

Green Energy Microgrid Dispatching: A Hybrid Framework of Improved Genetic Algorithm and Reinforcement Learning

Shengke Zhuo

Key Laboratory of Mathematical MIT and School of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, China
Email: sx2408037@nuaa.edu.cn

How to cite this paper: Zhuo, S.K. (2026) Green Energy Microgrid Dispatching: A Hybrid Framework of Improved Genetic Algorithm and Reinforcement Learning. *Journal of Applied Mathematics and Physics*, 14, 3716-3741.
<https://doi.org/10.4236/jamp.2026.149183>

Received: May 27, 2026

Accepted: September 26, 2026

Published: September 29, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).
<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

High renewable penetration turns microgrid dispatching from an optimization problem into an uncertainty management one. Solar and wind outputs flicker; load demand wavers without rhythm. Both economic efficiency and operational stability come under threat. Conventional methods fall into two camps: model-based approaches, brittle against parameter errors, and search-heavy algorithms, too sluggish for real-time shifts. Neither manages to reconcile global exploration with instantaneous responsiveness—the very balance that dynamic environments demand. To address these issues, this paper proposes a hybrid intelligent optimization framework that integrates an improved genetic algorithm and deep reinforcement learning, aiming to achieve real-time economic dispatching of microgrids with a high proportion of renewable energy. At the algorithmic level, an adaptive hybrid encoding strategy is designed to unify the representation of continuous control variables and discrete equipment states; further, a dynamic mutation mechanism based on deep reinforcement learning is introduced, enabling the algorithm to adjust the mutation probability in real time according to the degree of load disturbance, enhancing the system's robustness in sudden conditions. At the optimization mechanism level, a DRL-driven fitness evaluation system is constructed: the deep reinforcement learning agent obtains real-time reward signals through interaction with the environment and uses these signals to dynamically modify the fitness function of the genetic algorithm, thereby achieving collaborative optimization of global exploration and local fine search. Simulation results based on the NREL OpenEI Commercial Building and Microgrid Reference Datasets (2018-2022) and the IEEE 33-node system show that, compared with traditional genetic algorithms, Deep Q-Network (DQN), and particle swarm algorithms, the proposed method performs exceptionally well in multiple key

indicators: compared to traditional genetic algorithms, the average daily comprehensive operating cost is reduced by 16.3%; compared to particle swarm algorithms, the penetration rate of renewable energy increases by 8.7%; the convergence speed is increased by 20% (based on the genetic algorithm as the benchmark). Moreover, in load disturbance scenarios, the dynamic mutation strategy reduces the cost fluctuation amplitude by 25.0%, demonstrating excellent dynamic adaptability and robustness.

Keywords

Hybrid Optimization Framework, Improved Genetic Algorithm, Reinforcement Learning, Microgrid Scheduling, Renewable Energy Penetration Rate, Dynamic Mutation Strategy, Robust Optimization

1. Introduction

Decarbonization is now the common goal of major economies—China’s carbon-neutrality drive, the EU’s Green Deal, and the US rejoining the Paris Agreement all attest to that. In 2023, global renewable capacity overtook fossil fuels for the first time, and the IEA expects clean energy to supply nearly half of world electricity by 2030. China, the world’s largest energy producer and consumer, has long led in wind and solar installations.

But this expansion introduces a grid stability problem that conventional dispatch cannot easily solve. Unlike fossil plants, renewables are inherently intermittent—and increasingly so at sub-minute timescales. A passing cloud can slash solar output by over 50% in mere seconds, compressing uncertainty from the minute level to the second level. For microgrids, that shift turns routine power balancing into a high-frequency, real-time control challenge—one that fundamentally redefines operational reliability.

The large-scale integration of distributed renewable energy has profoundly transformed the traditional “generation-transmission-distribution-consumption” one-way radial structure. Microgrids, as the key technical carrier for integrating distributed energy and achieving autonomous management, embody their core value in the following aspects: aggregating dispersed resources to reduce the impact on the main grid, smooth switching between grid connection and island mode to ensure power supply for critical loads, and providing technical support for new business models such as demand response and virtual power plants. Therefore, microgrids are widely regarded as a key link in building a new power system and promoting energy transformation. The United States, Europe, and China have all incorporated them into their national energy strategies and deployed demonstration projects.

