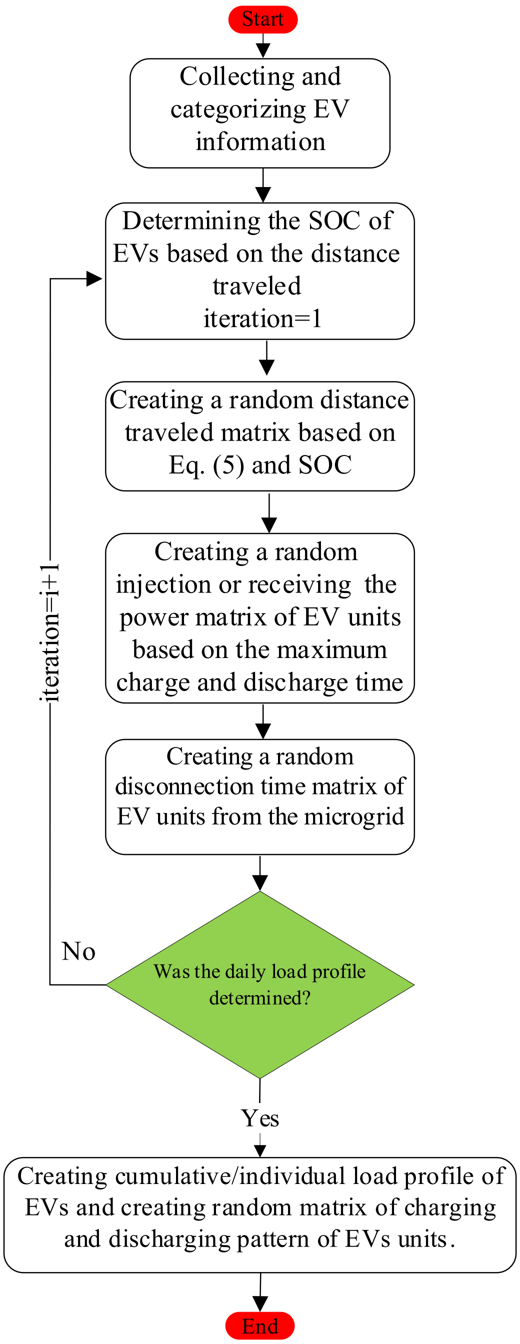

Fig. 3. Determination of EV units load profile.

According to Eq. (1) and Fig. 1, the Beta probability distribution function in the variable intervals c_1 and c_2 are presented. According to the Fig. 1 and the purple curve, as well as using the curve fitting algorithm, the ability to describe the behavior of sunlight based on radiation data is provided. According to the first part of Eq. (1), the probability behavior of the PV unit according to the solar radiation power variable based on the Beta probability distribution function is described. In the second part of Eq. (1), the power generation by the PV unit is presented based on the probability variable of sunlight (P_{solar}). Therefore, according to Eq. (1), it is possible to describe the probability behavior of the PV unit and predict the output power of the PV unit in different weather conditions.

2.2. EV units uncertainty model

Formulating the behavior of EV units is also based on probability data resulting from the behavior and consumption pattern of subscribers. According to the Fig. 2, how to model the behavior of EV units according to the Weibull probability distribution functions and its generalized models is presented (Sepehrzad et al., 2024), (Mobarakeh et al., 2023). Considering that EV units are either considered as load or power supplier in the power flow equations, it is important to describe the behavior of EV units. Therefore, three categories have been used to describe the behavior of units. According to Fig. (2-a), the first category includes data and probability variables that describe the EV connection behavior to the NMG. In this category, the behavior of EV units is such that the EV units are connected to the NMG either to absorb power or to inject power. According to Fig. (2-b), the second category includes data and probability variables that indicate the EV disconnection from the NMG. The separation of EVs from the NMG also provides the NMG operator with information about the EV units that have played a role in power injection or absorption.

$$PDF\left(f_{EV}^{AT}\right) = \frac{e^{-\left(1+\frac{k_{EV}^{AT}\cdot\left(d-\mu_{EV}^{AT}\right)}{\sigma_{EV}^{AT}}\right)^{-\left(\frac{1}{k_{EV}^{AT}}\right)}} \cdot \left(1+\frac{k_{EV}^{AT}\cdot\left(d-\mu_{EV}^{AT}\right)}{\sigma_{EV}^{AT}}\right)^{-\left(1+\frac{1}{k_{EV}^{AT}}\right)}}{\sigma_{EV}^{AT}} \tag{2}$$

$$PDF\left(f_{EV}^{DT}\right) = \frac{e^{-\left(1+\frac{k_{EV}^{DT}\cdot\left(d-\mu_{EV}^{DT}\right)}{\sigma_{EV}^{DT}}\right)^{-\left(\frac{1}{k_{EV}^{DT}}\right)}} \cdot \left(1+\frac{k_{EV}^{DT}\cdot\left(d-\mu_{EV}^{DT}\right)}{\sigma_{EV}^{DT}}\right)^{-\left(1+\frac{1}{k_{EV}^{DT}}\right)}}{\sigma_{EV}^{DT}} \tag{3}$$

$$\begin{cases} PDF(f_D) = \dfrac{c_2.e^{-\left(\frac{t}{c_1}\right)^{c_2}} \cdot \left(\frac{t}{c_1}\right)^{(c_2-1)}}{c_1}, t>0 \\ SOC(P_{BESS-EV}) = \left(\dfrac{D_{BESS-EV}^{\max} - D_{BESS-EV}}{D_{BESS-EV}^{\max}}\right).100 \end{cases} \tag{4}$$

$$\begin{cases} PDF(P_{solar}) = \dfrac{\Gamma(c_1+c_2).(1-P_{solar})^{c_2-1}.(P_{solar})^{c_1-1}}{\Gamma(c_1).\Gamma(c_2)}\Bigg|_{\frac{c_1=\frac{\mu_{solar}.c_2}{1-\mu_{solar}}}{c_2=\left(\frac{\mu_{solar}.(1-\mu_{solar}{}^2)-\sigma_{solar}{}^2}{\sigma_{solar}{}^2}\right), c_1, c_2>0}} \\ P_{PV} = N.\dfrac{(V_{MPP}.I_{MPP}).(V_{OC}-k_v T_{PV}).([I_{SC}+k_i(T_{PV}-25)])}{V_{OC}.I_{SC}}\Bigg|_{T_{PV}=T_{Amb}+\frac{P_{solar}(T_{nom}-20)}{0.8}} \end{cases} \tag{1}$$

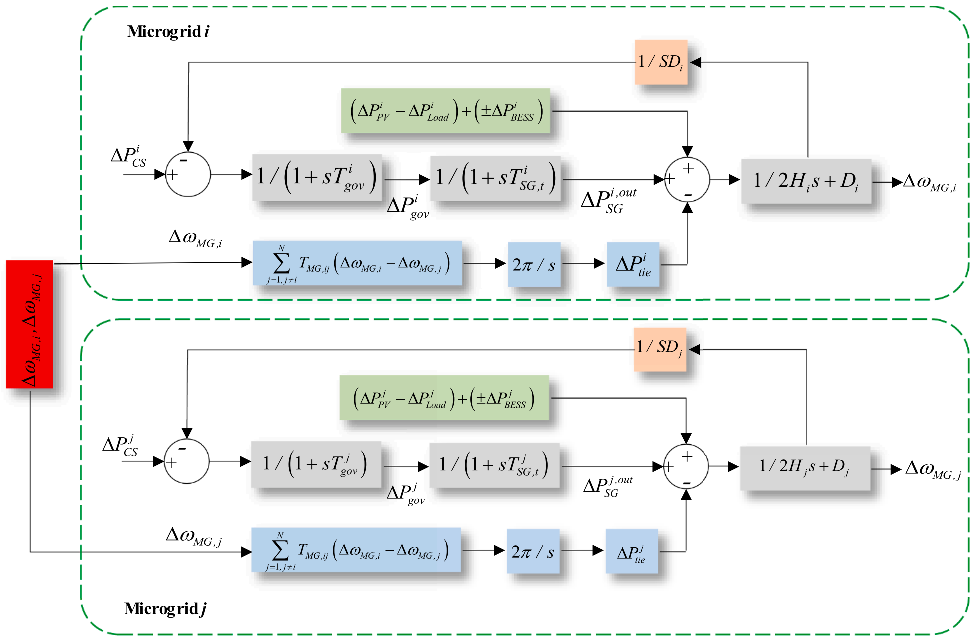

Fig. 4. NMG dynamic model.

$$P_{BESS-EV}(t)=\begin{cases}P^{\max}_{BESS-EV}\left(1-e^{\left[\frac{-(52.89)}{t_{ch}}\right]}\right),0<\mathrm{t}\leq \mathrm{t}_{ch}\\ P^{\max}_{BESS-EV}\left(\frac{5-t_{ch}}{5}\right),\mathrm{t}_{ch}<\mathrm{t}\leq 5\\ 0,\mathrm{t}>5\\ b^{ch}_{BESS-EV}+b^{dis}_{BESS-EV}=1,\forall t b^{ch}_{BESS-EV},b^{dis}_{BESS-EV}\in\{0,1\}\\ \quad P^{\min}_{BESS-EV}\leq P_{BESS-EV}(t)\leq P^{\max}_{BESS-EV}\Rightarrow P^{\min}_{BESS-EV}\\ \quad =(1-DOD).P^{\max}_{BESS-EV}\end{cases} \tag{5}$$

Therefore, according to this information, it is possible to consider the forecast for future hours in order to inject or absorb the power of EV units in the NMG. According to Fig. (2-c), the third category also includes data and probability variables due to the traveled distance by EV units. By using this curve, it can be find out the state of charge (SOC_{Bat}) of the EV units. Therefore, the NMG operator will have an estimate of the power injected into the grid by the EV or the power absorbed from the grid by the EV. According to the Fig. 2 and the blue curve, in order to reduce the prediction error, the generalized Weibull probability distribution function model has been considered. Therefore, according to Eqs (2-4), the behavior of EV units is presented in different categories. Then, in Eq. (5), the power of the EV unit with the Lithium-ion battery model (Nissan Altra which is $P^{\max}_{Bat-EV}=6.5kW$) is presented. The noteworthy point in Eq. (5) is the use of binary variables to describe the charging and discharging mode of EV units. Binary variables state that EV units are either in charging mode or in discharging mode and both modes do not occur together. Then, using the Monte Carlo method and according to Fig. 3, the load profile caused by the EV units can be obtained (Zhou et al., 2023).

2.3. EV units battery storage model

In this article, to validate and check the performance of the proposed strategy in the face of uncertainties, the battery of EV units has been considered as a BESS set in order to support the MG's frequency. Because the behavior of EV units is described based on the probability and random pattern, therefore, the performance of the proposed strategy based on uncertainty indicators and possible behavior of EV units has been analyzed. The equations of the BESS and the degradation index of the state of health (SOH) based on the SOC index and in the discrete time domain are:

$$E^{i}_{BESS,t}=E^{i}_{BESS,t-1}-\left(\frac{P^{i,dis}_{BESS,t}}{\eta_{dis}}-P^{i,ch}_{BESS,t}.\eta_{ch}\right)\Bigg|_{SOC^{i}_{BESS,t}=\frac{E^{i}_{BESS,t}}{E^{i,rate}_{BESS,t}}} \tag{6}$$

$$\begin{aligned}\Delta SOH_{BESS,t}&=\left(SOH_{BESS,t}-SOH_{BESS,t-1}\right)\\ &=\left(D.F_{BESS,t}.SOH_{BESS,t-1}\right),\begin{cases}D.F_{BESS,t}=\dfrac{d_{BESS,1}}{\left(\Delta SOC_{BESS,t}\right)^{d_{BESS,2}}}\\ \Delta SOC_{BESS,t}=SOC_{BESS,t}\\ -SOC_{BESS,t-1}\end{cases}\end{aligned} \tag{7}$$

SOC and SOH indices are programmed in [0 to 100%] range. Eqs. (6) and (7) are described according to Eqs. (2-5). The energy stored in BESS units is different and variable in different time periods. Therefore, Eq. (6) is for future times (t) and according to the energy stored in the BESS unit at the previous time (t-1). Therefore, in each time period, the initial amount of energy as well as the state of charge of BESS units are considered in order to predict the behavior of these units in the proposed controller. Then the Eq. (7) which is based on the $\Delta SOC_{BESS,t}$ and $D.F_{BESS,t}$ is presented. These indicators describe the changes in the BESS status of health ($\Delta SOH_{BESS,t}$).

