Two-level optimal scheduling strategy of demand response-based microgrids based on renewable energy forecasting

Sizhou Sun  
Anhui Polytechnic University

Yu Wang ( wangyu17007049@163.com )  
Anhui Polytechnic University

Hongtao Wang  
Ningde Normal University

Ying Meng  
Anhui Polytechnic University

Shilin Liu  
Anhui Polytechnic University

Research Article

**Keywords:** Renewable energy prediction, Two-level optimization scheduling strategy, Deep reinforcement learning, Demand response, Microgrid

**Posted Date:** July 5th, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-3074985/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Considering the influences caused by the uncertainty of renewable energy generation (REG) and load on the stable operation of microgrid (MG), a two-level optimal scheduling strategy, including upper-level model and lower-level model, of demand response-based MGs using improved deep reinforcement learning (DRL) is proposed in this study. In the two-level optimal scheduling strategy, energy optimal set points of different distributed generators in the upper-level model are optimized with the objective of the minimal operational cost of the MG, demand response based on dynamic electricity pricing mechanisms is employed to minimize the electricity cost of the consumers in the lower-level model, and the opportunity constraint is transformed into a mixed-integer linear programming to simplify the solution of the optimization scheduling model. To deal with the uncertainty of the renewable energy and load, a freshness priority experience replay deep reinforcement learning (FPER-DRL) is developed to deploy the DRL prediction model for prediction of REG and load power. Finally, the experimental results illustrate that compared with traditional scheduling models based on probability density functions, the proposed method in this paper has more accurate prediction results for load power and renewable energy output, the economic benefits of MG and power users have been also improved.

1. Introduction

With the development of industrialization, environmental pollution caused by the combustion of fossil fuels has become a major obstacle to global economic growth. Making full use of renewable energy is a new way to solve the above problems, and using clean energy is considered as a trend in the future [1, 2]. Therefore, vigorous development and utilization of renewable energy such as wind (WT) and photovoltaic (PV) is an important measure to cope with the energy crisis, climate change, and environmental pollution. Microgrid (MG) as an effective carrier of renewable energy, have attracted wide attentions from researchers. However, the randomness of renewable energy generation (REG) has a significant impact on reliability and safety of MG [3]. As is well known, an effective and accurate prediction method can improve the support for MG scheduling plans. Accurate prediction results can not only improve the utilization rate of renewable energy, but also make the power scheduling department to ensure the safe, stable and economic operation of the power grid after a high proportion of the integration of WT and PV energy. In addition, demand-side management is also considered as an effective measure to deal with the randomness of REG and load. How to coordinate REG and user demand reasonably is also a hot research topic [4]. Therefore, accurately predicting these uncertain factors and reducing the cost of MGs and consumers through demand response and optimal scheduling is an important issue.

1.1 Literature review

The uncertainty of WT, PV, and load forecasting errors has a significant impact on the scheduling and operation of MG. Therefore, how to express and quantify these uncertainty factors is key to improve the reliability and economy of MG. Common prediction methods include physical models [5, 6] and machine learning [7–9]. In physical prediction models, the predicted results are calculated based on meteorological
information obtained from numerical weather prediction (NWP) and geographical location, but the error of these information will gradually increase with the aging of hardware [10, 11]. In contrast, machine learning is robust and does not rely on data obtained from NWP. Ref. [12] proposed a new convolutional neural network (CNN) model for short-term PV output power prediction. Ref. [13] used an artificial neural network (ANN) integrated PV power generation prediction method. Ref. [14–16] used least squares support vector machine (LSSVM) to construct a mixed prediction model for predicting PV output, and the prediction accuracy of this mixed prediction model is usually higher than that of a single method [17, 18]. With the advent of the big data era, deep learning (DL) has gradually become a research hotspot. Ref. [19] proposed a hybrid ensemble learning prediction model based on the long-short-term memory (LSTM) network, aiming to predict industrial electricity load more robustly and accurately. Ref. [20] designed a deep learning method based on a combination of convolutional neural networks and gated control units to predict market electricity prices, REG, and load demand. However, these traditional supervised learning methods require good datasets for training, which are often limited in practical applications. Such machine learning-based methods often require a large amount of human intervention to provide accurate prediction results of renewable energy output and load power, thus, this method lack decision-making ability. Reinforcement learning (RL) is an artificial intelligence technology with strong decision-making capabilities, which is often applied in various optimization problems. Deep reinforcement learning (DRL) combining deep learning with reinforcement learning has good prediction performance in time series [21]. Ref. [22] proposed a modified deep learning model based on recurrent neural network (MDL-RNN) for day-ahead electricity price, load, and REG output forecasting to reduce the negative impact of uncertainty on the future environment.

In a MG, the uncertainty of REG and load power is a key factor that affects the economic and stable operation of the MG. Ref. [23] proposed an adaptive robust tri-level MG scheduling model, which used robust optimization strategy to decrease the operational cost of the MG. Ref. [24] used chance constraints to handle the uncertainty factors in the MG and builds a two-level distributed optimization model to reduce operational costs. Ref. [25] developed a deep recursive neural network model with long-short-term memory (LSTM) units to predict the total load and PV output of a community MG, then, optimized the model using particle swarm optimization to improve the system's economic and reliability performance. However, these methods have not utilized the flexibility of the load characteristics and previous researches have illustrated that demand response (DR) strategy has good capabilities in dealing with uncertainty of REG and load. Therefore, this paper also considers DR in the optimal scheduling of the MG.

