

# MANAGEMENT AND CONTROL OPTIMIZATION BASED ON DEEP LEARNING MODEL

---

**Jingjing Dai**

Nanjing University of Technology, Nanjing, Jiangsu, 210023, China

[c18905175401@163.com](mailto:c18905175401@163.com)



**Reception:** 19/10/2022 **Acceptance:** 26/12/2022 **Publication:** 19/01/2023

## Suggested citation:

D., Jingling (2023). **Management and control optimization based on deep learning model.** *3C Empresa. Investigación y pensamiento crítico*, 12(1), 37-49. <https://doi.org/10.17993/3cemp.2023.120151.37-49>

## ABSTRACT

*Microgrid technology is a key solution to improve distributed power consumption, complementary utilization of multiple energy sources, and power supply reliability. To guarantee the reliability of the microgrid system, a realistic strategy must be created. This work takes the microgrid as an object and uses simulation technology to construct a microgrid system. Then, using this simulation system and the double deep Q-learning, the goal is to minimize the 24-hour electricity consumption cost from the external power grid to meet the requirements of voltage deviation. Power balancing and energy storage loads for microgrid systems. Under the constraints of the electrical state and other constraints, the control variable is the energy storage's capacity for charging and discharging, and the optimization strategy of energy storage control is obtained through training. The results demonstrate that the DDQN algorithm will save 26.95% of the electricity purchase cost, which is significantly more than the MPPT algorithm's 12.43% savings. As a result, this work examines the efficacy of the charging and releasing approach for energy storage and confirms the potential of the suggested approach to reduce the cost of purchasing electricity.*

## KEYWORDS

*Deep reinforcement learning; DDQN; Microgrid technology; Optimization strategy; Electricity*

## PAPER INDEX

ABSTRACT

KEYWORDS

1. INTRODUCTION

2. SYSTEM MODEL

2.1. Microgrid Components

2.2. Deep reinforcement learning algorithm

3. RESULTS AND DISCUSSION

3.1. Network Settings

3.2. Numerical results

4. CONCLUSION

REFERENCES

# 1. INTRODUCTION

The power system's development of the energy structure has advanced significantly in recent years toward a clean and sustainable development due to the widespread availability of renewable energy. Because of structural changes on the energy supply side and the 2030 carbon peaking target, the amount of renewable energy in the power grid will increase even more[1]. The grid-connected use of large-scale renewable energy is faced with a significant problem because distributed renewable energy is highly volatile and intermittent due to the influence of natural factors [2]. A microgrid is a compact power generation [3]. Configuring an energy storage device is important to preserve the microgrid's power balance and lessen the effect that distributed energy output uncertainty has on the microgrid. A typical microgrid so typically consists of a variety of power-producing equipment, energy storage devices, loads, and other components [4]. At present, there are many studies on the microgrid, and the energy storage control strategy has attracted much attention as a hot spot for optimal regulation of energy dispatch.

Microgrid scheduling is a popular topic in research connected to microgrids and is a crucial tool for ensuring the safe, dependable, and cost-effective operation of microgrids [5]. Traditional microgrid optimization scheduling is usually based on optimization theories and methods. First, each component in the microgrid is modeled, then the model is simplified and processed, and finally, the model is solved by researching the related solution algorithm. Typically, the model's primary goal is to minimize operational costs, and there are also related studies that comprehensively consider the economic, environmental, and social benefits to establish a multi-objective optimization model [6]. Commonly used model modeling methods include mixed integer programming [7, 8], dynamic programming [9], model predictive control [10, 11], distributed optimization [12], Lyapunov optimization, etc.; commonly used model solving algorithms include Genetic algorithm [13, 14], particle swarm algorithm [15, 16], active evolution algorithm [17], Lagrangian relaxation method, etc.