2.4. Dynamic model and NMG frequency response

According to Fig. 4, the NMG model based on independent MGs that are connected to each other by a common bus is presented. Each MG includes a synchronous generator (SG), the BESS, electric load, the NMGLC and the PV panel. Each MG is connected to the NMGCC by a communication system. The frequency control algorithm based on the IPWFNN-DLRA approach is implemented in the NMGCC and the MG's frequency support control signals are uploaded to the NMGLC. The dynamic model and frequency response of the NMG are described according to Fig. 4 and the following equations (Reza Sepehrzad et al., 2022), (Reza Sepehrzad et al., 2022).

$$\Delta\omega'_{MG,i}=\frac{\left(\Delta P^{i,out}_{SG}+\Delta P^{i}_{PV}-\Delta P^{i}_{tie}-\Delta P^{i}_{Load}\right)+\left(\pm\Delta P^{i}_{BESS}\right)-\frac{\Delta\omega_{MG,i}.D_i}{2H_i}}{2H_i} \tag{8}$$

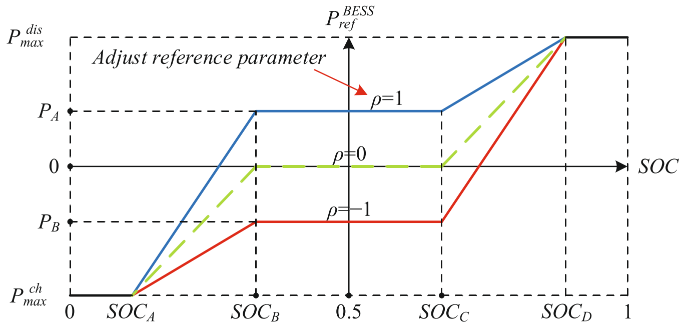

Fig. 5. Diagram of NMG frequency support model by BESS.

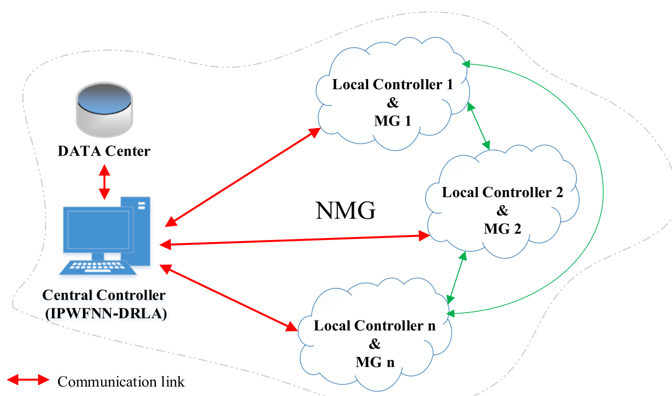

Fig. 6. NMG topology.

$$\Delta P'^{i,out}_{SG} = \frac{\Delta P^{i}_{gov}}{T^{i}_{SG,t}} - \frac{\Delta P^{i,out}_{SG}}{T^{i}_{SG,t}} \tag{9}$$

$$\Delta P'^{i}_{gov} = \frac{1}{T^{i}_{gov,t}}\left(\Delta P^{i}_{CS} - \frac{\Delta\omega_{MG,i}}{SD_i} - \Delta P^{i}_{gov}\right) \tag{10}$$

$$\Delta P'^{i}_{tie} = 2\pi\sum_{j=1,j\neq i}^{N} T_{MG,ij}\left(\Delta\omega_{MG,i} - \Delta\omega_{MG,j}\right) \tag{11}$$

$$T_{MG,ij} = \frac{|V_{MG,i}.V_{MG,j}|}{X_{MG,ij}}\cos\left(\delta^{0}_{MG,i} - \delta^{0}_{MG,j}\right) \tag{12}$$

Fig. 4 is the developed model of the NMGs frequency control structure, which includes the model of energy storage units in addition to renewable and non-RES. The NMG frequency control equations based on renewable, non-renewable sources and energy storage units are presented according to Eqs. (6-12). According to Fig. 4, the control signal caused by the frequency controller (ΔP^{i}_{CS}) is provided with the output feedback of the MG dynamic model ($\Delta\omega_{MG,i,j}$) and the error signal to update the control signals to each MG dynamic unit. Then, in addition to the synchronous generator, BESS units and PV units also participate in the dynamic frequency control of the NMG. In the dynamic model of NMG frequency support and control based on PV units and BESS units, the difference index of frequency changes of different MGs ($\Delta\omega_{MG,i} - \Delta\omega_{MG,j}$) has been used to reduce the global frequency synchronization error of NMGs. This difference causes an error signal that if the frequency of different MGs are not equal to each other, then compensation and synchronization of the global frequency is done by synchronous generators, PV units and also BESS units. NMG's frequency equations are described by focusing on SG units. The expressed differential equations of the dynamic behavior and frequency response of each MG are based on the linearized load frequency control (LLFC) model (Reza Sepehrzad et al., 2022), (Reza Sepehrzad et al., 2022).

3. NMG frequency support model by the BESS based on the IPWFNN-DRLA

3.1. Microgrid frequency support model by the BESS

Considering that changes in active power cause deviation and disturbance in MG's frequency index, therefore, BESS units will play an effective role in controlling changes in MG active power. In this article, BESS units are modeled with the aim of active power control (MG frequency support) and also support of SG and PV units. The BESS unit behavior based on SOC index recovery and NMG's frequency support is determined by the P-SOC characteristic curve. The P-SOC characteristic is described in Fig. 5. NMG's frequency support equation are:

$$P^{ref}_{BESS} = \frac{1}{2}\left((1+\rho).P_{BESS,A} + (1-\rho)P_{BESS,B}\right) \tag{13}$$

According to Fig. 5, two charging and discharging rate models have been considered for the BESS storage unit.

For example, according to the vertical axis of Fig. 5, the power, charge and discharge rate of the BESS unit are described between the values 0 to P_A and P_B, as well as between P_A and P_B to $P^{ch}_{\max}$ and $P^{dis}_{\max}$ When the BESS unit is charged and discharged in the power range between 0 to P_A and P_B, it experiences less electrical stress than when the BESS unit is charged and discharged in the power range between P_A and P_B to $P^{ch}_{\max}$ and $P^{dis}_{\max}$.

When the BESS unit is charged and discharged in the power range between 0 to P_A and P_B, the BESS unit is charged and discharged at a constant power rate in case of SOC changes. These conditions prevent electrical stresses on the BESS unit. The charging and discharging of the BESS unit is done with a constant rate of power in the range of SOC_B and SOC_C. However, these operating conditions with a fixed rate are not observed by all operators, and therefore the energy storage may be fully charged or completely discharged. Therefore, in the range of COC_C and SOC_D, as well as the range of SOC_A and SOC_B, the rate of power changes based on the changes in SOC of the BESS unit is described as a slope function. Changes in the SOC cause linear changes in the charge and discharge power.

For NMG frequency support, by optimizing the Fig. 5 according to theρ, the MG's frequency support structure can be created with the approach of supporting SG and PV units. Considering that the $\rho$is defined in the interval [1 and −1], it can be said that if $\rho = 1$, the BESS units will inject a power equivalent to $P^{ref}_{BESS} = P_{BESS,A}$ into the network. Also, if $\rho = -1$, the BESS units with a power equivalent to $P^{ref}_{BESS} = P_{BESS,B}$will receive power from the network. Also, in SOC_A and SOC_D, the BESS units follow powers of $P^{ch}_{\max}$and $P^{dis}_{\max}$. Also, to increase the BESS unit lifespan, the units with high SOH rate will contribute more to MG's power supply than the units with lower SOH rate.

Also, as stated previously, the SOC of the energy storage units is a measurable parameter. Therefore, the reference power for the participation of the BESS unit in supporting the frequency equivalent to 50% has been considered. This choice is due to the fact that it is possible for the user to obtain information about the participation of BESS units in injecting power into the NMG or absorbing power from the NMG. In general, BESS units with less than 50% power have the possibility of absorbing power from the NMG, and BESS units with more than 50% power have the possibility of injecting power into the NMG. Therefore, the reference power of the BESS unit has been considered in the reference value of 50% in order to determine the scenarios of the possibility of charging and discharging the BESS units. Therefore, the control index χ is defined to control the output power of BESS units, which are:

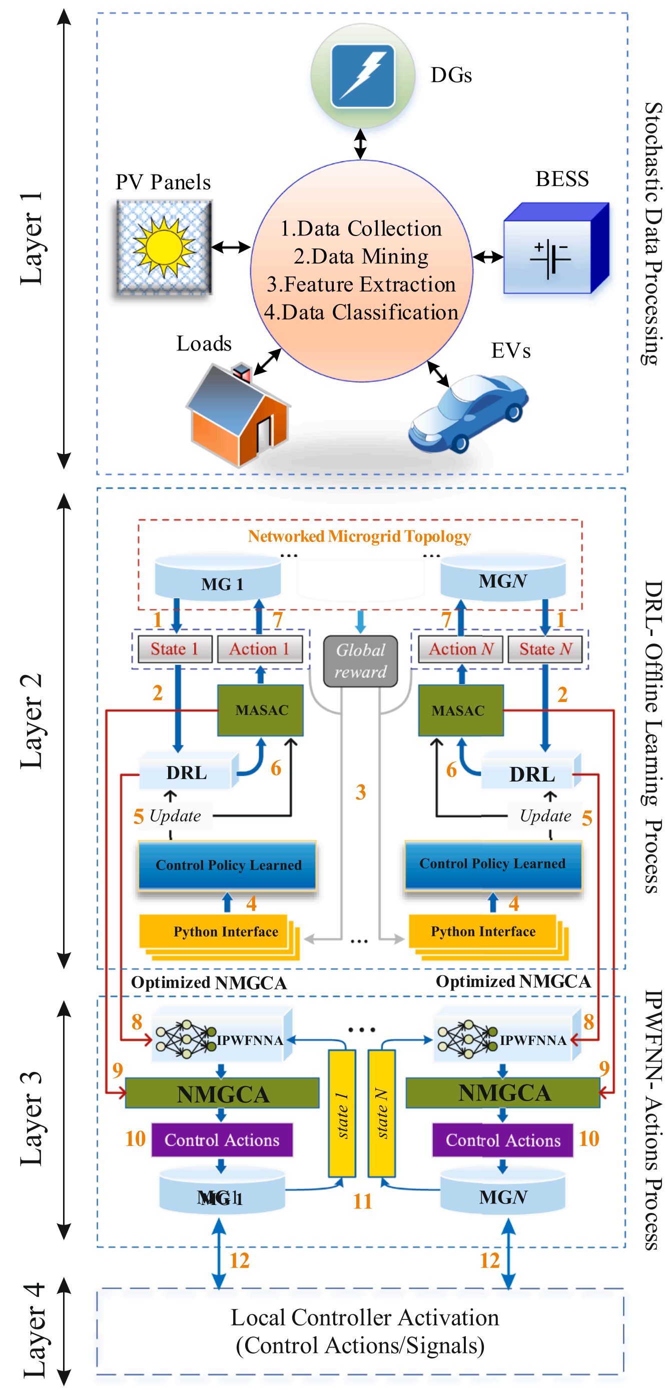

Fig. 7. Flowchart of the proposed control model based on the IPWFNN-DRLA algorithm.

$$P_{BESS} = \chi . P^{ref}_{BESS} \Big|_{\chi = \frac{1+\left(1+e^{\beta_1 \cdot (1-SOH_{BESS})}\right)}{\beta_2}} \tag{14}$$

3.2. NMG central control based on the IPWFNN-DRLA approach

Because the proposed algorithm implementation requires updating and having the real-time information of the MGs, therefore, according to Fig. 6, the MG's control model based on the NMGCC and NMGLC is presented in the context of the communication system. The MG's information is provided to the NMGCC by NMGLC every 5 ms. Then, the NMGCC calculates MG's frequency support and control algorithms based on the IPWFNN-DRLA approach and sends the results to the NMGLC. The NMGLC provides the NMGCC with information such as the voltage, frequency, active and reactive power, information on power generation sources and the status of BESS units. According to Fig. 7, the flowchart of the proposed control model based on IPWFNN-DRLA is presented. The structure and approach of DRL based on MDP are formulated and solved by the SAC algorithm (Haarnoja et al., 2018). The proposed strategy is developed in two structures of centralized training and decentralized operation. For this purpose, each MG has an NMGCA based on the IPWFNN algorithm (Ma et al., 2023), (Khodadadi et al., 2024). The learning model of the IPWFNN algorithm is based on the OBPL algorithm.