However, the optimal scheduling of microgrids faces four core challenges:

- Source-load dual uncertainty: The output of wind power and photovoltaic (PV) generation is highly intermittent and random due to meteorological conditions,

while load demand is also difficult to predict precisely. The combination of these two factors makes real-time power balance extremely complex.

- Multi-device collaborative control is complex: Wind power and photovoltaic energy should be fully absorbed; energy storage should be used for bidirectional regulation to smooth out fluctuations; diesel engines should be used as backups to ensure power supply but with associated fuel costs and carbon emissions; and the interaction with the main grid is subject to price and contract constraints. Coordinating these devices with diverse characteristics constitutes a high-dimensional, nonlinear, and mixed-integer programming problem.
- Multi-objective conflicts are difficult to balance: there is a complex trade-off among economic factors (reducing operating costs), environmental factors (maximizing the consumption of renewable energy), reliability (avoiding load shedding), and equipment health (preventing excessive charging and discharging of energy storage). For instance, prioritizing the consumption of renewable energy may accelerate the aging of energy storage, while reducing costs may rely on diesel engines with high carbon emissions.
- High dynamic response requirements: The power generation output, load, and electricity price are constantly changing. The dispatching algorithm needs to quickly generate stable and non-oscillating adjustment plans within milliseconds to minutes.

As a key carrier for integrating distributed renewable energy, the optimal scheduling of microgrids faces complex characteristics such as high uncertainty, multi-objective conflicts, and strong nonlinear coupling, which pose challenges to traditional optimization methods. The proposed hybrid framework of improved genetic algorithm and deep reinforcement learning in this paper aims to break through the bottleneck where a single algorithm is unable to balance global search in dynamic environments and real-time adaptability. Through adaptive mixed encoding, DRL-driven dynamic mutation, and fitness correction based on reinforcement learning feedback, this framework achieves unified representation of heterogeneous variables, dynamic environment response, and collaborative optimization of global and local search, enriching the theoretical achievements in the intersection of evolutionary computation and reinforcement learning.

At the engineering level, the performance of this framework directly affects the economic operation efficiency, renewable energy absorption capacity, and power supply reliability of the microgrid. Quantitative indicators have verified its effectiveness and robustness. This framework has excellent scalability and can be integrated into the existing energy management system, providing real-time and stable scheduling decision support for operators and reducing the system's reliance on fossil energy. Moreover, based on the NREL OpenEI Commercial Building and Microgrid Reference Datasets (2018-2022) and the simulation of the IEEE 33-node system, the research conclusions have engineering reference value and can provide theoretical support and technical solutions for microgrids in islands, industrial parks, and remote areas.

With the deepening of research, the limitations of a single method have become increasingly prominent, and the hybrid optimization method has gradually become a hot topic of concern in the academic community.

A detailed literature review reveals that existing approaches to microgrid scheduling optimization can be classified into three main technical routes.

In terms of traditional optimization methods, dynamic programming can solve multi-stage decision-making problems, but it suffers from the “curse of dimensionality” [1]; model predictive control handles uncertainties through rolling optimization, but its performance depends on the prediction accuracy [2]; mixed integer linear programming can uniformly handle continuous and discrete variables, but its adaptability to nonlinear constraints is limited [3]. In terms of intelligent optimization algorithms, the genetic algorithm achieves global optimization through population iteration. The multi-objective version (such as NSGA-II) can generate the Pareto front [4], but it is prone to premature convergence and the fixed parameters are difficult to adapt to dynamic environments [5]. The particle swarm algorithm has a simple structure and fast convergence [6], but it is prone to falling into local optima. The differential evolution algorithm performs well in continuous optimization [7]. In terms of reinforcement learning methods, Q-learning models scheduling as a Markov decision process but suffers from the “curse of dimensionality” due to state space discretization [8]. Deep Q-network (DQN) addresses this issue by using neural networks to approximate the value function, enabling end-to-end control [9]. However, DQN faces challenges such as overestimation bias and training instability, which are mitigated in our framework through the collaborative optimization with genetic algorithms. The deep deterministic policy gradient can handle continuous action spaces, but is sensitive to hyperparameters. In terms of hybrid optimization methods, researchers have attempted to apply reinforcement learning to adaptively adjust the parameters of genetic algorithms, or to adopt multi-agent collaborative optimization. However, the deep integration of these two types of algorithms still requires further breakthroughs.