The uncertainty of microgrid operation has greatly increased during the past few years because of the rising number of renewable energy sources. The common solution to such problems is to convert uncertain problems into deterministic problems for modeling and solving, mainly including scene-based stochastic optimization[18], Opportunity Constrained Optimization [19], Robust optimization [20], etc. However, these methods all have certain limitations, which are mainly reflected in: the refined modeling of each component of the microgrid is difficult, and the simplified model is difficult to describe the physical characteristics of the actual operation of the components, resulting in the optimization results may be sub-optimal; the established model It is generally nonlinear and non-convex, and its solution is a typical non-deterministic polynomial problem, and the solution efficiency is low; the establishment of the model needs to be completed according to the topology structure and operation mode, and the adaptability to the change of topology structure and the access of new power equipment Not strong; the pre-planned control based on "offline calculation and online matching" is difficult to adapt to complex and changeable system conditions.

Artificial intelligence has developed rapidly in recent years. As a typical timing control problem, the microgrid energy control problem conforms to the deep reinforcement learning solution framework, and there are many outstanding works at present. Hua et al. [21] used a deep learning algorithm with asynchronous advantages to solve the multi-microgrid energy dispatch control problem. Zhang et al. [22] gave the composite energy storage coordination control strategy of the microgrid through the deep reinforcement learning algorithm. Liu et al. [23] established a microgrid framework model based on energy buses and compared the advantages of a deep Q-learning algorithm in energy scheduling control problems compared with the heuristic algorithm. Du et al. [24] gave retail pricing strategies through Monte Carlo reinforcement learning algorithms to reduce the demand-side peak ratio and protect user privacy. Mocanu et al. [25] explored the use of deep learning to refine the usability of building energy management systems in a smart grid environment and was successfully validated on a large database.

When addressing the energy storage charging and releasing control technique with deep learning, in contrast to prior work, the training data is real-time data occurring while the microgrid simulation environment is running. By adding state information about voltage, current, phase angle, and other state information in the simulation environment to achieve a simulation that is closer to the real state, the results are more reliable. This paper firstly presents the composition of the microgrid objects studied, and then mainly introduces the deep reinforcement learning algorithm framework used and its application process in the energy storage control problem. Finally, the comparison with the existing method illustrates the usefulness of the algorithm flow in this research.

## 2. SYSTEM MODEL

### 2.1. MICROGRID COMPONENTS

A significant component of distributed energy in the microgrid system is photovoltaic power generation. The photovoltaic power can be given by equation (1)

$$P_{PV}(t) = R(t)S\eta_{PV} \quad (1)$$

$S$  is the area of the placed solar panels, and  $R(t)$  is the amount of solar radiation at time  $t$ , expressed in  $W/m^2$ . The power of photovoltaic energy generation at time  $t$  is calculated by multiplying the solar radiation power by the conversion efficiency  $\eta_{PV}$ , which is the product of the solar radiation power and the panel area.

The battery is a popular kind of high-efficiency energy storage, and its internal energy state satisfies the equation (2)

$$E_{bat}(t) = E_{bat}(t-1) + P_{bat}(t) \bar{t} \quad (2)$$

$P_{bat}(t)$  is the battery's charging and discharging capacity at time  $t$ ;  $\Delta t$  is the time interval between two charging and discharging operations.

**Load.** Load is the general term for the part that consumes electric energy in a microgrid system. For a fixed microgrid system, the load demand is determined by the local climate environment and the properties of the microgrid, so it is usually not adjustable. In the energy scheduling problem in this paper, the load curve is input to the microgrid as a fixed quantity. network system.