The framework of the proposed control model is developed in four layers according to Fig. 7. In the first layer, the operator of data collection, data mining, extraction of features, as well as classification of information based on the NMG structure and in the stochastic data processing framework are presented. Probability distribution functions and prediction functions are used in this layer. The second layer (include 12 steps shown in Fig. 7) includes the implementation of the offline learning algorithm and DRL. In this layer, in order to achieve learning actions, first, different situations based on load power variables and PVs panel output power (data of the first layer) are determined. Then control policies based on decision variables and states are determined. The control policies are updated by the DRL algorithm in each learning episode and iteration. The results of the control policies are converted into control actions by the MDP algorithm. The proposed actions in this section are based on the probability data from the first layer. Then, in the third layer, the IPWFNN-action process algorithm is executed. In this structure, by using the IPWFNN algorithm, it is possible to minimize the error caused by the actual reference power with the operating power value, as well as the reference power prediction error rate from the operating power prediction. Then, in the fourth layer, the control algorithms are loaded on the local controllers so that the control actions can be implemented.

3.2.1. MDP model

According to Fig. 7, each MG has an NMGCA that is responsible for calculating and implementing control policies (CP) based on the DRL approach. Each indicatorCPcauses an action (A) based on the observation state (S). Index S are considered as input variables. Therefore, the actions resulting from the control policies are$A = CP(S_t|DM_{CP})$whichDM_{CP}indicate the index decision-making variables of theCP. The purpose of the proposed algorithm is to improve the control policy learning approach (CP) based on accumulated reward (r). The MDP model is defined based on status, action, reward and constraints, which are:

1. **State:** In order to realize control policies and support the frequency of the MG in accordance with probability indicators such as the random behavior of loads and the probability power generation of PV units, previous and updated data and information of loads and PV units are needed. Therefore, the approach of the state, which is considered as the input of the proposed strategy, can be expressed as follows:

$$S_t^{MG,i} = \left\{ \Delta\omega_t^{MG,i}, P_t^{MG,i,L_1}, P_t^{MG,i,L_2}, \dots, P_t^{MG,i,L_{N_D}}, P_t^{MG,i,PV_1}, P_t^{MG,i,PV_2}, \dots, P_t^{MG,i,PV_{N}} \right\} \tag{15}$$

2. **Action:** The actions of each NMGCA unit are aimed at managing the output of DG and BESS units, which are:

$$A_t^{MG,i} = \left\{ \Delta P_{CS}^i, \rho_t^i \right\} \tag{16}$$

3. **Reward:** In the DRL algorithm, reward equations are expressed in order to analyze the performance of A_t, S_tindicators. Therefore, in this article, the reward equations based on minimizing the deviation of the frequency index from the reference values and minimizing the operating cost of SG units are expressed. The reward equations are:

$$r_t^i = W^{\omega}.S_d^{\omega} + W^{DG}.S_d^{DG} \begin{cases} S_d^{\omega} = -\Delta t.\sum\limits_{i\in N} \left| \omega_{MG,i}^{ref} - \omega_{MG,i}^{OP} \right| \\ S_d^{DG} = -\Delta t.\sum\limits_{i\in N} \left| a^i \left(P_t^{i,SG} \right)^2 + b^i.P_t^{i,SG} + c^i \right| \end{cases} \tag{17}$$

4. **Constraints:** Action constraints are defined according to Eqs. (18) and (19). The BESS unit constraints are expressed according to Eqs. (20-23). Also, the frequency constraint is expressed according to Eq. (24) in order to maintain the MG's stability and security.

$$0 \le P_{CS}^i \le P_{CS}^{i,\max} \tag{18}$$

$$-1 \le \rho_t^i \le 1 \tag{19}$$

$$0 \le SOC_{BESS,\min}^i \le SOC_{BESS,t}^i \le SOC_{BESS,\max}^i \le 1 \tag{20}$$

$$0 \le SOH_{BESS,\min}^i \le SOH_{BESS,t}^i \le SOH_{BESS,\max}^i \le 1 \tag{21}$$

$$0 \le P_t^{ch} \le P_{\max}^{ch} \tag{22}$$

$$0 \le P_t^{dis} \le P_{\max}^{dis} \tag{23}$$

$$\omega_{MG,i}^{\min} \le \omega_{MG,i} \le \omega_{MG,i}^{\max} \tag{24}$$

3.2.2. NMGCA model based on the IPWFNN algorithm

Fig. 6 shows the structure of the proposed strategy. According to Fig. 7, the proposed algorithm is planned and designed with two approaches: centralized training (DRL) and decentralized operation (IPWFNN). NMG indicators such as frequency, voltage, power, behavior of DGs and EVs are the inputs of each agent. Each agent can also search different action space to create optimal control policies. The safe operation of the proposed system is based on the safe model and continuous monitoring of the proposed system actions. According to Fig. 7, in each iteration of learning, in addition to storing the reward index, action and state indices are also stored in the buffer. The parameters of the operator are updated by the outputs of the buffer. After several iterations, optimal control policies based on the learning results from each agent in the NMGCA structure are generated. As stated, each NMGCA is programmed based on IPWFNN algorithm. The IPWFNN algorithm equations are designed in 5 layers, which are the input variables layer (L1), membership functions layer (L2), probability (Wavelet) layer (L3), fuzzification and defuzzification rules layer (L4) and output inference results layer (L5). The learning algorithm of each NMGCA is also expressed in four layers based on the OBPL algorithm. The layers of the OBPL algorithm are: the membership functions layer (L2), probability (Wavelet) layer (L3), fuzzification and defuzzification rules layer (L4) and output inference results layer (L5). Therefore, the equations describing NMGCA are:

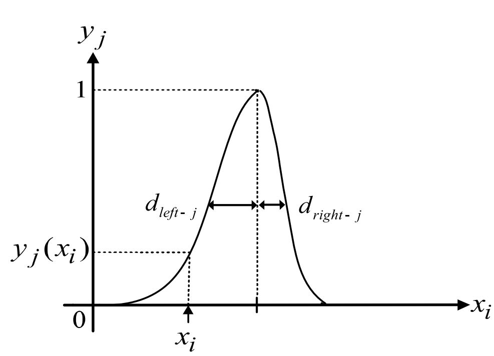

Fig. 8. Diagram of asymmetric Gaussian function based on IPWFNN algorithm structure.

Layer 1: Layer 1 includes the input variables of the algorithm. The most important input variables are active power changes and NMG's frequency deviation. Because the proposed structure is formulated based on learning algorithms, so to reduce the error, another input has been considered as the prediction error of the input variables.

$$I_i^{L_1} = \begin{cases} e_1(k) = e_P = P_{ref} - P_{OP} \\ e_1(k) = e_f = f_{ref} - f_{OP} \\ e_2(k) = e_{\text{Pre}} \end{cases} \tag{25}$$

Where$e_1$indicates the difference and error of the reference indices from the operating values and, $e_2$indicates the error of the predicted indices compared to the actual values in each iteration k. The layer 1 equations based on input and output variables are:

$$L_1 \begin{cases} NN_i^{L1}(k) = I_i^{L1} \\ O_i^{L1} = f_i^{L1}\left(NN_i^{L1}(k)\right) = NN_i^{L1}(k), \text{i} = 1, 2 \end{cases} \tag{26}$$

Layer 2: The equations of the membership functions (input of the second layer) are expressed based on the results of the first layer according to Fig. 8 and referring to Asymmetric Gaussian Function (AGF). The input and output equations of layer 2 are:

$$L_2 \begin{cases} NN_j^{L2}(k) = \begin{cases} -\dfrac{\left(I_i^{L2} - M_j^{L2}\right)^2}{\left(D_{l-j}^{L2}\right)^2}, -\infty < I_i^{L2} < M_j \\ -\dfrac{\left(I_i^{L2} - M_j^{L2}\right)^2}{\left(D_{r-j}^{L2}\right)^2}, \text{M}_j < I_i^{L2} < +\infty \end{cases} \\ O_j^{L2} = f_j^{L2}\left(NN_j^{L2}(k)\right) = e^{\left(NN_j^{L2}(k)\right)}, \text{j} = 1, 2, .., 6 \end{cases} \tag{27}$$

Layer 3: The probability equation of the third layer based on the Wavelet function and for each neuron is:

Table 2
Frequency response model coefficients (Zhou et al., 2023; Kumar & Karthikeyan, 2024; Bustos et al., 2023; Reza Sepehrzad et al., 2022; Zhao et al., 2023).

MG	T^i_{gov}	$T^i_{SG,t}$	H_i	D_i	SD_i
1	0.110	0.043	0.166	0.001	0.330
2	0.124	0.393	0.214	0.002	0.277
3	0.082	0.351	0.151	0.001	0.411
4	0.061	0.320	0.320	0.002	0.370

Table 3
DG/BESS coefficients (Bustos et al., 2023; Reza Sepehrzad et al., 2022; Zhao et al., 2023; Rosero et al., 2021; Wang et al., 2022).

MG	a	b	c	$d_{BESS,1,2}$	$\beta_{1,2}$
1	$1.5e^{-4}$	0.025	0.4	0.01/2	$9.2/8.2e^3$
2	$1.5e^{-4}$	0.076	0.6		
3	$1e^{-4}$	0.073	0.3		
4	$2e^{-4}$	0.072	0.7		

Table 4
IPWFNN-DRLA parameters (Onile et al., 2023; Wu et al., 2020).

Parameters	Value	Parameters	Value
Activation function (Hidden layers)	RelU	Discount factor	0.99
Sampling size	4000	Entropy index weight	0.02
Optimizer	Adam	Update rate	0.01
The learning rate	$1e^{-3}$	Loss function	Mean square error plus

Table 5
IPWFNN-DRLA training results.

Observation	samples	Mean square error
Validation	18,200	0.037
Training	87,000	0.034
Testing	18,200	0.037
Overall	106,000	0.036

Table 6
Comparison of the response time of the proposed strategy with WPFNN and ANN methods.

Frequency Controller	Response Time (*sec*)
IPWFNN	0.21
WPFNN (Ali et al., 2020)	0.54
ANN (Khodadadi et al., 2024)	1.02

Table 7
Comparing the training results and the computation burden of the proposed strategy with WPFNN and ANN methods.

Observation	No. of samples		
	IPWFNN	WPFNN (Ali et al., 2020)	ANN (Khodadadi et al., 2024)
Training	18,200	23,150	28,240
Validation	87,000	92,000	101,000
Testing	18,200	23,150	28,240
Overall	106,000	115,000	129,000

Table 8
Comparison of computation time of the proposed strategy with WPFNN and ANN methods.