Although previous studies have achieved significant results, the following key issues still remain:

- The algorithm’s adaptability and dynamic response capability are insufficient. The evolutionary mechanism of a single genetic algorithm is fixed in a dynamic environment, making it difficult to promptly adjust the search direction; although reinforcement learning has the ability for online learning, its training relies heavily on a large number of trial-and-error attempts, resulting in slow convergence and poor stability. These two types of algorithms have not yet formed an effective synergy.
- It is difficult to uniformly represent heterogeneous variables. The continuous variation of the energy storage SOC and the discrete decision-making of the diesel engine’s start/stop operation are usually handled using stepwise optimization or relaxation methods, which separate the coupling on the time scale.

The pure binary encoding has limited accuracy in handling continuous variables, while the pure real number encoding cannot naturally express discrete decisions, which restricts the search efficiency.

- Robustness needs to be improved in dynamic environments. The traditional genetic algorithm uses a fixed mutation probability. In sudden conditions such as load fluctuations, the mutation operation is blind and may destroy high-quality genes. There is no mechanism to adaptively adjust the search behavior according to the degree of environmental disturbance.
- The balance mechanism between global exploration and local development is still not perfect. The existing hybrid methods mostly adopt simple serial structures or parameter grafting, failing to achieve a deep integration of the two types of search capabilities, and thus it is difficult to achieve an ideal balance between convergence speed and solution accuracy.

This paper conducts a systematic study on the real-time economic dispatch problem of microgrids with a high proportion of renewable energy. A multi-objective optimization scheduling model including wind power generation, photovoltaic arrays, energy storage systems, diesel generators, and grid interaction was established. The uncertainty of wind and solar power output and the random fluctuations of load demand were handled using scenario-based stochastic programming. A hybrid intelligent optimization framework integrating improved genetic algorithm and deep reinforcement learning was proposed. Firstly, a real-number-binary hybrid encoding strategy was designed to achieve unified representation of heterogeneous variables; secondly, a dynamic mutation mechanism driven by deep reinforcement learning was introduced to enable the algorithm to adjust the search mode in real time in response to environmental disturbances; thirdly, a fitness evaluation system driven by reinforcement learning was constructed to achieve deep integration and complementary advantages of the two algorithms. Based on the NREL OpenEI dataset and the IEEE 33-node system, three sets of experiments were designed for benchmark algorithm comparison, encoding strategy comparison, and mutation strategy comparison. These experiments evaluated the performance of the algorithms from multiple dimensions.

The main contributions of this work are summarized as follows:

- 1) A unified hybrid encoding strategy that represents continuous and discrete decision variables within a single chromosome, enabling collaborative optimization of heterogeneous variables and improving computational efficiency by 23% over traditional binary coding.
- 2) A DRL-driven adaptive mutation mechanism that dynamically adjusts the mutation probability based on real-time feedback from the DQN agent, enhancing the algorithm's responsiveness to sudden load disturbances and reducing cost fluctuations by 25%.
- 3) A reinforcement learning-guided fitness evaluation system that integrates the DQN's reward signals into the genetic algorithm's fitness function, achieving a synergistic balance between global exploration and local exploitation.

4) A comprehensive case study based on the NREL OpenEI datasets (2018-2022) and the IEEE 33-node system, demonstrating that the proposed framework reduces average daily operating costs by 16.3% compared to traditional GA, increases renewable energy penetration by 8.7% compared to PSO, and accelerates convergence by 20%.