The external power supply used in this paper is replaced by the three-phase AC power supply module that comes with Simulink. The parameter settings of the module need to be given according to the actual simulation requirements. Then, it is connected with the internal module of the microgrid system through the varistor module to provide the required power for the system. In this study, a microgrid system that uses grid connectivity is taken into account. The three-phase dynamic load is also used to produce the PQ control effect since the energy storage system model that needs to be developed can be modified. In the simulation experiment, due to the randomness of the load, the microgrid load usually consists of two parts, one of which is a variable load and the other is a fixed load, which is directly acted by the resistance module in Simulink in the simulation model. The simulation modules of each distributed energy source are regulated by the PQ control strategy to guarantee the stability and controllability of the interaction between the microgrid system and the external power. Based on the simulation model, this paper conducts deep reinforcement learning training. As a result, the energy storage system's charging and releasing control approach is optimized. Compared to traditional reinforcement learning training based on mathematical formulas, the influence of more abundant state variables on the control objective can be comprehensively considered.

## 2.2. DEEP REINFORCEMENT LEARNING ALGORITHM

Reinforcement learning is aimed at maximizing the expected return. The mapping relationship between the environment's state variable and the agent's action variable is discovered through the agent and the environment's ongoing interaction. The agent provides an optimized action policy. Deep reinforcement learning uses deep neural networks to create a correspondence between state variables and action variables. Due to the powerful expressive ability of deep neural networks, deep reinforcement learning can deal with more complex and practical policy decision-making problems. In the disciplines of optimum control, robot control, and autonomous driving, deep reinforcement learning has made significant strides in recent years. Deep reinforcement learning is derived from MDP. The traditional MDP process consists of 4 elements, consisting of  $(S, A, R, \mu)$ , which stand for the set of environmental conditions, the set of actions the agent is capable of taking, the reward function, and the policy set, respectively.

Solving algorithms for reinforcement learning optimization strategies includes three categories: value function-based solutions, policy gradient-based solutions, and search and supervision-based solutions [26]. This paper focuses on the solution

method based on the value function. Among them, the dynamic programming algorithm is suitable for solving the situation with models and dimensions. The Monte Carlo technique has the drawback of requiring entire state sequences, which is challenging in many systems. Without the entire state sequence, the time series difference approach can estimate the value function. The classic time series difference methods include the SARSA and the Q-Learning algorithm. Both approaches keep a Q-table to address tiny reinforcement. The size of the Q-table that needs to be kept increases when the state and action spaces are continuous or discrete at very large scales, which will bring difficulties to storage. However, the development of deep neural networks has solved this problem. Using a deep neural network instead of Q-table results in a deep reinforcement learning algorithm that is more suitable for complex problems. A typical algorithm is the deep Q-Learning (deep Q network, DQN) algorithm.

The Q-Learning algorithm is responsible for formula-based updates to the defined Q function (3).

$$Q(S, A) = Q(S, A) + \alpha (R + \gamma \max_a Q(S', a) - Q(S, A)) \quad (3)$$

$Q(S, A)$  is a function of action value;  $\alpha$  is the learning rate. When the update formula converges, the optimal control strategy for reinforcement learning can be obtained.

DQN replaces the Q function in Q-Learning with a deep neural network  $Q(S, A | \omega)$ .

Formula (3) is then used to determine the current goal Q value, and the Q network updates the parameter  $\omega$  of the neural network.

The DDQN algorithm is based on DQN with two improvements. The first is to introduce two networks, one is the target network  $Q'(S, A | \omega')$  to calculate the target Q value, and the other is the update network  $Q(S, A | \omega)$  to update the Q value, which can reduce the Dependencies between small target Q-values and updated network parameters. The target network has the same structure as the Q-value network, and the timing is synchronized with the parameters from the Q-value network.

The second is to decouple the action selection of the target Q and the calculation process of the target Q in the target formula (3), thereby reducing the overestimation caused by the greedy algorithm. The specific method is that when calculating the target Q, the maximum Q corresponding to the action is no longer found in the target Q network, but the action corresponding to the maximum Q value is first found in the update network:

$$a(S | \omega) = \operatorname{argmax}_{a \in A} Q(S, A | \omega) \quad (4)$$

Then use this action  $a(S | \omega)$  to calculate the target Q value:

$$Q' = R + \gamma Q'(S', a | S') \quad (5)$$

Through the above two improvements, the DDQN algorithm solves the strong dependence and overestimation problems of the traditional DQN algorithm, and other algorithm processes are the same as DQN.