No. of samples	Computation time		
	IPWFNN	WPFNN (Ali et al., 2020)	ANN (Khodadadi et al., 2024)
1000	0.21	0.54	1.02
2000	0.34	1.14	2.16
5000	0.84	2.43	3.76
8000	1.12	4.74	6.87

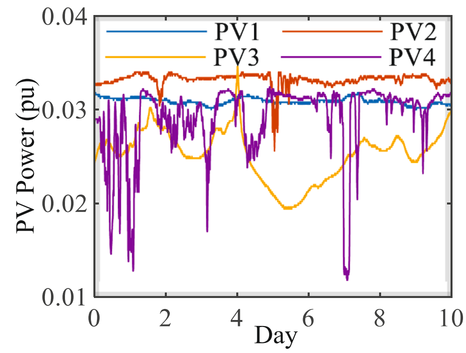

Fig. 9. PV power generation profile.

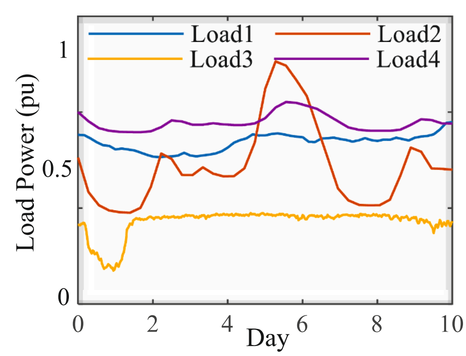

Fig. 10. Load power profile.

$$\lambda_q^{L3}(k) = \sum w_{iq}^{L3} . \frac{\left[1 - \frac{\left(I_i^{L1}(k) - M_{iq}^{L3}\right)^2}{\left(D_{iq}^{L3}\right)^2}\right] . e^{\left[-\frac{\left(I_i^{L1}(i) - M_{iq}^{L3}\right)^2}{2\left(D_{iq}^{L3}\right)^2}\right]}}{\sqrt{\left|D_{iq}^{L3}\right|}}, \mathrm{q} = 1, 2, \ldots, 9 \tag{28}$$

Therefore, the input and output equations of layer 3 based on the Wavelet function are:

$$L_3 \begin{cases} NN_x^{L3}(k) = \lambda_q^3(k) \\ O_x^{L3} = f_x^{L3}\left(NN_x^{L3}(k)\right) = \left(NN_x^{L3}(k)\right), \mathrm{x} = 1, 2, .., 9 \end{cases} \tag{29}$$

Layer 4: Layer 4 includes fuzzy rules based on membership functions formulated in the second layer (Ma et al., 2023), (Khodadadi et al., 2024). The fuzzy rules based on the normalized weight of the membership functions of the second layer are:

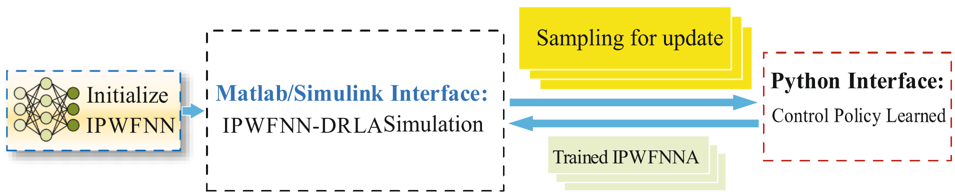

Fig. 11. IPWFNN-DRLA training framework.

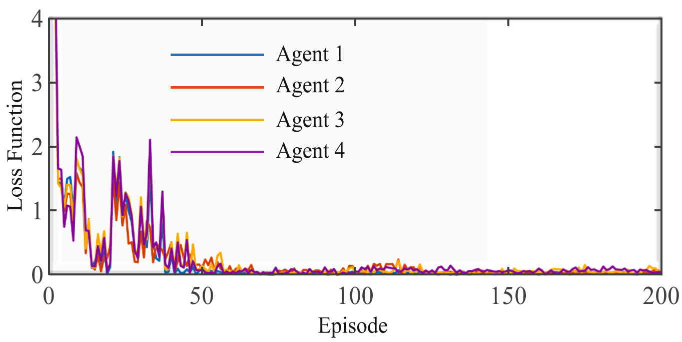

Fig. 12. Loss function of MASAC algorithm in different learning episodes.

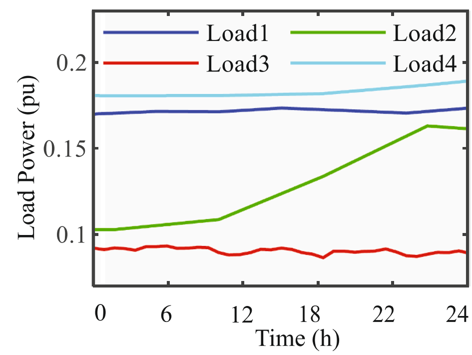

Fig. 15. Load power profile in Case 1.

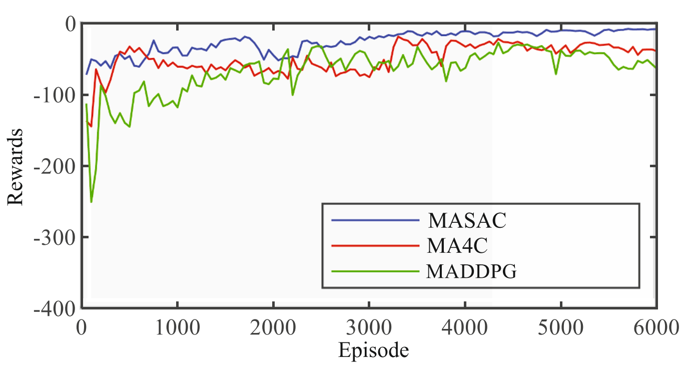

Fig. 13. Average reward of MASAC algorithm in different learning episodes.

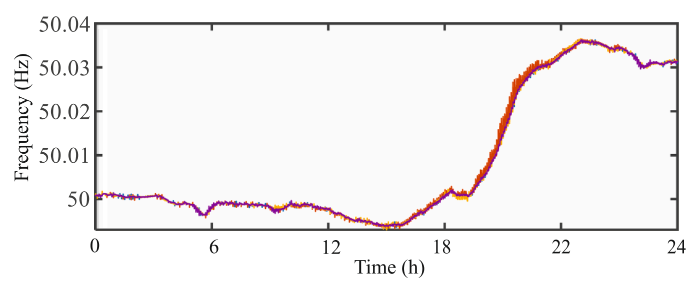

Fig. 16. Microgrid frequency control based on IPWFNN-DRLA in Case 1.

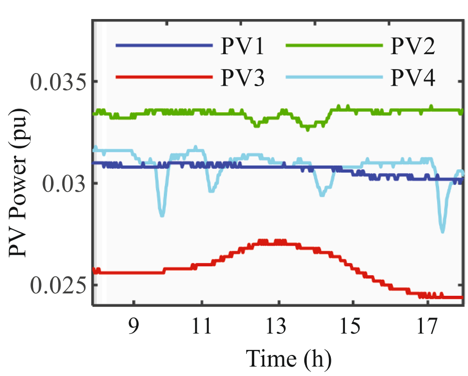

Fig. 14. PV power generation profile in Case 1.

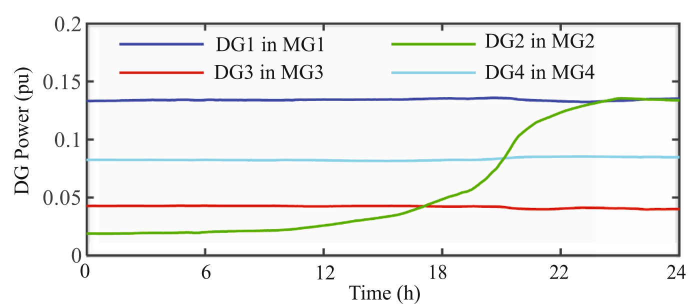

Fig. 17. DGs response to power changes in Case 1.

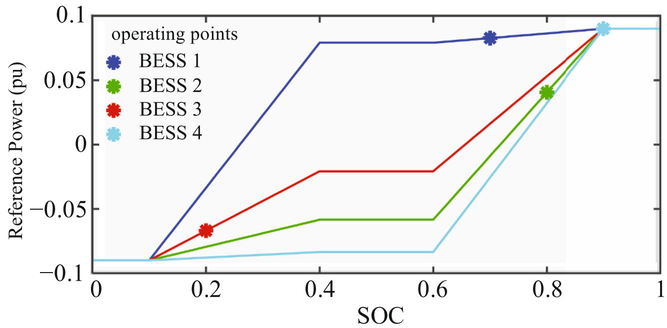

Fig. 18. The BESSs participation in the NMG's power changes control (Case 1).

$$L_4 \begin{cases} O_{jz}^{L4} = \prod_j w_{jz}^{L4}.O_j^{L2}, z = 1,2,..,9 \\ NN_z^{L4}(k) = O_{jz}^{L4}.O_z^{L3} \\ O_z^{L4}(k) = f_z^{L4}\left(NN_z^{L4}(k)\right) = \left(NN_z^{L4}(k)\right) \end{cases} \tag{30}$$

Layer 5: In layer 5, the inference results based on the defuzzification operator are determined and the operation currents of the converters are calculated. The results of this layer are sent to the controllers of each NMGCAs after several learning episodes in the continuous and online solution space.

$$L_5 \begin{cases} NN_g^{L5}(k) = \sum_{z=1} w_{zg}^{L5}.O_z^{L4}(k) \\ O_g^{L5}(k) = f_g^{L5}\left(NN_g^{L5}(k)\right) = \left(NN_g^{L5}(k)\right), \begin{cases} O_g^{L5}(k) = i_{q-P}^{OP} \\ O_g^{L5}(k) = i_{d-Q}^{OP} \end{cases} \end{cases} \tag{31}$$

3.2.3. OBPL model

The OBPL algorithm equations are formulated according to the proposed strategy assumptions and equivalent the following objective function (F):

$$F = \begin{cases} \frac{(P_{ref} - P_{OP})}{2} = \frac{e_P(i)^2}{2} \\ \frac{(f_{ref} - f_{OP})}{2} = \frac{e_f(i)^2}{2} \end{cases} \tag{32}$$

Eq. (32) shows the tracking error of the active power and frequency deviation in the learning algorithm with the aim of reducing this error. Therefore, the reverse learning algorithm based on the learning rate and the propagation process according to the results of layers 5 and 4 are:

$$E_{L_5} \Rightarrow \sigma_g^{L5} = -\frac{\partial(F)}{\partial O_g^{L5}(k)} = -\frac{\partial(F)}{\partial P_{OP}}.\frac{\partial P_{OP}}{\partial O_g^{L5}(k)} \tag{33}$$

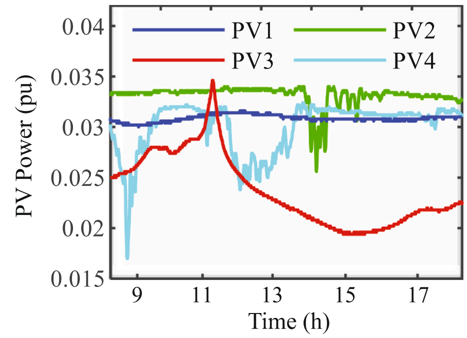

Fig. 20. PV power generation profile in Case 2.

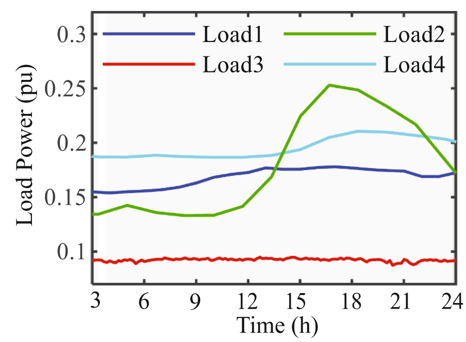

Fig. 21. Load power profile in Case 2.