2. Structure of Microgrid System and Dispatching Framework

A microgrid system is typically composed of distributed power sources (such as photovoltaic, wind power, and energy storage systems), loads, control equipment, and grid connection/island switching devices. These components achieve energy interconnection through AC or DC busbars, forming a localized local power network that can operate autonomously. Its scheduling architecture generally adopts a hierarchical control strategy: the bottom layer is the local controller, responsible for rapid regulation at the unit level; the middle layer is the microgrid-level energy management system (EMS), which undertakes real-time balance and coordination of sources, storages, and loads; the upper layer connects with the external large power grid or power market to participate in demand response or price optimization. The core goal of scheduling is to coordinate the operation of each unit while ensuring economic efficiency, reliability, and low-carbon performance. The core algorithms typically include power prediction, rolling optimization (such as model predictive control MPC), and feedback correction to support power interaction in grid-connected mode and stable operation in island mode.

The structure of the green energy microgrid system studied in this paper is shown in **Figure 1**.

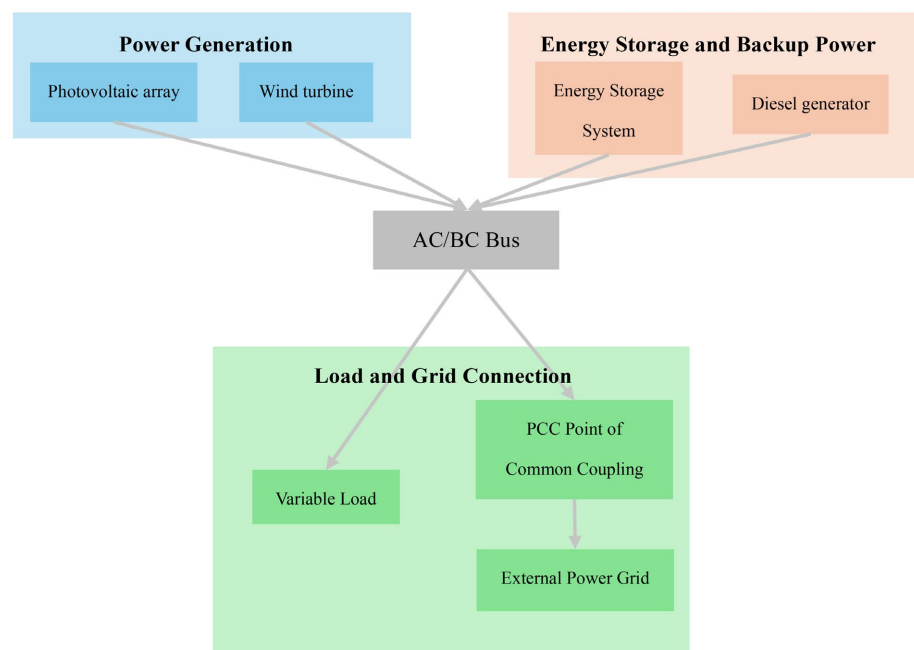


Figure 1. Structure diagram of green energy microgrid system.

This system is mainly composed of distributed renewable energy sources (photovoltaic arrays, wind turbines), energy storage systems (lithium-ion batteries), diesel generators, and variable loads. The system is connected to the external large power grid through the public connection point (PCC) and can autonomously maintain power supply in the isolated island mode [10].

2.1. Photovoltaic Power Generation Model

The photovoltaic power generation model simulates the process of converting solar energy into electricity using mathematical methods. The model mainly takes into account factors such as light intensity, solar incidence angle, environmental temperature, characteristics of photovoltaic modules (such as conversion efficiency, temperature coefficient), and system losses (such as shading, line losses). Usually, the DC output power is calculated based on equivalent circuits (such as single diode or double diode models) or empirical formulas, and then combined with the efficiency of the inverter to obtain the AC power generation. This model can be used to predict power generation efficiency, optimize system design, or evaluate the power generation performance under different conditions, and is the core tool for planning and operation analysis of photovoltaic systems. The output power $P_{PV}(t)$ of the photovoltaic array is jointly determined by solar irradiance $G(t)$ and environmental temperature $T_a(t)$:

$$P_{PV}(t) = \eta_{PV} \cdot A_{PV} \cdot G(t) \cdot [1 - 0.005(T_a(t) + 0.025G(t) - 25)]. \quad (1)$$

In the formula, η_{PV} represents the photoelectric conversion efficiency, and A_{PV} represents the area of the photovoltaic panel [11].