### 3. RESULTS AND DISCUSSION

#### 3.1. NETWORK SETTINGS

The implementation of the DDQN algorithm in solving the microgrid energy storage control problem is demonstrated in this part, along with a comparison to the MPPT control to demonstrate the efficacy of the DDQN method.

To speed up the simulation speed, in the simulation experiment verification, this paper does not consider other distributed generation modules such as wind power generation.

The upper and lower bound constraints in constraint (8) are set as:

$$P_{bat}^{min} = -4 \times 10^5 W, \quad P_{bat}^{max} = 4 \times 10^5 W, \quad S_{soc}^{min} = 0.19, \quad S_{soc}^{max} = 0.85$$

Another important setting is the setting of the Q-value estimation neural network in the DDQN algorithm. According to experience and continuous debugging and verification, the neural network used in this paper is a fully connected layer with 4 layers and 24 nodes, and the relu function is employed as the middle layer activation function.

#### 3.2. NUMERICAL RESULTS

When training the DDQN algorithm, it is necessary to normalize each component of the state variable and then substitute it into the target network of DDQN for training. Given that the control objective of this paper is to reduce the cost of purchasing electricity, this paper multiplies the electricity price state variable by a 2-fold amplification factor during training to increase the impact of electricity price fluctuations on the control objective. The training time step is 300s, so a training cycle is 288 steps. The results of this experiment were obtained under 100,000 training cycles (Table 1).

Table 1. Hyperparameter settings for target network training

| Hyperparameter         | value  |
|------------------------|--------|
| Batch size             | 300    |
| Learning rate          | 0.001  |
| No. of training epochs | 100000 |
| No. of layers          | 4      |



|                |                                   |
|----------------|-----------------------------------|
| No. of neurons | 24                                |
| Optimization   | Adaptive Moment Estimation (Adam) |

---

Combined with the real-time electricity price table in Fig.1, it can be seen that in 24 hours, the electricity price from 16 to 21 hours is the highest, followed by the electricity price from 5 to 6 hours. The energy storage charging and discharging strategy obtained by the DDQN algorithm is, between 0 and 2h, when the photovoltaic power generation does not work, the energy storage is discharged first, from 50% to below 20%, after the SOC of the energy storage reaches the lower limit, the External grid power generation. The energy storage begins to charge around 7 to 11 hours after the photovoltaic system starts to produce power, at which point any excess power is stored and the SOC rises from 20% to roughly 80%. The energy storage will stop working after being fully charged between 11 and 17 hours, and the excess photovoltaic power generation will be fed back to the external power grid. Then at a peak time of electricity consumption from 17 to 22 hours, when the electricity price increases, the cost of electricity increases, the energy storage is discharged during this period for the use of the microgrid load, and the SOC of the energy storage at this time is similar to that at 5 hours. level, that is, below 20%. In summary, this control strategy is to store solar energy during the day for use at night, to achieve the purpose of saving electricity purchase costs. Therefore, logically, this control strategy is reasonable. Compared with the control strategy obtained by the DDQN algorithm, the control strategy obtained by the DQN algorithm in the same training period is to discharge the energy storage in 0-1h, and then the energy storage has been in an idle state until about 14h to start charging the energy storage, and the same charging To a similar level to DDQN, that is, the SOC is about 80%, and when the price of electricity starts to increase in 18 hours, the energy storage is then released. The data results showed that the control strategy obtained by the DDQN algorithm, compared with the DQN algorithm, considers the power consumption characteristics at different times in more detail, and allocates the charge and discharge capacity more reasonably, thus achieving a better control effect than the DQN algorithm.