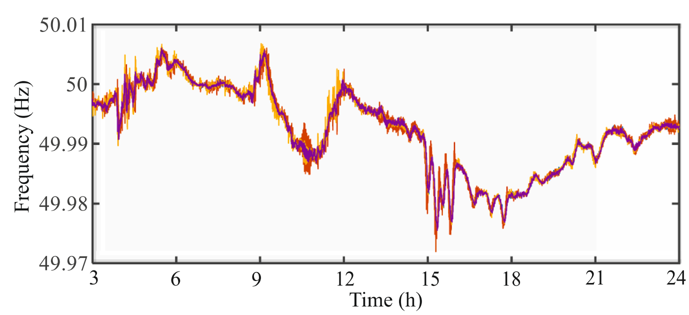

Fig. 22. MG frequency control based on IPWFNN-DRLA in Case 2.

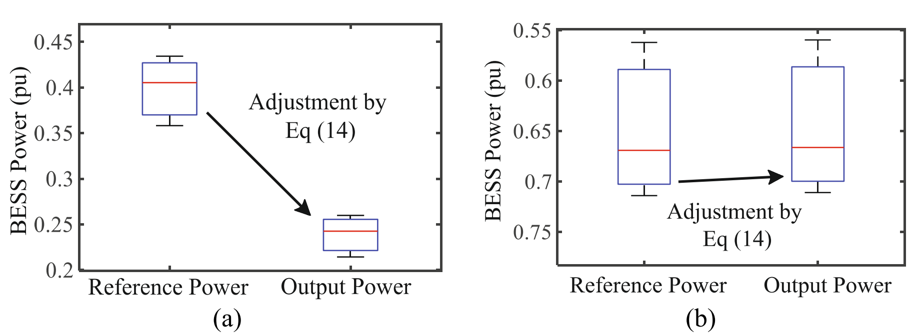

Fig. 19. The BESS reference/Output power in Case 1. (a) $BESS_4$ with SOC=20% &SOH=20%. (b) $BESS_1$ with SOC=90% &SOH=80%.

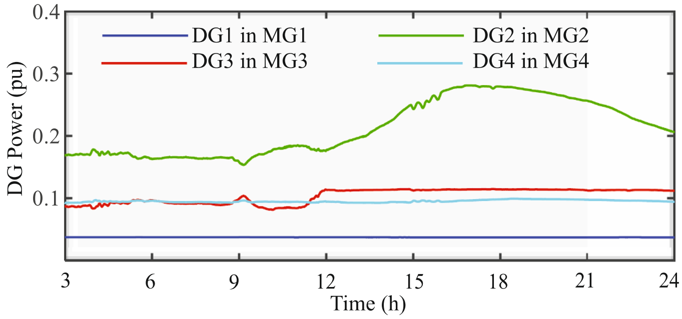

Fig. 23. DGs response to power changes in Case 2.

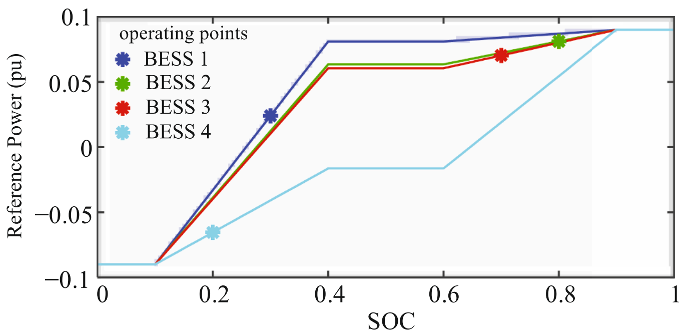

Fig. 24. The BESSs participation in the NMG's power changes control (Case 2).

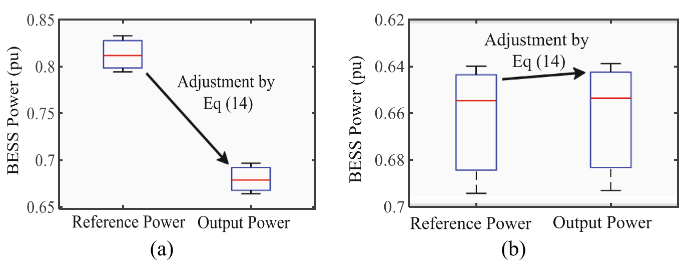

Fig. 25. The BESS reference/Output power in Case 2. (a) $BESS_4$ with SOC=60% &SOH=70%. (b) $BESS_1$ with SOC=90% &SOH=90%.

$$w_{zg}^{L5}(k+1) = w_{zg}^{L5}(k) + \Delta w_{zg}^{L5}\Big|_{\Delta w_{zg}^{L5} = -\mu_{zg}\frac{\partial(F)}{\partial w_{zg}^{L5}} = -\mu_{zg}\cdot\frac{\partial(F)}{\partial O_g^{L5}(k)}\cdot\frac{\partial O_g^{L5}(k)}{\partial w_{zg}^{L5}(k)} = \mu_{zg}.\sigma_g^{L5}.O_z^{L4}} \tag{34}$$

The tracking error in layer 4 based on the normalized weight of the learning rate and the propagation process are:

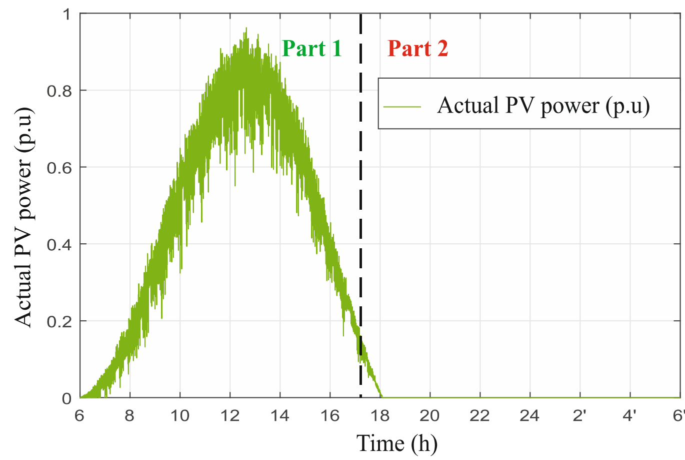

Fig. 27. PV power generation in 2 part.

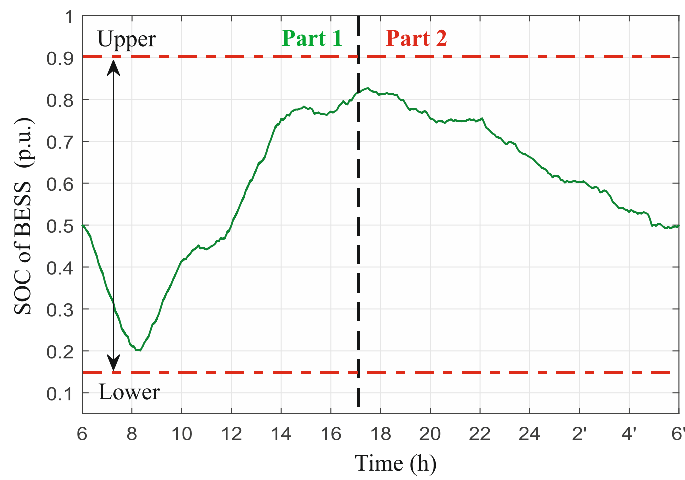

Fig. 28. The $BESS_3$ participation in the MG_3 power changes control.

$$E_{L_4} \Rightarrow \begin{cases} \sigma_z^{L4} = -\dfrac{\partial(F)}{\partial O_z^{L4}(k)} = -\left[\dfrac{\partial(F)}{\partial O_g^{L5}(k)}\right]\dfrac{\partial O_g^{L5}(k)}{\partial O_z^{4}(k)} = \sigma_g^{L5}.w_{zg}^{L5}(k) \\ \sigma_{jz}^{L4} = -\dfrac{\partial(F)}{\partial O_{jz}^{L4}(k)} = -\left[\dfrac{\partial(F)}{\partial O_g^{L5}(k)}\cdot\dfrac{\partial O_g^{L5}(k)}{\partial O_z^{L4}(k)}\right]\dfrac{\partial O_z^{L4}(k)}{\partial O_{jz}^{L4}(k)} = \sigma_z^{L4}.O_z^{L3}(k) \end{cases} \tag{35}$$

The tracking error in layer 3 based on the normalized weight of the learning rate and the propagation process are:

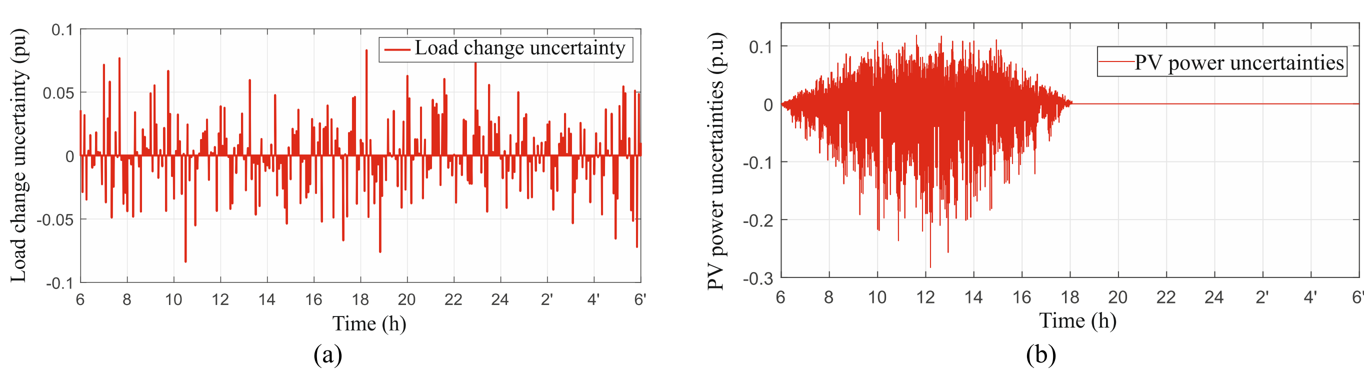

Fig. 26. (a) Load change uncertainty. (b) PV power uncertainty.

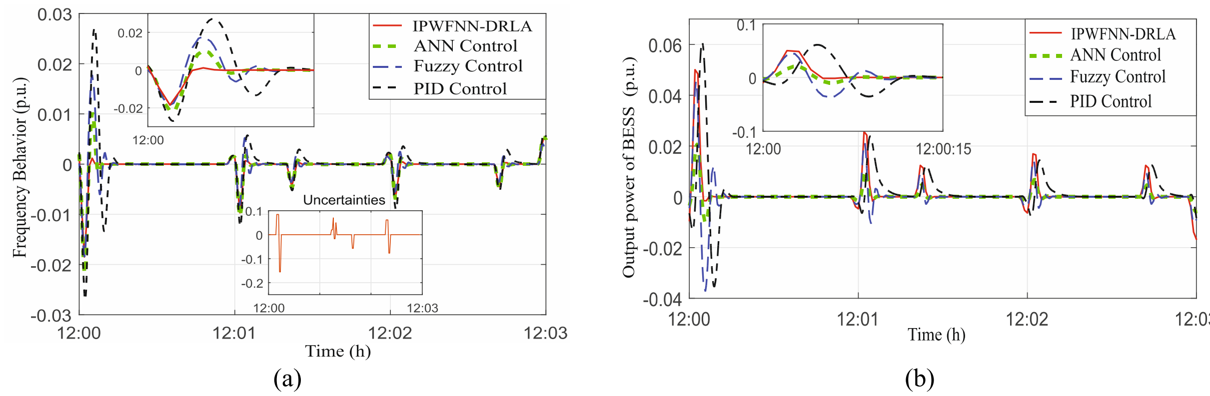

Fig. 29. (a) Frequency control comparison of MG_3 based on the IPWFNN-DRLA method with other methods. (b) The BESS participation in MG_3 frequency control.

Table 9
Experimental setup specifications.