1.2 Main works and novelty

In summary, the main works and contributions of this paper can be drawn as follows:

(a) A freshness priority experience replay deep reinforcement learning (FPER-DRL) method is developed for predicting REG and load demand. This approach can effectively improve the forecasting efficiency and accuracy, thus, providing more reliable data information for the scheduling of MG.
(b) Considering the characteristics of multi-objective optimization involving multiple stakeholders (the economic benefits of MG and consumers), a two-level optimization scheduling strategy based on demand response is developed with the aims of reducing the operational cost of MG and the electricity cost for consumers.

(c) In the scheduling model, a new dynamic pricing mechanism is developed to guide consumer electricity usage behavior and improve the integration rate of renewable energy, which can better improve the economic benefits.

1.3 Main structures

The remainders of the paper are organized as follows. Section 2 introduces the prediction methods of REG and load demand. Section 3 constructs a two-level optimization scheduling model and its constraints for multiple stakeholders. Section 4 describes the solving method and process of the model. Section 5 shows the effectiveness of the proposed method through experiments. Finally, Section 6 presents the conclusion.

2. Related methodology

2.1 Deep certainty policy gradient

Deep reinforcement learning (DRL) is an effective solution that deals with system perception and sequential decision problems [26]. Deep Deterministic Policy Gradient (DDPG) is a model-free, non-policy of the DRL framework. DDPG consists of two parts: the main actor network $a = \pi(s|\theta)$ and the main critic network $Q(s, a|\theta)$. The function of the main actor network is interacted with the environment, provide precise output actions $a$ based on the state $s$ (instead of a probability distribution), and then obtain the next state $s'$ and reward $r$. Update the actor network with the target of maximizing the total reward, which is expressed as Eq. (1).

\[
J(\theta^\mu) = E_{\theta^\mu}[r_1 + \gamma r_2 + \gamma^2 r_3 + \cdots + \gamma^{i-1} r_i]
\]

where $\gamma$ represents the discount factor.

The objective function is optimized to improve the total reward $J$ through stochastic gradient descent. The testing results in Ref. [27] illustrated that the learning process of DDPG was similar as Q-learning, where the gradient of the objective function is equivalent to the expected gradient of the Q-value function. Therefore, the gradient can be expressed approximately as Eq. (2).

\[
\nabla_{\theta^\mu} J \approx \frac{1}{N} \sum_i [\nabla_a Q(s, a|\theta^Q)|_{s=s_i, a=\pi(s_i)} \nabla_{\theta^\mu} \pi(s|\theta^\mu)|_{s = s_i}]
\]
Update the critic network in terms of the minimal values of the loss function. The loss function is shown as Eq. (3).

\[
L(\theta^Q) = E_{(s, a, r, s') \sim B}\left\{ [r + \gamma Q'(s', \pi'(s'|\theta^\mu')) - Q(s, a|\theta^Q)]^2 \right\}
\]

where \(B\) denotes the replay buffer. \(Q'\) denotes the target \(Q\)-value function. \(\theta^\mu'\) and \(\theta^Q'\) represent the parameters of the target actor network and the target critic network.

In order to minimize the loss function, DDPG updates the parameters of the critic network \(\theta^Q\) through gradient descent. On the other hand, to maximize the total reward and improve efficiency, the parameters of the actor network \(\theta^\mu\) are updated through gradient ascent, which can be described as Eq. (4).

\[
\begin{align*}
\theta^Q &\leftarrow \theta^Q - lr_c \nabla_{\theta^Q} L(\theta^Q) \\
\theta^\mu &\leftarrow \theta^\mu - lr_a \nabla_{\theta^\mu} L(\theta^\mu)
\end{align*}
\]

where \(lr_c\) and \(lr_a\) denote the learning rate.

After updating \(\theta^Q\) and \(\theta^\mu\), DDPG updates the target critic and actor networks through an update factor \(\tau\), expressed as Eq. (5).

\[
\begin{align*}
\theta^Q' &\leftarrow \tau \theta^Q + (1 - \tau) \theta^Q' \\
\theta^\mu' &\leftarrow \tau \theta^\mu + (1 - \tau) \theta^\mu'
\end{align*}
\]

\(\tau \ll 1\)

### 2.2 Freshness priority experience replay deep reinforcement learning (FPER-DRL)

In the training process of DDPG, the selection of the reward function is the main concern, and a suitable reward function can improve the prediction performance of the model. Therefore, there are some changes to the reward function in this paper and a novel reward function storage set is constructed, which contains the mean absolute error (MAE), mean absolute percentage error (MAPE), mean error (MSE), mean square root error (RMSE), absolute error \(\sigma\) and coefficient of determination \(R^2\), which are expressed as Eq. (6).
\[
\begin{align*}
-\text{MAE} &= -\frac{1}{N} \sum_{i=1}^{N} | P_{i,t} - \hat{P}_{i,t} | \\
-\text{MAPE} &= -\frac{1}{N} \sum_{i=1}^{N} | (P_{i,t} - \hat{P}_{i,t}) / P_{i,t} | \\
-\text{MSE} &= -\frac{1}{N} \sum_{i=1}^{N} (P_{i,t} - \hat{P}_{i,t})^2 \\
-\text{RMSE} &= -\sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_{i,t} - \hat{P}_{i,t})^2} \\
-|\sigma| &= -| P_{i,t} - \hat{P}_{i,t} | \\
R^2 &= 1 - \frac{\sum_{i=1}^{N} (P_{i,t} - \hat{P}_{i,t})^2}{\sum_{i=1}^{N} (P_{i,t} - \bar{P}_{i,t})^2}
\end{align*}
\]

(6)

where \( N \) is the number of samples, \( P_{i,t} \) is the actual value, \( \hat{P}_{i,t} \) is the prediction result, \( \bar{P}_{i,t} \) denotes the prediction result.