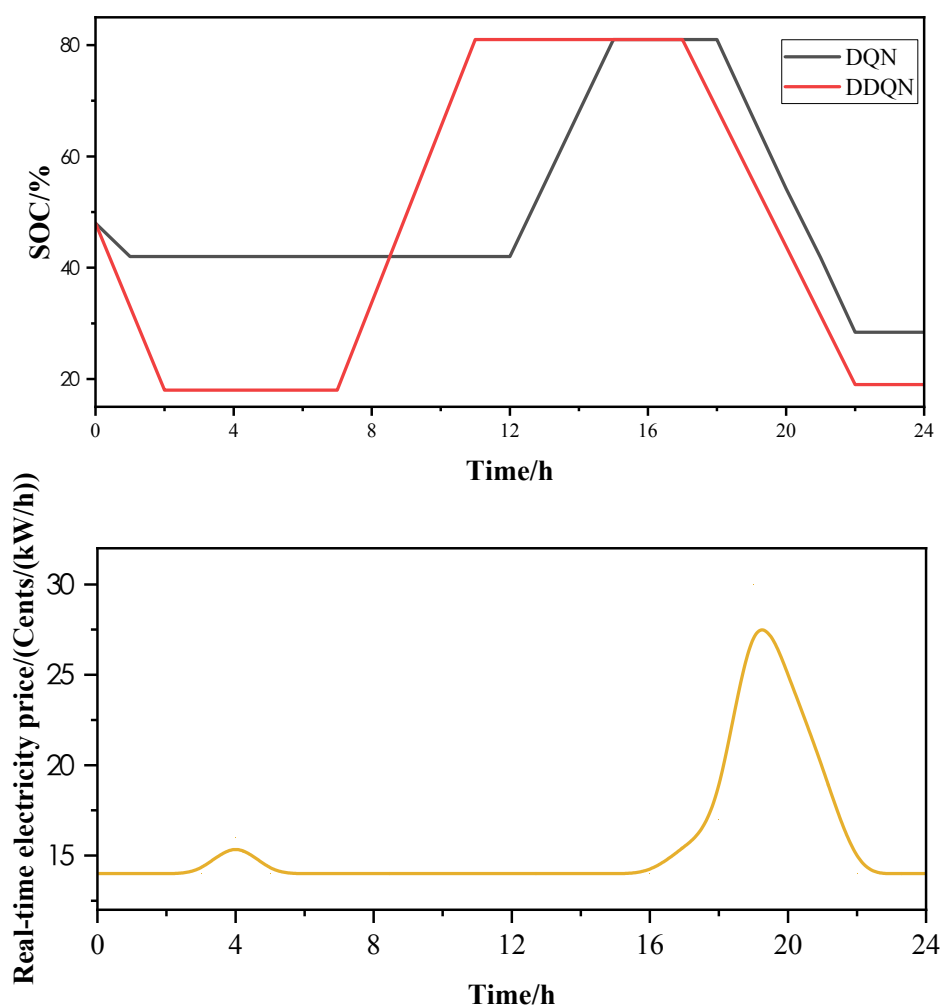


Fig.1 Comparison of energy storage SOC and real-time electricity price

The previous results give a qualitative analysis of the energy storage control strategy, and Fig. 2 and Table 2 give quantitative data results to illustrate that the control strategy trained by the DDQN algorithm is more optimized than the traditional control method and the DQN algorithm. The traditional MPPT control based on photovoltaic power generation is a simple state machine control method, that is, it judges the positive or negative power difference between the load and the power in the microgrid, and when the energy storage meets the given SOC state, according to the photovoltaic maximum power tracking principle. It is possible to obtain the energy storage's charging and discharging power. See the literature [27] for more information on the specific algorithm process.