Parameter	Specification	Parameter	Specification
PV type	HT Series TOPConH—NT11/ 66GDF 585–605	Converters efficiency	95% with THD<5%
Max- power of PV	607 W	LC filter	1.79mH/30μF
Voltage of PV at Max-power	40.6 V	BESS efficiency	0.95
Current of PV's at Max-power	14.96 A	Digital signal processor (DSP) type	TMS320F28335
PV efficiency	23.4%	switching frequency of PWM	12 kHz
OP 5707 XG	1. CPU Intel Xeon 8 cores, 3.8 GHz. 2. FPGA Xilinx® Virtex®−7 FPGA, 485T. 3. High speed communication 16 x SFP socket, 1 to 5 Gbps, duplex multimode optical fiber 50/125 µm with support for Xilinx® Aurora (1–5 Gbps).		
Communication software	EXata CPS v1.1 (1. Firewalls, 2. Intrusion Detection System (IDS), 3. Anti-Virus System (AVS), 4. Security Logs and Audit Trails)		
Software for NMG simulator	MATLAB 2019b simulations compiled to RT LAB 2021.3 & HYPERSIM 2021.3		
Communication protocols	1. IEC 61,850, 2. MODBUS, 3. IEEE C37 118		
PC type	Intel®Core™ i7–7600 U, 3.4 GHz CPU, and 16.00 GB RAM		

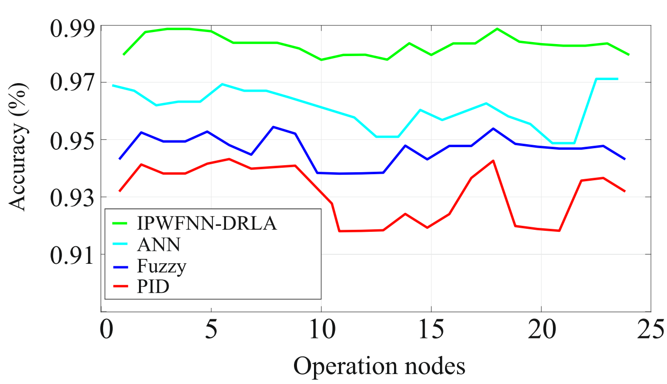

Fig. 30. Comparison of the accuracy of the proposed approach with other methods.

$$\sigma_x^{L3} = -\frac{\partial(F)}{\partial O_z^{L3}(k)} = -\left[\frac{\partial(F)}{\partial O_g^{L5}(k)} \cdot \frac{\partial O_g^{L5}(k)}{\partial O^{L4}(k)}\right] \frac{\partial O_z^{L4}(I)}{\partial O_x^{L3}(I)} = \sigma_z^{L4}.O_{jz}^{L4}(I) \tag{36}$$

$$w_{ik}^{L3}(I+1) = w_{ik}^{L3}(I) + \Delta w_{ik}^{L3}\Big|_{\Delta w_{ik}^{L3} = -\mu_{ik}\frac{\partial(O.F)}{\partial w_{ik}^{L3}} = -\mu_{ik}.\frac{\partial(O.F)}{\partial y_o^{L5}(N)}.\frac{\partial y_o^{L5}(I)}{\partial y_l^{L4}(I)}.\frac{\partial y_l^{L4}(I)}{\partial y_p^{L3}(I)}.\frac{\partial y_z^{L3}(I)}{\partial w_{ik}^{L3}(I)} = \mu_{ik}.\sigma_k^{L3}.W_{ik}^{L3}} \tag{37}$$

The tracking error in layer 2 based on the normalized weight of the learning rate, the propagation process and updating the asymmetric Gaussian function average values are:

$$\sigma_j^{L2} = \frac{\partial(F)}{\partial O_j^{L2}(k)} = -\left[\frac{\partial(F)}{\partial O_{jz}^{L4}(k)}\right] \frac{\partial O_{jl}^{L4}(k)}{\partial O_j^{L2}(k)} \cdot \frac{\partial O_j^{L2}(k)}{\partial O_x^{L3}(k)} = \sum_{jz} \sigma_{jz}^{L4} O_{jz}^{L4}(k) \tag{38}$$

$$\Delta M_j^{L2} = -\mu_{ij}\frac{\partial(F)}{\partial O_j^{L2}} = -\mu_{ij}\left[\frac{\partial(F)}{\partial O_j^{2}(k)}\right] \frac{\partial O_j^{L2}(k)}{\partial M_j^{L2}(k)} = \begin{cases} \dfrac{\left(I_i^{L2} - M_j^{L2}\right)^2 . \left(2\mu_{ij}\sigma_j^{L2}\right)}{\left(D_{l-j}^{L2}\right)^2} - \infty < I_i^{L2} < M_j \\ \dfrac{\left(I_i^{L2} - M_j^{L2}\right)^2 . \left(2\mu_{ij}\sigma_j^{L2}\right)}{\left(D_{r-j}^{L2}\right)^2} \mathrm{M}_j < I_i^{L2} < +\infty \end{cases} \tag{39}$$

The left and right standard deviations of the AGF are updated according to the following equations.

$$\Delta D_{l-j}^{L2} = -\mu_l\frac{\partial(F)}{\partial D_{l-j}^{L2}} = -\mu_l\left[\frac{\partial(F)}{\partial O_j^{L2}(k)}\right] \frac{\partial O_j^{L2}(k)}{\partial D_{l-j}^{L2}(k)} = -\frac{\left(2\mu_l\sigma_j^{L2}\right).\left(I_i^{L2} - M_j^{L2}\right)^2}{\left(D_{l-j}^{L2}\right)^3} \tag{40}$$

$$\Delta D_{r-j}^{L2} = -\mu_r\frac{\partial(F)}{\partial D_{r-j}^{L2}} = -\mu_r\left[\frac{\partial(F)}{\partial O_j^{L2}(k)}\right] \frac{\partial O_j^{L2}(k)}{\partial D_{ri-j}^{L2}(k)} = -\frac{\left(2\mu_r\sigma_j^{L2}\right).\left(I_i^{L2} - M_j^{L2}\right)^2}{\left(D_{r-j}^{L2}\right)^3} \tag{41}$$

To increase the learning rate and improve the learning speed, the adaptive error can be expressed as the following equation:

$$E^{Adaptive}\begin{cases} \sigma_g^5 = \left(P_{ref} - P_{OP}\right) + \left(P_{\mathrm{Pre}}^{ref} - P_{\mathrm{Pre}}^{OP}\right) \\ \sigma_g^5 = \left(f_{ref} - f_{OP}\right) + \left(f_{\mathrm{Pre}}^{ref} - f_{\mathrm{Pre}}^{OP}\right) \end{cases} \tag{42}$$

Using discrete time stability analysis based on the Lyapunov function brings the performance of the proposed strategy and the guarantee of the learning algorithm (Reza Sepehrzad et al., 2022).

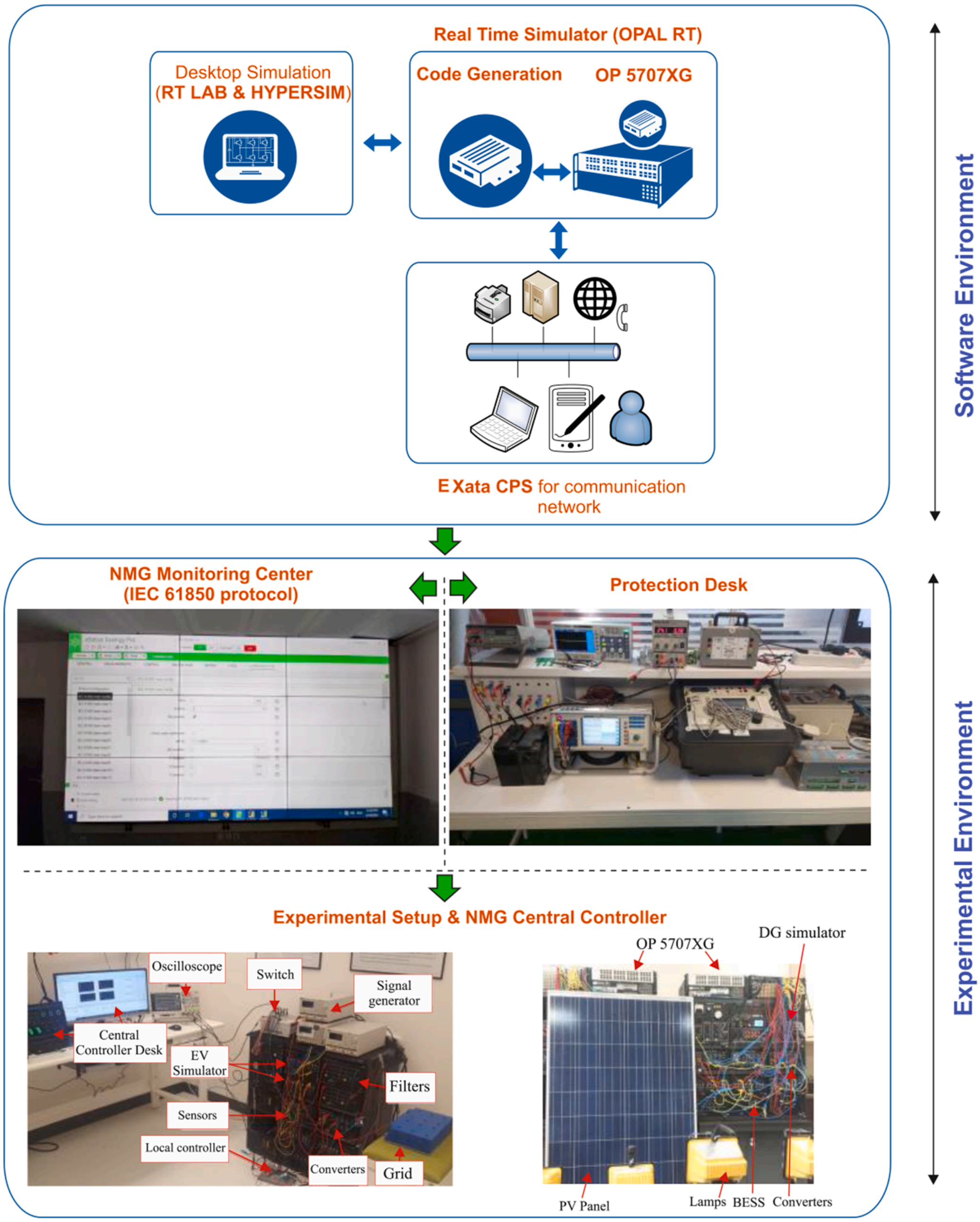

Fig. 31. Software and experimental setup environment.

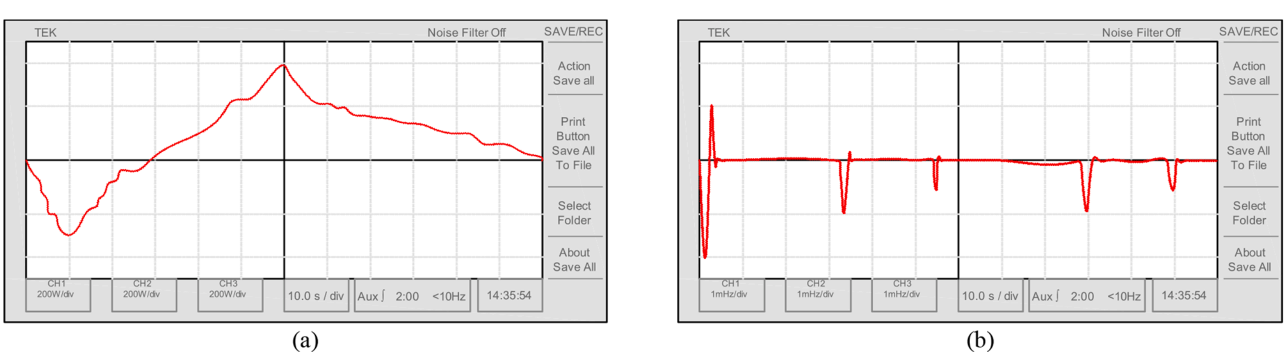

Fig. 32. Experimental results of case study 3. (a) $BESS_3$ participation in the MG_3 power changes control. (b) Frequency control.