For there are more hyperparameters in the DRL algorithms that need to be optimized, which usually influence the efficiency [28], thus, their efficiency is often low, and hyperparameter optimization can be extremely expensive. To simplify DRL models and improve operational efficiency, a sequence-based Bayesian optimization algorithm - Metis is applied to determine the deployment of DRL models and tune the hyperparameters [29]. The actor part makes prediction using the input data, and then the critic part evaluates the prediction results and the optimal solutions are obtained.

Considering that the complex training process of the DRL algorithms and the large number of hyperparameters that need to be optimized, the freshness prioritized experience replay (FPER) has been proposed to improve DDPG [30]. FPER introduces a freshness discount factor \( \mu \) to evaluate each experience, changing the sampling probability of different experiences. This significantly reduces training time and improves algorithm efficiency. Traditional prioritized experience replay (PER) speeds up the agent's selection of correct actions by replaying extreme experiences, the value of experience is usually evaluated using TD-error, which is expressed as Eq. (7).

\[
\delta_i = R_i + \varsigma Q(S_{i+1}, \arg \max_{A_{i+1}} Q(S_{i+1}, A_{i+1})) - Q(S_i, A_i)
\]
Experiences with larger positive TD-errors and smaller negative TD-errors are considered valuable. Therefore, these experiences have a higher probability of being sampled during replay. However, many experiences can also lead to a decrease in sampling efficiency. To address this problem, PER introduces a replay buffer to limit the storage of experiences. However, for two similar experiences, the experience with fewer repetitions may squeeze out valuable information. In these cases, the introduction of a freshness discount factor can improve performance. FPER defines the relationship between the freshness of an experience and its replay frequency, that freshness gradually decreases with the replay frequency. The priority of experience $p_i$ in FPER is expressed as Eq. (8).

$$p_i = \mu^{C(i)} |\delta_i| + \varepsilon$$

where $\mu$ denotes the freshness discount factor, $C(i)$ is the number of times that experience $i$ has been replayed and $\varepsilon$ is a positive constant.

The probability of sampling experience is expressed as Eq. (9).

$$P(i) = \frac{p_i^\alpha}{\sum_k p_k^\alpha}$$

where $\alpha$ represents the priority level.

To prevent multiple sampling of high TD-error experiences, the sampling weights are introduce in the traditional method, which is expressed as Eq. (10).

$$W_i = \left(\frac{1}{N} \cdot \frac{1}{P(i)}\right)^\beta$$

where $N$ denotes the buffer size and the parameter $\beta$ represents the degree of correction.

Assumed a replay buffer of storage length $N$, each experience will be updated at each iteration, and the valuable experience will be replaced by a new one. To prevent such events, the concept of a lifetime pointer is introduced in the algorithm, the new experience replaces the lowest absolute TD-error experience and does not substitute for a valuable experience.

As shown in Fig. 1, the lifetime pointer always points to the position of the lowest absolute TD-error, and worthless experiences will be squeezed out when new ones enter the replay buffer, this makes the experience of high TD-errors to be replayed multiple times.

The Pseudo of the integration algorithm for FPER-DRL is shown in Algorithm I.
**Algorithm I: DRL with freshness prioritized experience replay**

**Input**: learning rate $\eta$, replay buffer size $N$, freshness discounted factor $\mu$, exponent $\alpha$, sampling weight parameters $\beta$.

**Initialization**: actor network $a = \pi(s|\theta^\mu)$, critic network $Q(s, a|\theta^Q)$, replay memory $H = \emptyset$, count array $C$, lifetime pointer $L$, $\Delta = 0$, reward function.

1: **for** trail = 1, ..., $M$ **do**

2: Select new neural network architecture $\theta^\mu$, $\theta^Q$ according to the hyperparameters.

3: Select reward function.

4: Select an initial state $s_t$

5: **for** $t = 1, ..., H$ **do**

6: Select action $a_t$ according to the new policy.

7: Obtain reward $r_t$ and new state $s_{t+1}$

8: Store experience ($s_b$, $a_b$, $r_b$, $s_{t+1}$) in position $L$ with maximal priority $p_t = \max_{i<t} p_i$

9: Reset $C(L) = 0$

10: **if** $t > N$ **then**

11: $L$ points to the position with lowest absolute TD-error

12: **for** $i = 1, ..., K$ **do**

13: Sample experience $j \sim P(j) = pa_i / \sum_i pa_i$

14: Update count array $C(i) \leftarrow C(i) + 1$

15: Compute importance-sampling weight $W_i$ and TD-error $\delta_i$.