Fig. 2 and Table 2 show the purchased electricity and the corresponding expenses in different situations. For the case without energy storage, when the electricity purchased is 4344.6 kWh, the cost is \$840.1; for the MPPT control method, when the electricity purchased is 4003.6 kWh, the cost is \$735.7; for the DQN control method, when the electricity purchased for 3833.5 kWh, the cost is \$626.9; for the DDQN

control method, when the purchased electricity is 3762.4 kWh, the cost is \$613.7. From the data in Fig. 2 and Table 2, compared with the MPPT algorithm, the control strategy obtained by the reinforcement learning algorithm training saves the electricity purchase cost, and under the same training period, the DDQN algorithm is cheaper than the DQN algorithm[28-29]. Compared with the microgrid operation without energy storage, the DDQN algorithm will save 26.95% of the electricity purchase cost, which is much greater than 12.43% of the MPPT algorithm, which fully demonstrates the effectiveness of the DDQN algorithm.

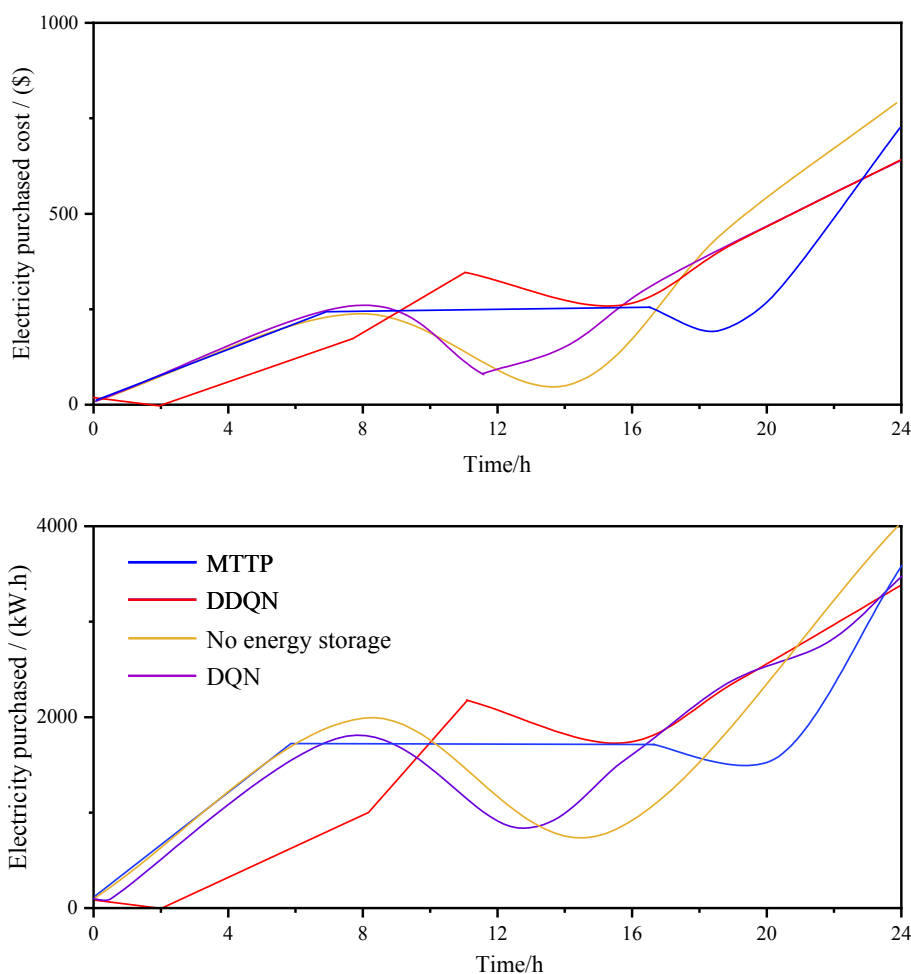


Fig 2. Comparison results of DDQN, DQN, MPPT, and no energy storage.