3.3. Multi-agent soft actor-critic (MASAC) model

The actor-critic scheme, which includes the actor network for generating actions and the critic network for evaluating the actions of emerging RL algorithms, is widely used. Among the evaluation algorithms based on the actor-critic scheme, the SAC algorithm is presented with the aim of maximizing the objective function entropy with the expected efficiency. In this framework, the objective function entropy shows the randomness of actions based on control policies. Therefore, this algorithm is consistent with the proposed framework in this modeling. By using the SAC algorithm, the control policies related to each NMGCA unit can be optimized based on the maximum reward and random actions. The optimal control policies based on random actions are:

$$CP^{*} = \underset{CP}{\operatorname{argmax}} \sum_{t} E_{(S_t,A_t)}[r_t + w_{\Gamma}.\Gamma(CP(S_t|DM_{CP}))] \tag{43}$$

Eq. (43) is expressed based on the entropy function (Γ). Therefore, the entropy function is:

$$\Gamma(CP(S_t|DM_{CP}\Big)) = \underset{A\sim CP(S_t|DM_{CP})}{E}[-\log CP(S_t|DM_{CP})] \tag{44}$$

Using Eq. (44), the state-action function can be defined according to Eq. (45). Then, the goals of the critical network are described by referring to the Bellman's regularized equation and based on entropies in Eq. (46).

$$Z_{tar}(S_t,A_t) = \underset{S_t,A_t,r_t\in R}{E}\left[\sum_{t=0}^{T}\theta^t r_t + \sum_{t=0}^{T}\theta^t.\Gamma(CP(S_t|DM_{CP}))\right] \tag{45}$$

$$Z_{tar}(S_t,A_t) = \underset{\substack{S_{t+1}\in R \\ A_{t+1}=\Gamma(CP(S_t|DM_{CP}))}}{E}[r_t + \theta(Z_{tar}(S_{t+1},A_{t+1})) + w_{\Gamma}.\Gamma(CP(S_t|DM_{CP}))] \tag{46}$$

Eq. (47) also describes the training of the critic network based on the loss function in each time slot. According to Eq. (48), the actor network in the SAC algorithm can be described based on the Gaussian asymmetric probability distribution function and trigonometric functions.

$$J_Z(\varphi) = \underset{S_t,A_t\in R}{E}\left[\frac{(Z_{tar}(S_t,A_t|\varphi) - Z_{tar}(S_t,A_t))^2}{2}\right] \tag{47}$$

$$A_t = CP(S_t,\zeta_t) = \tanh(\mu_{DM_{CP}}(S_t) + \sigma_{DM_{CP}}(S_t).\zeta_t) \tag{48}$$

$$J_{CP}(DM_{CP}) = \underset{\substack{S_t\in R \\ \zeta_t\in N}}{E}[\log CP(S_t,\zeta_t|DM_{CP}) - J_Z(S_t, CP(S_t,\zeta_t|DM_P)|\varphi)] \tag{49}$$

Eq. (49) is defined in order to update the φindex. Therefore, the proposed strategy based on DRL and SAC algorithm approach has the ability to discover all the optimal paths during the execution of the learning algorithm. Therefore, the proposed strategy has a higher performance and efficiency compared to the deep deterministic policy gradient (DDPG) method.

4. Simulation and experimental results

4.1. Simulation requirements and learning algorithms analysis

In this section, the simulation requirements and parameters are presented. The simulation structure includes 4 MGs consisting of SG, BESS, PV panel and load. NMG topology is implemented in MATLAB 2019b software environment. In order to maintain the security of operation, the constraint of frequency changes equivalent to [49.7 and 50.4] according to IEEE 1547 standard has been considered for the proposed system. The frequency response parameters of the NMG model based on different MG components are presented in Tables 2 (Reza Sepehrzad et al., 2022), (Reza Sepehrzad et al., 2022) and 3 (Bustos et al., 2023), (Wang et al., 2022). Tables 4 and 5 show the specifications of the proposed strategy and the results of the learning algorithm (Onile et al., 2023), (Wu et al., 2020). In Tables 6–8 the comparison of the proposed strategy with other algorithms is presented. In the following, experimental results are also presented to validate the simulation results.

Also, according to Figs. 9 and 10, the information of the average probability power generated by PV panels and the load profile in a period of 10 days have been determined (Liu et al., 2019), (Reza Sepehrzad et al., 2022). In this study, in addition to real-time information, in order to improve the learning algorithm, the history of information has also been considered as input to the proposed strategy. Fig. 11 also shows each learning episode based on the IPWFNN-DRLA. At first, the IPWFNN algorithm is initialized based on past information. Then the results of this level are provided to the DRL algorithm for the learning process. After the completion of the first episode, a small batch of these results will be normalized in the Python Interface environment by the Pytorch library to update the IPWFNN algorithm and improve the control policies based on the NMGCA structure and will be provided to the DRL unit again. After each update, the next stages of learning are executed.

According to Fig. 12, the loss function of the multi-agent SAC (MASAC) algorithm is presented in different learning episodes of the DRL algorithm. As it is clear from this figure, the MASAC loss function reaches the minimum deviation after 50 learning episodes. Reducing the loss function increases the learning rate and reduces the learning deviation of the DRL algorithm in each episode. Reducing the loss function causes the convergence of the DRL algorithm in determining learning control policies and improves the response speed of the proposed strategy. Fig. 13 also shows the average reward of the MASAC algorithm in different learning episodes. The performance of the proposed strategy is compared with other methods such as multi-agent deep deterministic policy gradient (MADDPG) and multi-agent asynchronous advantage actor-critic (MA4C) and the results of this comparison show the high convergence speed and less fluctuation of the proposed strategy compared to other methods.

4.2. Results

4.2.1. Case study 1

Three case studies have been defined to validate and check the performance of the proposed strategy. Case 1 includes load power changes and performance analysis of the proposed strategy in NMG frequency control. The basic power in this study is equal to 1 MW and the basic power for BESS units is assumed to be equal to 0.1 MW. In case 1, each MG includes BESS units with SOC characteristics of 90%, 80%, 70%, and 20%. Also, the SOH characteristic of each BESS unit is 80%, 90%, 60%, and 20%, respectively. Power generation by PV units as well as load behavior is shown in Figs. 14 and 15.

According to Fig. 15, at $t = 9$, with the increase of the load power in MG_2, in addition to the SG unit, the BESS units also participate in providing the required power of the load. Therefore, according to Fig. 16, the frequency control of the MG is presented in the condition that the load fluctuation occurs. According to Fig. 17, the participation of BESS and SG units has caused continuous frequency control based on the IPWFNN-DRLA within the acceptable range of frequency changes.

Due to the load power change and SG_2 power changes, other MGs are also affected by this power stress. Therefore, to control these changes, BESS units have participated in controlling and eliminating these fluctuations and have controlled the disturbance without changing the power of SG units. The behavior of BESS units and the power change rate of each unit in each MG are shown in Fig. 18. Considering that BESS1,2 units have significant SOC and SOH characteristics, therefore, the participation rate of these two units in controlling power fluctuations is higher than other BESS units. Therefore, because the $BESS_4$ unit, which has low SOC and SOH characteristics, is considered as the last unit in the prioritization of power supply. According to the IPWFNN-DRLA method, in addition to the optimal control of the frequency in the acceptable range, the selfprotection algorithm is also implemented in order to increase the lifespan of the BESS units. According to Fig. 19 (a and b) and Eq. (14), the contribution of BESS units based on SOC and SOH characteristics is determined. For example, because the $BESS_4$ unit has low SOC and SOH characteristics, it is the last priority in providing power and controlling power changes. $BESS_1$ unit also has high SOC and SOH characteristics, so it has a high priority in controlling MG power changes. According to Fig. 19(a), after adjusting the power of the $BESS_4$ unit, the output power of this unit is lower than the reference power of the $BESS_4$ unit. While the $BESS_1$ output power follows its reference power and the output power is almost consistent with the reference power of the $BESS_1$ unit.

4.2.2. Case study 2

In case 2, each MG includes BESS units with SOC characteristics of 90%, 80%, 80%, and 60%. Also, the SOH characteristic of each BESS unit is 90%, 80%, 80%, and 70%, respectively. Power generation by PV units as well as load behavior are shown in Figs. 20 and 21. In this case study, the performance of the IPWFNN-DRLA method in MG frequency control based on load power changes and sudden power changes of PV units is presented.

According to Figs. 20 and 21, with the increase in load and decrease in generation power by PV units, the global frequency of the MG faces fluctuations in addition to reduction. According to Fig. 22, the IPWFNN-DRLA method has not only controlled the frequency fluctuations of the MG in the acceptable range, but also controlled the stability of the MG in the conditions of severe power changes. Fig. 23 also shows the behavior of SG units in response to MG power changes. Because most load changes occurred in MG_2, therefore the changes of SG_2 unit are more than other SG units. Also, Fig. 24 shows the BESS units participation in controlling the power fluctuations caused by the PVs power generation uncertainty. As in the previous case study, BESS units with high SOC and SOH characteristics are more involved in MG power supply than other units. Although MG_3 faces severe power changes of PV units, the participation of the $BESS_3$ unit based on the IPWFNN-DRLA method controls the power changes. According to Fig. 25 (a and b) and Eq. (14), the contribution of BESS units based on SOC and SOH characteristics is determined. For example, because the $BESS_4$ unit has low SOC and SOH characteristics, it is the last priority in providing power and controlling power changes. $BESS_1$ unit also has high SOC and SOH characteristics, so it has a high priority in controlling MG power changes. According to Fig. 25(a), after adjusting the power of the $BESS_4$ unit, the output power of this unit is lower than the reference power of the $BESS_4$ unit. While the $BESS_1$ output power follows its reference power and the output power is almost consistent with the reference power of the $BESS_1$ unit.

4.2.3. Case study 3

In this case study, the investigation of the uncertainties of load and PV units in the conditions of low and high changes and the performance of the BESS unit in controlling the uncertainties have been investigated. Also, to verify the performance of the proposed algorithm, the proposed strategy has been compared with other methods such as ANN, fuzzy and PID controller and its results have been presented. According to Fig. 26 (a and b), the uncertainties of load and PV unit in MG_3 are presented. As it is clear from these two figures, the uncertainties are presented with different ranges. Fig. 27 also shows the pattern of power generation by the PV unit, which is presented in two parts. The first part shows the power supply and the second part is when this unit is not able to generate power. According to Figs. 27 and 28, in the first part where the PV unit is able to generate power, the BESS unit is able to be charged according to the IPWFNN-DRLA method. While in the second part, due to the lack of power generation by the PV unit, the BESS unit injects the stored power into the grid. According to Fig. 28, the behavior of the BESS unit in the first part is in charge mode, while in the second part it is in discharge mode. Different uncertainties cause active power changes in the MG. Therefore, changes in active power also lead to changes in MG frequency. In this case study, the performance of the proposed strategy in controlling the frequency caused by the uncertainties of active power changes with different amplitudes is expressed. Fig. 29(a) shows the performance of the proposed strategy in controlling the frequency behavior of MG_3 based on power uncertainty with different amplitudes. Fig. 29(b) also shows the participation of the $BESS_3$ unit in MG frequency control and participation in the control of uncertainties. According to Fig. 29, compared to other methods, the proposed strategy has a high dynamic speed, so the MG experiences fewer frequency fluctuations caused by power changes.