16: Update experience priority $p_i \leftarrow \mu C(i) |\delta_i| + \epsilon$

17: Accumulate weight-change $\Delta \leftarrow \Delta + W_i \delta_i \nabla_\phi Q(s, a, \phi)$

18: **end for**

19: Update main critic network according to the minimal loss function value.

20: Update main actor network.

21: Update target network.

22: **end if**

23: **end for**
Algorithm I: DRL with freshness prioritized experience replay

24: Collect MAPE values and update them in Metis Tuner.

25: \textbf{end for}

26: Determine the optimal neural network architecture $\theta^\mu, \theta^Q$ and the optimal policy $\pi$ in the terms of minimal MAPE value.

27: return $\theta^\mu, \theta^Q, \pi$

2.3 Multi-input multi-output strategy based on FPER-DRL

Multi-period forecasting is mainly divided into single-step forecasting and multi-step forecasting. Studies have shown that although single-step forecasting is more accurate than multi-step forecasting, it cannot timely update the real data of each period during the forecasting process. Therefore, it cannot provide complete and timely prediction results for the day-ahead scheduling in practical engineering applications. Common multi-step forecasting methods include direct strategy, recursive strategy, direct-recursive strategy, multi-input multi-output strategy, and direct-multi-output strategy. The direct strategy involves creating a new model for each prediction step, which consumes a significant amount of time. The recursive strategy cannot guarantee accuracy as the number of prediction steps increases. The direct recursive strategy can overcome the limitations of the previous two approaches, but it requires a significant computational effort. The direct multiple-output strategy is complex and requires a substantial amount of data to avoid the problem of overfitting. Considering the amount of historical data and the forecasting method used in this study, a multi-input multi-output strategy is adopted for forecasting research. The structure of the multi-input multi-output strategy is shown in Fig. 2.

From Fig. 2, it can be seen that the multi-input and multi-output strategy preserves the dependencies between predicted values and avoids the conditional independence assumption of the direct strategy or the cumulative error of the recursive strategy. Research has shown that this strategy has been successfully applied in multiple time series predictions [31].

2.4 Uncertainty modelling of prediction errors

2.4.1 The models of prediction errors

The TLS distribution is an extension of the $T$ distribution, and the probability distribution of the prediction errors of FPER-DRL follows the TLS distribution [32]. The expression can be described as Eq. (11).

$$f(x; \mu, \zeta, v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\zeta \sqrt{\pi v} \Gamma\left(\frac{v}{2}\right)} \left[1 + \left(\frac{x-\mu}{\zeta\sqrt{v}}\right)^2\right]^{-v+1/2}$$
where $\Gamma$ denotes the gamma function. $\mu$, $\zeta$ and $\nu$ denote the mean, standard deviation and shape factor respectively.

### 2.4.2 The revision of forecast results

In order to obtain more accurate prediction results, the prediction value is obtained by FPER-DRL, then, the prediction error expectation value obtained by the distribution of the prediction error, and finally correct the prediction knot using TLS distribution, which is expressed as Eq. (12).

\[
R(P_{i,t}^{pred}) = P_{i,t}^{pred} - E(\sigma_{i,t}^i), i \in \{WT, PV, Load\}
\]

where $R(P_{i,t}^{pred})$ denotes the corrected forecast result of the uncertainty factor. $P_{i,t}^{pred}$ and $E(\sigma_{i,t}^i)$ represent the forecasting result and forecasting error, respectively.

### 3. The two-level optimal model for a grid-connected micro-grid

By analyzing the characteristics of optimizing economic benefits for MGs and consumers, a two-level scheduling model for demand response-based MGs containing renewable energy sources is developed. The upper-level model takes the economic benefits of the MG as objective function, while the lower-level model uses the economic benefits of the consumers as the objective function. Among the upper-level model and lower-level model, the electricity plan determined by the MG and the electricity price mechanism are used as a bridge.

The grid-connected MG includes wind turbine (WT), photovoltaic (PV), micro gas turbine (MT), energy storage system, main grid and energy management system (EMS), which is shown as Fig. 3.

#### 3.1 Dynamic electricity price mechanism based on demand response

In order to coordinate the scheduling problems between the upper-level model and the lower-level model, reduce the pressure of power demand during the peak periods of the power system and the cost of electricity for customers., a new pricing mechanism as a bridge between the two-level model is developed.

Compared with the traditional time-of-use electric price, the dynamic electric price mechanism can respond the dynamic supply and demand of the system in real time, and the flexibility of the load can be better utilized. The dynamic electricity price mechanism proposed in this study can reduce the operational cost of the upper and low levels while guiding consumers to adjust their electricity consumption plans. The specific description is expressed as Eq. (13).

\[
RTP_K = a_k P_{load,t}^2 + b_k P_{load,t} + c_k
\]
where \( P_{\text{load},t} \) is the net load of the MG at time period \( t \). \( a_k, b_k \) and \( c_k \) are real-time electric price coefficients that values at different time periods depend on the demand dynamics of the customer.