Table 2 Data results for DDQN, DQN, MPPT, and no energy storage

| Method            | Electricity purchased/(kW.h) | Cost / \$ |
|-------------------|------------------------------|-----------|
| No energy storage | 4344.6                       | 840.1     |
| MPPT              | 4003.6                       | 735.7     |
| DQN               | 3833.5                       | 626.9     |
| DDQN              | 3762.4                       | 613.7     |

## 4. CONCLUSION

This research presents an optimum method for the microgrid energy storage control problem based on the simulation model and the deep learning algorithm. The specific conclusions of this study are as includes:

(1) Due to the advantages of deep reinforcement learning in solving model-free problems, this study does not need to know the link between the control objectives, control variables, and state information of grid-connected microgrids. The Q-value function approximated by the neural network can be trained to identify the correlation between the control goal and the state variables using the micro-state network's data.

(2) In the same operation cycle, the energy storage discharge time of the control strategy of the DQN algorithm is 0-1h and 18-22h, while the discharge time of DDQN is 0-2h and 17-22h. The data results demonstrated that the DDQN algorithm's control strategy effect is superior to that of the DQN algorithm.

(3) The results of numerical verification show that the DDQN algorithm will save 26.95% of the electricity purchase cost, which is much greater than 12.43% of the MPPT algorithm, which fully demonstrates the effectiveness of the DDQN algorithm. The method proposed can be extended to a variety of microgrid energy control scenarios, such as adding or deleting different distributed generation modules in the simulation model, changing different control objectives, setting different control variables, and so on.

## REFERENCES

- (1) Yunhui, J., **Optimal Allocation of Distributed Energy Storage Capacity in Power Grid With High Proportion of New Energy**. *IOP Conference Series: Earth and Environmental Science*, 2021. 827(1).
- (2) Na, L., W. Chengdong and G. Cailian, **Research and Application of Energy Storage Battery on Stabilizing Fluctuation Characteristics of Photovoltaic Power System**. *International Journal of Hybrid Information Technology*, 2016. 9(5).
- (3) LvZhipeng, Y.X.S.J., **Overview on micro-grid technology**. *Proceedings of the CSEE*, 2014. 1(31): p. 57-70(in Chinese).
- (4) Amjad, A., et al., **Overview of Current Microgrid Policies, Incentives and Barriers in the European Union, United States and China**. *Sustainability*, 2017. 9(7): p. 1146.

- (5) Domínguez-Barbero, D., et al., **Optimising a Microgrid System by Deep Reinforcement Learning Techniques**. *Energies*, 2020. 13(11): p. 2830.
- (6) Liheng Liu, Miaomiao Niu, Dongliang Zhang, Li Liu and Dietmar Frank. **Optimal allocation of microgrid using a differential multi-agent multi-objective evolution algorithm**. *Applied Mathematics and Nonlinear Sciences*, 2021. 6(2): p. 111-124.
- (7) **Energy; Data on Energy Detailed by Researchers at Department of Energy (Optimization Models for Islanded Micro-Grids: A Comparative Analysis between Linear Programming and Mixed Integer Programming)**. *Energy & Ecology*, 2017.
- (8) Dolara, A., et al., **Optimization Models for Islanded Micro-Grids: A Comparative Analysis between Linear Programming and Mixed Integer Programming**. *Energies*, 2017. 10(2).
- (9) Gazijahani, F.S. and J. Salehi, **Stochastic multi-objective framework for optimal dynamic planning of interconnected microgrids**. *IET Renewable Power Generation*, 2017. 11(14).
- (10) **Energy - Renewable Energy; New Renewable Energy Study Findings Reported from Shenzhen University (Optimal Scheduling of Multiple Multi-energy Supply Microgrids Considering Future Prediction Impacts Based On Model Predictive Control)**. *Journal of Robotics & Machine Learning*, 2020.
- (11) Duchaud, J., et al., **Trade-Off between Precision and Resolution of a Solar Power Forecasting Algorithm for Micro-Grid Optimal Control**. *Energies*, 2020. 13(14).
- (12) Wang, Y., et al., **A Wasserstein based two-stage distributionally robust optimization model for optimal operation of CCHP micro-grid under uncertainties**. *International Journal of Electrical Power and Energy Systems*, 2020. 119(C).
- (13) Fossati, J.P., et al., **Optimal scheduling of a microgrid with a fuzzy logic controlled storage system**. *International Journal of Electrical Power and Energy Systems*, 2015. 68.
- (14) Provata, E., et al., **Development of optimization algorithms for the Leaf Community microgrid**. *Renewable Energy*, 2015. 74.
- (15) Fengdao, Z., B. Siyu and W. Dan, **Optimal Dispatching of Microgrid Based on Improved Particle Swarm Optimization**. *Journal of Physics: Conference Series*, 2021. 1871(1).
- (16) Yuvaraja, T. and G. M, **Artificial intelligence and particle swarm optimization algorithm for optimization problem in microgrids**. *Asian Journal of Pharmaceutical and Clinical Research*, 2015. 8(3).
- (17) Bifei, T. and C. Haoyong, **Stochastic Multi-Objective Optimized Dispatch of Combined Cooling, Heating, and Power Microgrids Based on Hybrid Evolutionary Optimization Algorithm**. *IEEE Access*, 2019. 7.
- (18) Luo, Y., et al., **Optimal configuration of hybrid-energy microgrid considering the correlation and randomness of the wind power and photovoltaic power**. *IET Renewable Power Generation*, 2020. 14(4).
- (19) Sefidgar-Dezfouli, A., M. Joorabian and E. Mashhour, **A multiple chance-constrained model for optimal scheduling of microgrids considering**