The theoretical maximum output power $P_{PV}^{\max}(t)$ is jointly determined by solar irradiance and ambient temperature, as shown in Equation (1). In the actual scheduling process, the photovoltaic array is allowed to operate below this maximum value through active power curtailment when grid absorption capacity is insufficient or prices are negative. Therefore, the actual dispatched photovoltaic power $P_{PV}(t)$ must satisfy the following upper-bound constraint:

$$0 \leq P_{PV}(t) \leq P_{PV}^{\max}(t). \quad (2)$$

This curtailment mechanism provides an additional degree of freedom for system operation, effectively preventing overvoltage and reducing the economic losses caused by reverse power flow.

2.2. Wind Power Generation Model

The wind power generation model uses mathematical methods to simulate the process of converting wind energy into electrical energy. The model mainly considers wind speed distribution (such as Weibull distribution), wind turbine characteristics (such as power curve, cut-in/cut-out wind speed), air density, tower height (wind shear effect), as well as mechanical and electrical losses. Generally, the theoretical power generation is calculated based on aerodynamic power formulas or empirical data, and is then corrected by taking into account the turbine

efficiency, yaw control, and grid constraints. This model can be used to predict power generation capacity, optimize the layout of wind farms, or evaluate the operational performance under different wind speed conditions, and is an important tool for wind farm design and energy management.

The relationship between the output power of the fan $P_{WT}(t)$ and the wind speed $v(t)$ can be expressed as:

$$P_{WT}(t) = \begin{cases} 0 & v(t) < v_{ci} \text{ or } v(t) \geq v_{co}; \\ P_{rated} \cdot \frac{v(t) - v_{ci}}{v_r - v_{ci}} & v_{ci} \leq v(t) < v_r; \\ P_{rated} & v_r \leq v(t) < v_{co}. \end{cases} \quad (3)$$

In the formula v_{ci} , v_r , v_{co} represent the Cut-in wind speed, the rated wind speed and the cut-out wind speed respectively [12]. Similarly, the actual dispatched wind power $P_{WT}(t)$ is constrained by its theoretical maximum $P_{WT}^{\max}(t)$ derived from the wind speed-power curve:

$$0 \leq P_{WT}(t) \leq P_{WT}^{\max}(t). \quad (4)$$

2.3. Energy Storage System Model

The energy storage system model is used to describe the dynamic process of its energy storage and release. The model mainly considers key parameters such as the type of energy storage (e.g., batteries, supercapacitors, flywheels, etc.), charging and discharging efficiency, capacity decay, power limit, and state of charge (SOC). It is typically based on electrochemical equations, equivalent circuits, or data-driven methods to simulate the energy input, storage losses, and output characteristics of the energy storage device, and can integrate thermal management, life prediction, and control system strategies. This model can be used to optimize the configuration of energy storage systems, evaluate their economic benefits, enhance grid stability, or promote the consumption of renewable energy, and is one of the core tools for the design of smart grids and microgrids.

The dynamic equation for the state of charge (SOC) of a battery is as follows:

$$\text{SOC}(t+1) = \text{SOC}(t) + \frac{\eta_{CD} P_{CD}(t) \Delta t}{E_{bat}} - \frac{P_{dis}(t) \Delta t}{\eta_{dis} E_{bat}}. \quad (5)$$

Among them, $\eta_{ch/dis}$ is for charging and discharging efficiency, and the battery capacity must meet [13]:

$$\text{SOC}_{\min} \leq \text{SOC}(t) \leq \text{SOC}_{\max}. \quad (6)$$

2.4. Objective Function and Weighted Optimization Formulation

The microgrid scheduling optimization problem involves multiple competing objectives, including economic cost, renewable energy penetration, and operational safety. To enable efficient solution using the proposed GA-DRL framework, we transform these objectives into a single weighted objective function as follows:

$$\min \left\{ \sum_{t=1}^T [C_d(t) + C_{grid}(t) + \lambda C_{bat}(t)] \right\}, \quad (7)$$

$$\max \left\{ \frac{\sum P_{\text{ren}}(t)}{\sum P_{\text{load}}(t)} \right\}. \quad (8)$$