Three electricity price levels are introduced in the pricing mechanism, which is expressed as Eq. (14).

\[
IBR_k = \begin{cases} 
RTP_k & 0 \leq P_{\text{load},t} \leq \delta^1_k \\
\lambda_1 \cdot RTP_k & \delta^1_k \leq P_{\text{load},t} \leq \delta^2_k \\
\lambda_2 \cdot RTP_k & P_{\text{load},t} \geq \delta^2_k 
\end{cases}
\]  

(14)

The expression for the new dynamic tariff \( Pr(P_{\text{load},t}) \) is expressed as Eq. (15).

\[
Pr(P_{\text{load},t}) = \begin{cases} 
RTP_{re} & P_{\text{load},t} \leq 0 \\
RTP_k & 0 < P_{\text{load},t} \leq \delta^1_k \\
\lambda_1 \cdot RTP_k & \delta^1_k \leq P_{\text{load},t} \leq \delta^2_k \\
\lambda_2 \cdot RTP_k & P_{\text{load},t} > \delta^2_k 
\end{cases}
\]  

(15)

where \( \delta^1_k \) and \( \delta^2_k \) denote bounds between different electricity prices. \( \lambda_1 \) and \( \lambda_2 \) are price multipliers, \( \lambda_2 > \lambda_1 > 1 \). When \( P_{\text{load},t} \leq 0 \), it means that there is surplus renewable energy output and the MG can sell the excessive electricity to the main grid, and \( RTP_{re} \) represents the price of a sale to the main grid.

As shown in Fig. 4, MGs offer preferential electricity prices that guide consumers to change their electricity consumption plans and actively involve in demand response. Some of the load is shifted out of peak periods and moved into low periods. Consumers reduce the electricity cost while relieving pressure on the power system.

3.2 The upper-level model

The upper-level model is optimized according to the economic and environmental benefits of MG. The MG needs to be reasonably modeled before optimization.

3.2.1 Objective function

The upper-level model includes operation and maintenance costs of distributed power sources in MGs, the cost of sale and purchase of electricity between the MG and the main grid, the cost of losses of energy storage units and the cost of pollutant emissions from individual distributed generation units.
\[
\min F_{UP} = \sum_{t=1}^{T} \left[ \sum_{n=1}^{N} \left( C_{n,t}^{MT} + C_{n,t}^{WT} + C_{n,t}^{PV} \right) + C_{t}^{grid} + C_{t}^{R} \right]
\]

where \( T \) denotes the total scheduling time period, \( CMT_{n,t}, CWT_{n,t} \) and \( CPV_{n,t} \) represent the generation cost of MT, WT and PV, respectively, \( N \) is the number of distributed generation subunits of each type, \( C_{grid} t \) and \( CR t \) denote the main grid power purchase cost and energy storage cost respectively.

### 3.2.2 Constraint conditions

**a) Supply and demand balance constraints**

The power of the MG and the load of the users need to satisfy the following constraints at each moment, expressed as Eq. (17).

\[
E(P_{t}^{RGs}) + P_{t}^{DG} + P_{t}^{ESS} = R(P_{EL,t}^{pred})
\]

where \( E(P_{t}^{RGs}) \) represents the supply and demand of the MG, \( PDG t \) is the output power of the distributed generation unit, \( PESS t \) denotes the power of ESS charging and discharging.

**b) Power generation constraint**

The output power of each distributed generation unit must be within a safe power range, expressed as Eq. (18).

\[
P_{\text{min},t}^{DG} \leq P_{t}^{DG} \leq P_{\text{max},t}^{DG}
\]

where \( PDG t \) denotes the power of output unit.

**c) Battery constraints of ESS**

ESS needs to meet the following constraints when exchanging power with MGs, which are expressed as Eq. (19), (20), (21) and (22), respectively.

\[
C_{t+1}^{ESS} = C_{t}^{ESS} + (\eta_{ch} P_{CH,t}^{ESS} - P_{DC,t}^{ESS} / \eta_{dc}) \Delta t, \forall t
\]

\[
0 \leq P_{DC,t}^{ESS} \leq P_{DC,\text{max}}^{ESS}
\]

\[
0 \leq P_{CH,t}^{ESS} \leq P_{CH,\text{max}}^{ESS}
\]

\( \forall t \) (20)
where $CESS_t$ denotes the storage capacity of ESS, $\eta_{ch}$ and $\eta_{dc}$ are the charging and discharging efficiency, $PESS\ CH,t$ and $PESS\ DC,t$ are the charging and discharging power of ESS, and $CESS\ 24$ and $CESS\ D$ denote the initial capacity of ESS and the capacity at the end of the scheduling cycle, respectively.

### 3.3 The lower-level model

The study shows that DR can make the power profiles smoother by shifting or interrupting the load during peak periods of electricity consumption.

#### 3.3.1 Objective function

The goal of the lower-level model is to minimize consumer spending, expressed as Eq. (23).

$$\min F_{down} = P_r(P_{load}) \sum_{t=1}^{T} (P_{shift,t} + P_{unshift,t})$$

where $P_{shift,t}$ and $P_{unshift,t}$ denote the output power of time-shiftable loads and the output power of non-time-shiftable loads, respectively.