- normal and emergency operation.** *International Journal of Electrical Power and Energy Systems*, 2019. 112.
- (20) **Energy - Electric Power; Study Findings on Electric Power Are Outlined in Reports from North China Electric Power University (A Wasserstein Based Two-stage Distributionally Robust Optimization Model for Optimal Operation of Cchp Micro-grid Under Uncertainties).** *Energy Weekly News*, 2020.
- (21) Hua, H., et al., **Optimal energy management strategies for energy Internet via deep reinforcement learning approach.** *Applied Energy*, 2019. 239(APR.1): p. 598-609.
- (22) Dongxia, Z.Z.Q.C., **A coordinated control method for hybrid energy storage system in microgrid based on deep reinforcement learning.** *Power System Technolog*, 2019. 6(43): p. 1914-1921(in Chinese).
- (23) Xiaosheng, L.J.C.J., **Energy management and optimization of multi-energy grid based on deep reinforcement learning.** *Power System Technology*, 2020. 10(44): p. 3794-3803(in Chinese).
- (24) Du, Y. and F. Li, **Intelligent Multi-Microgrid Energy Management Based on Deep Neural Network and Model-Free Reinforcement Learning.** *IEEE Transactions on Smart Grid*, 2019. PP(99): p. 1-1.
- (25) Mocanu, E., et al., **On-line Building Energy Optimization using Deep Reinforcement Learning.** *IEEE Transactions on Smart Grid*, 2017: p. 1-1.
- (26) Xionglin, L.J.G.F., **Survey of deep reinforcement learning based on value function and policy gradient.** *Chinese Journal of Computers*, 2019. 6(42): p. 1406-1438(in Chinese).
- (27) Xiong, L., W. Peng and P.C. Loh, **A Hybrid AC/DC Microgrid and Its Coordination Control.** *IEEE Transactions on Smart Grid*, 2011. 2(2): p. 278-286.
- (28) Medina, R., Breña, J. L., y Esenarro, D. (2021). **Efficient and sustainable improvement of a system of production and commercialization of Essential Molle Oil (Schinus Molle).** *3C Empresa. Investigación y pensamiento crítico*, 10(4), 43-75. <https://doi.org/10.17993/3cemp.2021.100448.43-75>
- (29) Meng Siyu & Zhang Xue.(2021).**Translog function in government development of low-carbon economy.** *Applied Mathematics and Nonlinear Sciences* (1). <https://doi.org/10.2478/AMNS.2021.2.00138>