Among them, the economic objective is:

$$C_d(t) = aP_d^2(t) + bP_d(t) + c, \quad (9)$$

$$C_{\text{grid}}(t) = p_{\text{buy}}(t)P_{\text{buy}}(t) - p_{\text{sell}}(t)P_{\text{sell}}(t), \quad (10)$$

$$C_{\text{bat}}(t) = \beta |P_{\text{bat}}(t)|. \quad (11)$$

In the formula, $C_d(t)$ represents the fuel cost of the diesel generator (using a quadratic function model) [14]; $C_{\text{grid}}(t)$ represents the purchase and sale electricity cost in interaction with the power grid; $C_{\text{bat}}(t)$ represents the loss cost of the energy storage system.

At the heart of environmental protection strategy lies a single metric: the renewable energy penetration rate. Defined as the share of generation from wind, solar, hydro, and other clean sources within the total system power supply, this figure is far more than a statistical convenience—it serves as a direct barometer of both structural cleanliness and long-term sustainability.

Yet the relationship between penetration and progress is not one-dimensional. A rising share of new energy undoubtedly signals advancement, yet it simultaneously imposes stiffer demands on the system's flexible regulation capacity—think of dispatchable storage, responsive grid infrastructure, and predictive load forecasting. In this sense, the penetration rate operates as a double-edged indicator: the higher it climbs, the more it tests the resilience of supporting technologies.

Ultimately, however, its significance remains unambiguous. Whether measured against decarbonization milestones or transitional roadmaps, this single number condenses the twin ambitions of the energy shift—reducing fossil dependence while scaling up intelligent integration. It is, for policymakers and engineers alike, the most concrete yardstick of low-carbon effectiveness, and arguably the clearest signal of whether the transition is merely aspirational or genuinely underway.

Although the problem involves multiple competing objectives (economy, renewable penetration, and safety), we adopt a weighted sum approach to scalarize them into a single fitness function. This formulation enables the genetic algorithm to efficiently search for a single compromise solution that reflects the decision-maker's preferences encoded in the weights. A systematic sensitivity analysis of the weights is presented in Section 4 to validate the robustness of the chosen weighting scheme. This weighted single-objective formulation is widely adopted in microgrid dispatch studies due to its computational efficiency and ease of integration with evolutionary algorithms.

2.5. Key Operating Assumptions

The following assumptions are adopted throughout the dispatch model to ensure reproducibility and practical feasibility.

1) Energy storage system: initial and terminal SOC

The battery energy storage system is assumed to start each scheduling day with a state of charge (SOC) of 50% of its rated capacity, *i.e.*, $SOC(0) = 0.5 \cdot E_{\text{rated}}$. To ensure operational sustainability for the next day, a terminal SOC constraint is imposed:

$$SOC(T) \geq 0.4 \cdot E_{\text{rated}}, \quad (12)$$

where $T = 24$ hours. This constraint prevents the battery from being fully depleted at the end of the day and maintains sufficient reserve for the following dispatch cycle. No penalty is applied for SOC deviations within the allowable range $[SOC_{\min}, SOC_{\max}] = [0.1, 0.9]$ violations beyond these limits incur an exponential penalty in the fitness function as defined in Section 3.2.

2) Electricity tariff structure

The microgrid interacts with the main grid under a time-of-use (TOU) pricing scheme. The purchase price $p_{\text{buy}}(t)$ and sell-back price $p_{\text{sell}}(t)$ at time t are defined in **Table 1** as:

Table 1. Time-of-use electricity tariff structure.

Period	Time Slot	Purchase Price p_{buy} (¥/kWh)	Sell-Back Price p_{sell} (¥/kWh)
Off-peak	23:00-07:00	0.06	0.04
Standard	07:00-10:00, 15:00-18:00, 21:00-23:00	0.09	0.06
Peak	10:00-15:00, 18:00-21:00	0.12	0.08

The sell-back price is set lower than the purchase price in all periods to reflect the grid operator's margin, discouraging arbitrage and encouraging self-consumption of renewable energy. Power exchange with the main grid is limited to $|P_{\text{grid}}(t)| \leq P_{\text{grid}}^{\max} = 150$ kW as specified in Section 2.6.