#### 3.3.2 Constraint conditions

a) Power supply and demand balance

The load needs to satisfy the power balance constraint, which is expressed as Eq. (24).

$$P_{shift,t} + P_{unshift,t} = P_{load,t}, \forall t$$

where $P_{load,t}$ denotes the total consumer load.

b) Constraint of time-shiftable loads

The battery constraint of EVs is similar as ESS, which is expressed as Eq. (25).

$$P_{shift,min,t} \leq P_{shift,t} \leq P_{shift,max,t}, \forall t$$
where $P_{\text{shift min},t}$ and $P_{\text{shift max},t}$ denote the minimum and maximum output power of the time-shiftable load respectively.

c) Spinning reserve constraint

In extreme cases, MG need to provide sufficient spinning reserve to reduce the impact of stochastic factors on the system [33]. The spinning reserve constraint of the MG system is provided by the main power grid, energy storage systems (ESS), which can be represented as an opportunity constraint, expressed as Eq. (26).

$$
Pr \left\{ R_{\text{grid},t} + P_{\text{Res},t} \geq E(\sigma_{E}^{EL}) - (\sigma_{t}^{L} - \sigma_{t}^{WT} - \sigma_{t}^{PV}) \right\} \geq \alpha, \forall t
$$

where $R_{\text{grid},t}$ denotes the spinning reserve provided by the main grid, $P_{\text{Res},t}$ represents the spinning reserve provided by the generator and ESS, $\alpha$ is the confidence level, and $E(\sigma_{E}^{EL} t)$ denotes the expected value of the EL prediction error.

4. Solving method

4.1 Serialization modeling of random variables

The uncertainty will gradually increase with the increase of renewable energy's role in power generation. Sequential arithmetic theory (SOT) has been successfully applied in various engineering fields as an effective method to deal with multiple complex uncertainties. The advantage of this method is that the inverse function of the cumulative distribution function does not need to be solved, opportunity constraints can be transformed into deterministic equivalence classes by sequential operations [34]. The probability sequence can be expressed as Eq. (27) [35].

$$
\sum_{i=0}^{N_{a}} a(i) = 1, a(i) \geq 0, i = 0, 1, 2, \ldots, N_{a}
$$

where $a(i)$ represents the discrete sequence of length $N_{a}$.

After discrete processing of the continuous probability release, $a(i_{a,t})$, $b(i_{b,t})$, $c(i_{c,t})$ can represent the power of PV, WT, and load, respectively. Taking WT energy as an example, the sequence length of the WT power output $N_{a,t}$ which is expressed as Eq. (28).

$$
N_{a,t} = \lfloor P_{\text{max},t}^{WT}/q \rfloor
$$
where \( q \) denotes the discretization step. \( PWT \ max, t \) is the maximum output power of the WT.

The probability sequence of PV output, which are expressed as Eqs. (29) and (30).

\[
a(i_{a,t}) = \begin{cases} 
\int_{i_{a,t} \cdot q/2 + \sigma_{WT}^{\min,t}}^{q/2 + \sigma_{WT}^{\min,t}} f_{o}(\sigma_{WT}) \, d\sigma_{PV}, & i_{a,t} = 0 \\
\int_{i_{a,t} \cdot q/2 + \sigma_{WT}^{\min,t}}^{i_{a,t} \cdot q/2 + \sigma_{WT}^{\min,t}} f_{o}(\sigma_{WT}) \, d\sigma_{PV}, & i_{a,t} > 0, i_{a,t} \neq N_{a,t} \\
\int_{i_{a,t} \cdot q + \sigma_{WT}^{min,t}}^{i_{a,t} \cdot q + \sigma_{WT}^{max,t}} f_{o}(\sigma_{WT}) \, d\sigma_{PV}, & i_{a,t} = N_{a,t}
\end{cases}
\]

\[
N_{a,t} = \left[ \frac{\sigma_{WT,t}^{max} - \sigma_{WT,t}^{min}}{q} \right]
\]

where \( f_{o}(\sigma_{WT}) \) is the probability density function of WT power prediction error, \( \sigma_{WT \ min,t} \) and \( \sigma_{WT \ max,t} \) represent the minimum and maximum values of WT output power prediction error.

The probability sequences of the PV power output prediction error \( N_{b,t} \) and the load prediction error and \( N_{d,t} \) denote \( b(i_{b,t}) \) and \( d(i_{d,t}) \) respectively.

In this paper, the output power of PV and WT and load power is defined as equivalent load (EL). The probability sequence of the prediction error \( c(i_{c,t}) \) of the output of PV and WT, which is expressed as Eq. (31) [36].

\[
c(i_{c,t}) = a(i_{a,t}) \oplus b(i_{b,t}) = \sum_{i_{a,t} + i_{b,t} = i_{c,t}} a(i_{a,t})b(i_{b,t}), i_{c,t} = 0, 1, \ldots, N_{a,t} + N_{b,t}
\]

where \( b(i_{b,t}) \) denotes the probability series of the prediction error with the PV power.

Through the subtraction-type-convolution operation of SOT, the probability sequence \( e(i_{e,t}) \) of EL power, which is expressed as Eq. (32) [37].
where \(N_{e,t}\) is the length of \(e(i_{e,t})\).