4.3. Experimental results

In this section, to validate the simulation results, experimental results are also presented. According to Table 9, the equipment information and configuration of the prototype experimental set-up are provided. According to Fig. 31, the prototype experimental set-up is presented in two parts: the software and the experimental environment. The software environment includes RT LAB 2021.3 or HYPERSIM 2021.3 software, Real-Time simulator and EXata CPS v1.1 software. Also, the experimental environment includes a monitoring center based on the IEC61850 protocol, a protection desk, as well as an experimental setup and measurement equipment. In the simulation environment, first, the simulation structure of MATLAB 2019b software is compiled into RT LAB 2021.3 or HYPERSIM 2021.3 software, and then the implementation structure of the proposed algorithm is configured using the real-time simulator (OP5707 XG). By using EXata CPS v1.1 software, the structure of the communication system is configured. Also, by using this software, it is possible to create security firewalls to prevent cyber attacks. In the experimental environment section, various equipment is provided in sections such as monitoring, protection, experimental set-up, and central controller section. To validate the simulation results, only the experimental results of case study 3 have been evaluated and presented according to Fig. 32. The most important indicators presented in Fig. 32 are the performance of storage units in charge and discharge mode, as well as the performance of the proposed approach in frequency control. The presented results confirm the accuracy of the simulation results. The experimental results show that not only the proposed approach is highly accurate, but the high response speed of the proposed approach is also confirmed. According to Fig. 30 the experimental results have followed the simulation results with more than 98% accuracy.

5. Conclusion

In this study, the intelligent frequency control strategy of islanded NMG was presented taking into account the uncertainties of PV units and load power changes based on the IPWFNN-DRLA method. The proposed control model based on the NMGCC and NMGLC was presented in the context of the communication system. MG's information was provided to the NMGCC by NMGLC every 5 ms. Therefore, the NMGCC calculated MG frequency support and control algorithms based on the IPWFNN-DRLA approach and sent the results to the NMGLC. The NMGLC provided the NMGCC with information such as the voltage, frequency, active and reactive power, information on power generation sources and the status of BESS units. The DRL's structure and approach based on MDP were formulated and solved by the SAC algorithm. The proposed strategy was developed in two structures: centralized training and decentralized operation. For this purpose, each MG had an NMGCA based on the IPWFNN algorithm. The learning model of the IPWFNN algorithm was formulated based on the OBPL algorithm.

In this study, to verify the performance of the proposed strategy, different case studies based on load power changes and the uncertainty caused by the power generation of PV units were also analyzed. First, NMG frequency control based on proportional power sharing between BESS units and power generation sources and control of power uncertainties was presented. The results of this strategy and approach were compared to other methods such as the ANN, Fuzzy and PID controller. The obtained results showed the efficiency and high speed of the proposed strategy compared to other methods, especially for various uncertainties with different ranges. Because the proposed strategy had a high dynamic speed compared to other methods, the MG experienced fewer frequency fluctuations caused by power changes. The computation accuracy of the proposed approach is more than 98% in different operating scenarios and compared to other conventional and advanced methods. Also, the proposed approach has experienced a 7.82% reduction of the computation burden and 61.1% reduction of the computation time compared to other methods.

The most important future research in the framework of this study is the proposed strategy analysis in large-scale networks. Large-scale networks are considered as a system with big data that require the use of intelligent data mining and feature extraction strategies. For this purpose, the proposed strategy require suitable hardware infrastructures for big data processing to reduce the calculations time and volume.

CRediT authorship contribution statement

Reza Sepehrzad: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Amir Saman Godazi Langeroudi:** Data curation, Investigation, Resources, Software, Writing – review & editing. **Amin Khodadadi:** Data curation, Investigation, Resources, Validation. **Sara Adinehpour:** Data curation, Formal analysis, Methodology, Resources, Visualization. **Ahmed Al-Durra:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing. **Amjad Anvari-Moghaddam:** Conceptualization, Investigation, Supervision, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Reference

Abid, M. S., et al. (2024). Techno-economic and environmental assessment of renewable energy sources, virtual synchronous generators, and electric vehicle charging stations in microgrids. *Applied Energy, 35*, Article 122028.
Ajagekar, A., Decardi-Nelson, B., & You, F. (2024). Energy management for demand response in networked greenhouses with multi-agent deep reinforcement learning. *Applied Energy, 35*, Article 122349.
Ali, M., et al. (2020). Load forecasting through estimated parametrized based fuzzy inference system in smart grids. *IEEE Transactions on Fuzzy Systems, 29*, 156–165.
Almaleck, P., et al. (2024). Electrical consumption forecasting in sports venues: A proposed approach based on neural networks and ARIMAX Models. *Sustainable Cities and Society, 10*, Article 105019.
Babayomi, O., Li, Y., & Zhang, Z. (2022). Distributed consensus-based reactive power sharing in microgrids: A predictive virtual capacitance control technique. *International Journal of Electrical Power & Energy Systems, 14*, Article 108139.
Bustos, R., et al. (2023). Hierarchical energy management system for multi-microgrid coordination with demand-side management. *Applied Energy, 34*, Article 121145.
Cui, W., Jiang, Y., & Zhang, B. (2022). Reinforcement learning for optimal primary frequency control: A Lyapunov approach. *IEEE Transactions on Power Systems, 38*, 1676–1688.
Cui, W., & Zhang, B. (2021). Lyapunov-regularized reinforcement learning for power system transient stability. *IEEE Control Systems Letters*, 974–979.
Dorokhova, M., et al. (2021). Deep reinforcement learning control of electric vehicle charging in the presence of photovoltaic generation. *Applied Energy, 30*, Article 117504.
Eid, A., Mohammed, O., & El-Kishky, H. (2022). Efficient operation of battery energy storage systems, electric-vehicle charging stations and renewable energy sources linked to distribution systems. *Journal of Energy Storage, 5*, Article 105644.
Ferahtia, S., et al. (2022). Adaptive Droop based Control Strategy for DC Microgrid Including Multiple Batteries Energy Storage Systems. *Journal of Energy Storage, 4*, Article 103983.
Haarnoja, T., et al. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*. PMLR.
Hassanzadeh, M. E., et al. (2023). Hierarchical optimal allocation of BESS using APT-FPSO based on stochastic programming model considering voltage sensitivity and eigenvalues analyses. *International Journal of Electrical Power & Energy Systems, 15*, Article 109291.
Huang, Y., et al. (2023). Multi-objective optimization of campus microgrid system considering electric vehicle charging load integrated to power grid. *Sustainable Cities and Society, 9*, Article 104778.
Khalid, J., et al. (2022). Efficient load frequency control of renewable integrated power system: A twin delayed DDPG-based deep reinforcement learning approach. *IEEE Access : Practical Innovations, Open Solutions, 1*, 51561–51574.
Khodadadi, A., et al. (2024). Data-driven hierarchical energy management in multi-integrated energy systems considering integrated demand response programs and energy storage system participation based on MADRL approach. *Sustainable Cities and Society*. , Article 105264. https://doi.org/10.1016/j.scs.2024, 105264.
Kumar, R. P., & Karthikeyan, G. (2024). A multi-objective optimization solution for distributed generation energy management in microgrids with hybrid energy sources and battery storage system. *Journal of Energy Storage, 7*, Article 109702.
Li, J., Yu, T., & Zhang, X. (2022). Coordinated load frequency control of multi-area integrated energy system using multi-agent deep reinforcement learning. *Applied Energy, 30*, Article 117900.
Liu, J., Yao, Q., & Hu, Y. (2019). Model predictive control for load frequency of hybrid power system with wind power and thermal power. *Energy, 17*, 555–565.
Lovering, J. R. (2023). A techno-economic evaluation of microreactors for off-grid and microgrid applications. *Sustainable Cities and Society, 9*, Article 104620.
Ma, Y., Hu, F., & Hu, Y. (2023). Energy efficiency improvement of intelligent fuel cell/battery hybrid vehicles through an integrated management strategy. *Energy, 26*, Article 125794.
Mobarakeh, A. I., et al. (2023). Optimal planning and operation of energy hub by considering demand response algorithms and uncertainties based on problem-solving approach in discrete and continuous space. *Electric Power Systems Research, 21*, Article 108859.
Mukhopadhyay, B., & Das, D. (2021). Optimal multi-objective expansion planning of a droop-regulated islanded microgrid. *Energy, 21*, Article 119415.
Onile, A. E., et al. (2023). Energy efficient behavior modeling for demand side recommender system in solar microgrid applications using multi-agent reinforcement learning model. *Sustainable Cities and Society, 9*, Article 104392.
Parast, M. E., et al. (2023). Micrgrids and distribution system resilience assessment: A multi-objective robust-stochastic optimization approach considering social behavior using a deep learning modeling. *Sustainable Cities and Societ*, Article 104686.
Rosero, D. G., Díaz, N. L., & Trujillo, C. L. (2021). Cloud and machine learning experiments applied to the energy management in a microgrid cluster. *Applied Energy, 30*, Article 117770.
Sepehrzad, R., et al. (2021a). Intelligent energy management and multi-objective power distribution control in hybrid micro-grids based on the advanced fuzzy-PSO method. *ISA transactions, 11*, 199–213.
Sepehrzad, R., et al. (2021b). An efficient multilevel interconnect control algorithm in AC/DC micro-grids using hybrid energy storage system. *Electric Power Systems Research, 191*, Article 106869.
Sepehrzad, R., et al. (2022c). Optimal energy management of distributed generation in micro-grid to control the voltage and frequency based on PSO-adaptive virtual impedance method. *Electric Power Systems Research, 20*, Article 107881.
Sepehrzad, R., et al. (2022b). Intelligent hierarchical energy and power management to control the voltage and frequency of micro-grids based on power uncertainties and communication latency. *Electric Power Systems Research, 20*, Article 107567.
Sepehrzad, R., et al. (2022a). Islanded micro-grid frequency control based on the optimal-intelligent lyapunov algorithm considering power dynamic and communication uncertainties. *Electric Power Systems Research, 20*, Article 107917.
Sepehrzad, R., et al. (2023). Experimental and developed DC microgrid energy management integrated with battery energy storage based on multiple dynamic matrix model predictive control. *Journal of Energy Storage, 7*, Article 109282.
Sepehrzad, R., et al. (2024). Two-Stage experimental intelligent dynamic energy management of microgrid in smart cities based on demand response programs and energy storage system participation. *International Journal of Electrical Power & Energy Systems, 15*, Article 109613.
Sharma, P., et al. (2022). A critical and comparative review of energy management strategies for microgrids. *Applied Energy, 32*, Article 120028.
Wang, W., Huang, Y., Yang, M., Chen, C., Zhang, Y., & Xu, X. (2022). Renewable energy sources planning considering approximate dynamic network reconfiguration and nonlinear correlations of uncertainties in distribution network. *International Journal of Electrical Power & Energy Systems, 13*, Article 107791.
Wu, J., et al. (2020). Battery thermal-and health-constrained energy management for hybrid electric bus based on soft actor-critic DRL algorithm. *IEEE Transactions on Industrial Informatics, 17*, 3751–3761.
Yamashita, D. Y., et al. (2022). Autonomous observer of hydrogen storage to enhance a model predictive control structure for building microgrids. *Journal of Energy Storage, 5*, Article 105072.
Yan, R., et al. (2022). A multiagent quantum deep reinforcement learning method for distributed frequency control of islanded microgrids. *IEEE Transactions on Control of Network Systems, 9*, 1622–1632.
Yan, Z., & Xu, Y. (2018). Data-driven load frequency control for stochastic power systems: A deep reinforcement learning method with continuous action search. *IEEE Transactions on Power Systems, 34*, 1653–1656.

Yan, Z., & Xu, Y. (2020). A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system. *IEEE Transactions on Power Systems, 35*, 4599–4608.
Yang, M., Cui, Y., & Wang, J. (2023). Multi-Objective optimal scheduling of island microgrids considering the uncertainty of renewable energy output. *International Journal of Electrical Power & Energy Systems, 14*, Article 108619.
Zhang, B., et al. (2022). Physical-model-free intelligent energy management for a grid-connected hybrid wind-microturbine-PV-EV energy system via deep reinforcement learning approach. *Renewable Energ*.
Zhao, J., Wang, W., & Guo, C. (2023). Hierarchical optimal configuration of multi-energy microgrids system considering energy management in electricity market environment. *International Journal of Electrical Power & Energy Systems, 14*, Article 108572.
Zhou, X., et al. (2023). A three-stage mechanism for flexibility-oriented energy management of renewable-based community microgrids with high penetration of smart homes and electric vehicles. *Sustainable Cities and Society, 9*, Article 104946.