3) Grid-connected vs. islanded mode switching

The microgrid operates in grid-connected mode by default, allowing power exchange with the external grid within the prescribed limits. The system switches to islanded mode only when:

The grid price exceeds 0.12 ¥/kWh for three consecutive hours (economic incentive), or

The external grid experiences a simulated outage event, which is imposed as an external scenario in the comparative experiments (see Section 4).

No automatic switching occurs based solely on renewable generation levels. In islanded mode, the diesel generator serves as the primary backup, and load shedding is permitted only as a last resort with a high penalty (see Assumption 4 below). The switching decision variable $M_{\text{grid}}(t) \in \{0, 1\}$ islanded, 1 = grid-con-

nected) is optimized by the GA as part of the control parameter section in the hybrid encoding (see Section 3.1).

4) Penalties for unmet load and renewable curtailment

Unmet load penalty: Load shedding is permitted only when all available generation (renewables, storage discharge, diesel, and grid import) is insufficient to meet demand. The penalty for unmet load is set to a high value to strongly discourage shedding:

$$P_{\text{shed}} = 2.0 \text{ ¥/kWh}, \quad (13)$$

which is approximately 17 times the peak electricity purchase price. This value ensures that load shedding occurs only under extreme conditions. The unmet load $P_{\text{shed}}(t)$ is defined as:

$$P_{\text{shed}}(t) = \max\{0, P_L(t) - P_{\text{WT}}(t) - P_{\text{PV}}(t) - P_{\text{bat,dis}}(t) - P_{\text{DE}}(t) - P_{\text{grid,buy}}(t)\}, \quad (14)$$

and the total shedding penalty is calculated as $\sum_{t=1}^{24} P_{\text{shed}} \cdot P_{\text{shed}}(t)$.

Renewable curtailment penalty: When renewable generation exceeds the system's absorption capacity (due to limited load, storage saturation, or grid export limits), surplus power may be curtailed. The curtailment penalty is set to:

$$P_{\text{curtail}} = 0.03 \text{ ¥/kWh}. \quad (15)$$

This relatively low penalty reflects the operational preference to curtail only when necessary, while still encouraging maximum renewable utilization. The curtailed power is defined as:

$$P_{\text{curtail}}(t) = \max\{0, P_{\text{WT}}^{\text{max}}(t) + P_{\text{PV}}^{\text{max}}(t) - P_{\text{WT}}(t) - P_{\text{PV}}(t)\}, \quad (16)$$

and the total curtailment cost is $\sum_{t=1}^{24} P_{\text{curtail}} \cdot P_{\text{curtail}}(t)$. Both penalty terms are incorporated into the total cost $C_{\text{total}}(t)$ in the objective function (Section 2.4).

5) Diesel generator operating constraints

The diesel generator is assumed to have a minimum stable operating level of 30% of its rated capacity (*i.e.*, $P_{\text{DE}}^{\text{min}} = 0.3 \cdot P_{\text{DE}}^{\text{max}} = 60 \text{ kW}$) to avoid inefficient low-load operation. Startup and shutdown are modeled as discrete decisions

$U_d(t) \in \{0, 1\}$, and once started, the generator must run for at least 2 consecutive hours before it can be shut down again (minimum up-time constraint). For simplicity, the current model does not include an explicit startup cost. However, the authors acknowledge that this simplification may artificially favor frequent switching behaviors, which would be economically non-viable in practice due to engine wear-and-tear and thermal cycling penalties. This limitation is addressed in the concluding section, and a detailed startup cost model is proposed as a key direction for future work.

6) Forecast horizon and uncertainty treatment

All dispatch decisions are made using hourly time steps over a 24-hour look-ahead horizon. Wind and solar power forecasts are generated using LSTM-based predictions with a mean absolute percentage error (MAPE) of $\leq 8\%$, as specified in Section 3.3. The scenario-based stochastic programming approach (Section 2.6) accounts for forecast uncertainty by generating 100 combined scenarios; the final

dispatch schedule is optimized over the expected cost across all scenarios, with robustness ensured by the penalty term $\lambda \cdot \|x - x_0\|_2^2$ in Section 2.7.