The prediction error and probability sequence of EL power are shown in Table 1. The expected value of \(e(i_{e,t})\) is expressed as Eq. (33).

\[
E(P_t^{EL}) = \sum_{i_{e,t}=0}^{N_{e,t}} [(\sigma_{min}^{EL,t} + i_{e,t}q) \cdot e(i_{e,t})]
\]

### Table 1
**Probability sequence of EL power prediction error**

<table>
<thead>
<tr>
<th>Error(kW)</th>
<th>(\sigma EL_{min,t})</th>
<th>(q + \sigma EL_{min,t})</th>
<th>(\ldots)</th>
<th>(i_{e,t}q + \sigma EL_{min,t})</th>
<th>(\ldots)</th>
<th>(N_{e,t}q + \sigma EL_{min,t})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>(e(0))</td>
<td>(e(1))</td>
<td>(\ldots)</td>
<td>(e(i_{e,t}))</td>
<td>(\ldots)</td>
<td>(e(N_{e,t}))</td>
</tr>
</tbody>
</table>

In Table 1, the prediction error \(i_{e,t}q + \sigma EL_{min,t}\) of EL always has the corresponding probability \(e(i_{e,t})\).

### 4.2 Handling of chance constraint

In order to facilitate the treatment of probability sequences, a 0–1 variable \(Z_{i_{e,t}}\) is set in this paper, expressed as Eq. (34).

\[
Z_{i_{e,t}} = \begin{cases} 1, & \sum_{n=1}^{MG} R_{n,t} + P_{Res,s,t} \geq E(\sigma_t^{EL}) - (\sigma_{min,t}^{EL} + i_{e,t}q), \forall t \\ 0, & otherwise \end{cases}
\]

As shown in Table 1, the probability of prediction error \(i_{e,t}q + \sigma EL_{min,t}\) of EL power is \(e(i_{e,t})\), thus, Eq. (28) can be replaced by Eq. (35)

\[
\sum_{i_{e,t}=0}^{N_{e,t}} Z_{i_{e,t}}e(i_{e,t}) \geq \alpha
\]
4.3 Determination of dispatching scheme

Through the iterative optimization of the upper and lower levels, multi-group scheduling scheme and the costs of the upper and lower levels for each group of schemes can be obtained. In order to determine the optimal scheduling solution, the fitness function is calculated according to Eq. (36) and the final optimal scheduling solution can be determined using the minimum cost.

\[ F_{finally} = \min \sqrt{(F_{up})^2 + (F_{down})^2} \]

4.4 Solving process

The prediction model can be transformed into a mixed integer programming form after chance constraint processing. The working flow chart of the optimal scheduling model proposed in this study is shown in Fig. 5.

From Fig. 5, the detailed solution process of the proposed two-level scheduling model is shown as follows:

Step 1: Input the historical data of WT, PV and load.

Step 2: Construct a prediction model based on the input data.

Step 3: Use FPER-DRL to obtain the probability distribution of the WT, PV and load prediction values and errors, then, generate a probability distribution.

Step 4: Use SOT to convert opportunity constraints into deterministic equivalence classes and obtain the MILP model.

Step 5: Input initial electricity prices and operational constraints.

Step 6: Solve the model using CPLEX method.

Step 7: Check whether a solution has been found. If yes, execute the next step. Otherwise, update the load and confidence level, and go back to step 5.

Step 8: Obtain the optimal scheduling plan of the upper-level model and use a dynamic pricing mechanism to determine the dynamic electricity prices.

Step 9: Construct a lower-level scheduling model.

Step 10: Input dynamic electricity prices and parameters, and solve the model using CPLEX.
Step 11: Obtain the final scheduling plan based on Eq. (38).

Step 12: Check whether the termination condition is met. If yes, proceed to the next step. Otherwise, update the user plan and go back to step 5.

Step 13: Output the final scheduling plan and economic costs.

5. Case studies

5.1 Description of testing system

To evaluate the accuracy of the proposed prediction method, 80% of the original dataset is used to train the model and the others are used to test the model. The parameters of the power output unit and energy storage unit in the scheduling model are as follows: The minimum output and the rated power of MT are 10kW/h and 65kW/h respectively. The maximum capacity of the energy storage unit is 160kW/h, the maximum charge/discharge power is 40kW/h, and the range of use is 20% ~ 90%.

5.2 The results of Renewable energy and load forecasting

As we all know, accurate prediction results have a strong influence on the optimal scheduling of MG. However, renewable energy output and customer loads are uncertain and stochastic. To address the prediction problem of these uncertainties, the FPER-DRL is applied to construct a specific model framework with optimized hyperparameters using the historical data of WT, PV and load.

The forecasting results of WT, PV and load by different models are shown in Fig. 6, Fig. 7 and Fig. 8. Comparing the results of the proposed prediction method with that of the other traditional methods, including multivariate linear regression (MLR) and LSTM. It intuitively reflects the fitting degree of the prediction results with the actual data.

As shown in Fig. 6, Fig. 7 and Fig. 8, the prediction results of MLR and LSTM will have stronger prediction errors when the data have large fluctuations. In contrast, the prediction method in this paper has a small prediction error. In conclusion, the prediction curve of the FPER-DRL method fits the actual power curve better, it can provide reliable data support for MG optimization scheduling.

Figure 9 displays the MAPE and RMSE evaluation metrics of the proposed method for WT, PV, and load prediction at different prediction steps.

As shown in Fig. 9, the prediction models show different predictions on the experimental results of WT, PV, and load. The load has fewer influencing factors, resulting in smaller errors and higher accuracy compared with WT and PV, which indicates that the randomness of WT and PV is stronger than loads. Meanwhile, due to the inevitable accumulation of prediction errors in multi-step prediction, the prediction error increases gradually with the increase of prediction steps.
5.3 Two-level optimal scheduling based on prediction results

Based on the aforementioned prediction method for REG and load, the predicted results are used for optimal scheduling of two-level model. The upper-level and lower-level models use real-time electricity prices based on DR and consumers' electricity plans as a bridge, and through iterative optimization, obtain an optimization result for MG's operational costs and consumers' electricity costs. Figure 10 and Fig. 11 respectively show the results of scheduling schemes with and without considering DR. Figure 10 shows the load shifting after the DR strategy is applied.

As shown in Fig. 10 and Fig. 11, some conclusions: can be obtained. (1) The proposed dynamic pricing mechanism can effectively guide consumers to actively participate in MG scheduling, which maximizes the consumption of renewable energy output and reducing MG operational costs. (2) Due to the implementation of demand response plans, the power output of MT and the cost of MG purchasing electricity from the main grid have been significantly reduced. (3) Regarding ESS, the scheduling plan reasonably utilizes the energy storage capacity of ESS, reducing the waste of REG.

From Fig. 12, it can be seen that after the implementation of demand response, the user's load is shifted out during periods of high electricity consumption and shifted in the non-peak periods, making the load power curve smoother. In summary, the implementation of demand response realizes load transfer, encourages users to actively participate in MG scheduling, and improves the reliability and economic efficiency of MG operation.

To address the impact of renewable energy's randomness with MG, spinning reserve configuration is one of the important methods to ensure the stability of MG's power. However, increasing the spinning reserve will lead to higher operational costs of MG, which will not be beneficial to the economic operation of MG. Figure 13 shows the optimization of MG's operational costs, consumers' electricity costs, and the optimization at different confidence levels.

As shown in Fig. 13, choosing the appropriate confidence level has a significant impact for the economic and reliable operation of the system. As the confidence level increases, MG needs more spinning reserves to deal with the adverse effects of renewable energy sources' randomness, the system's reliability also improves, but the economic efficiency will inevitably decrease. On the other hand, if the confidence level decreases, the system's reliability will decrease and the operational cost will increase. In addition, the proposed two-level optimization achieves a good compromise solution compared with the independent optimization of the upper-level and low-level models. This demonstrates that this method can balance the benefits of multiple stakeholders.

6. Conclusions
This article proposes a FPER-DRL method for predicting REG and load, which is then used in a two-level optimization scheduling model to reduce operational costs of MGs and electricity costs for consumers. Based on this, a dynamic pricing mechanism DR plan is designed to guide consumers’ behavior of electricity consumption by utilizing the flexibility of load. Experimental results demonstrate that:

1. Compared with the traditional time series prediction method, the FPER-DRL designed in this paper is more accurate and more practical.
2. The two-level optimization scheduling model based on DR, balances the interests of MG and consumers through real-time feedback. The implementation of DR strategy allows consumers to participate in the optimization scheduling of MG, which reduces the impact of renewable energy on MG and improves the stability of the power system, and simultaneously enhances the economics of multiple stakeholders.
3. The developed dynamic pricing mechanism indirectly encourages consumers to actively participate in DR programs. This greatly improves the utilization of renewable energy and reduces the power burden for the MG during peak electricity consumption periods.

Declarations

Acknowledgments

The authors would like to thank the anonymous reviewers for their constructive comments.

Author contributions

Sizhou Sun: Conceptualization, Investigation, Writing; Yu Wang: Methodology, Software, Data curation; Hongtao Wang and Ying Meng: Validation and Investigation; Shilin Liu: Supervision.

Funding

This work was supported by Anhui Provincial Key Research and Development Program (202004a05020014), the Open Research Fund of Energy Internet Engineering Research Center of Anhui Provincial Department of Education, Anhui Polytechnic University (2021EIRC04ZD and 2021EIRC09YB). Scientific Research Project of Anhui Polytechnic University (Xjky2022041). The Open Research Fund of Anhui Key Laboratory of Detection Technology and Energy Saving Devices, Anhui Polytechnic University (JCKJ2021B08).

Availability of data and materials

Most of the data used in the study is publicly available, and information about the sources of the data is provided in the references.

Ethics approval and consent to participate

Not applicable.
Consent for publication

Not applicable.

Competing interests

The authors declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work. There is no conflict of interest exiting in the submission of this manuscript which has been approved by all authors for publication.

Author details

1 Anhui Key Laboratory of Electric Drive and Control, Anhui Polytechnic University, Wuhu, 241000, China;

2 School of Information Mechanical and Electrical Engineering, Ningde Normal University, Fujian 352100, China.

References


Figures
**Figure 1**

Replay buffer with lifetime pointer

```
1 2 ... N-1 N  \rightarrow  N+1 N+2 ... N+H-1 N+H
```

**Figure 2**

The structure of the Multi-Input Multi-Output strategy.
Figure 3

The structure of grid-connected MG

Figure 4

The curve of the real-time electricity price mechanism

Figure 5
The working flow chart of MG scheduling model based on data prediction

Figure 6

Forecasting results of loads
Figure 7

Forecasting results of WT
Figure 8

Forecasting results of PV
Figure 9

FPER-DDPG error metrics for different prediction steps in load, WT and PV

Figure 10

Scheduling results without DR
Figure 11

Scheduling results with DR
Figure 12

Load shifting under DR
Figure 13

Operational costs of upper-level and low-level under different confidence levels