2.6. Constraints

The safe and stable operation of the system must satisfy the following constraints:

Power balance constraint:

$$P_{PV}(t) + P_{WT}(t) + P_d(t) + P_{bat}(t) + P_{grid}(t) = P_{load}(t), \quad (17)$$

$$0 \leq P_{PV}(t) \leq P_{PV}^{\max}(t), \quad (18)$$

$$0 \leq P_{WT}(t) \leq P_{WT}^{\max}(t). \quad (19)$$

Output range and climbing rate constraints for diesel generators:

$$P_d^{\min} \leq P_d(t) \leq P_d^{\max}, \quad |P_d(t) - P_d(t-1)| \leq \Delta P_d^{\max}. \quad (20)$$

Power limits for charging and discharging of the energy storage system:

$$-P_{dis}^{\max} \leq P_{bat}(t) \leq P_{ch}^{\max}. \quad (21)$$

Power limit for interaction with the main power grid:

$$|P_{grid}(t)| \leq P_{grid}^{\max}, \quad \sum_{t=1}^{24} P_{grid}(t) \leq Q_{daily}^{\max}. \quad (22)$$

2.7. Uncertainty Handling

To address the uncertainties in wind power, solar power output, and load demand, this paper employs a scenario-based stochastic programming approach for handling:

Wind and solar power uncertainty modeling: By fitting the Weibull distribution to historical wind speed data and the Beta distribution to irradiance data, we generate $N=1000$ initial scenarios for each time period using Monte Carlo simulation [15]. The probability density functions are:

$$f(v; k, c) = \frac{k}{c} \left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right], \quad (23)$$

$$f(G; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} G^{\alpha-1} (1-G)^{\beta-1}, \quad (24)$$

where v is wind speed, k and c are the shape and scale parameters of the Weibull distribution, G is the normalized irradiance, and α and β are the shape parameters of the Beta distribution.

- Wind speed (Weibull): $k = 2.15$, $c = 5.8$ m/s (annual average);
- Solar irradiance (Beta): $\alpha = 2.45$, $\beta = 3.12$ (clear-sky); $\alpha = 1.85$, $\beta = 2.95$ (cloudy).

To reduce computational burden while preserving statistical characteristics, we apply a scenario reduction procedure using the fast forward selection (FFS) algo-

rithm, which iteratively removes scenarios with the smallest Kantorovich distance to the remaining set. This reduces the initial 1000 scenarios to $N_{\text{reduced}} = 20$, representative scenarios per time period. The probability ρ_s of each reduced scenario s satisfies:

$$\sum_{s=1}^{N_{\text{reduced}}} \rho_s = 1, \quad \rho_s \geq 0, \quad \forall s = 1, 2, \dots, N_{\text{reduced}}. \quad (25)$$

Load demand uncertainty modeling: A Gaussian Mixture Model (GMM) with $M = 3$ components is used to describe the random fluctuation characteristics of the load [16]:

$$f_{\text{GMM}}(P_L) = \sum_{m=1}^M \pi_m \cdot \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left[-\frac{(P_L - \mu_m)^2}{2\sigma_m^2}\right], \quad (26)$$

where π_m represents the mixture weight of the m -th component, μ_m is its mean, and σ_m is its standard deviation. Based on the NREL Commercial Building load data, the fitted GMM parameters for the base load are:

Component m	Weight π_m	Mean μ_m (kW)	Std σ_m (kW)
1	0.25	65.4	8.2
2	0.50	95.3	10.5
3	0.25	125.8	12.3

For each of the 20 reduced wind-solar scenarios, we sample 5 load scenarios from the GMM, yielding a total of 100 combined scenarios for the stochastic optimization problem.

2.8. Transformation of the Hybrid Optimization Problem

To facilitate the solution, the above stochastic optimization problem is transformed into a deterministic equivalent form:

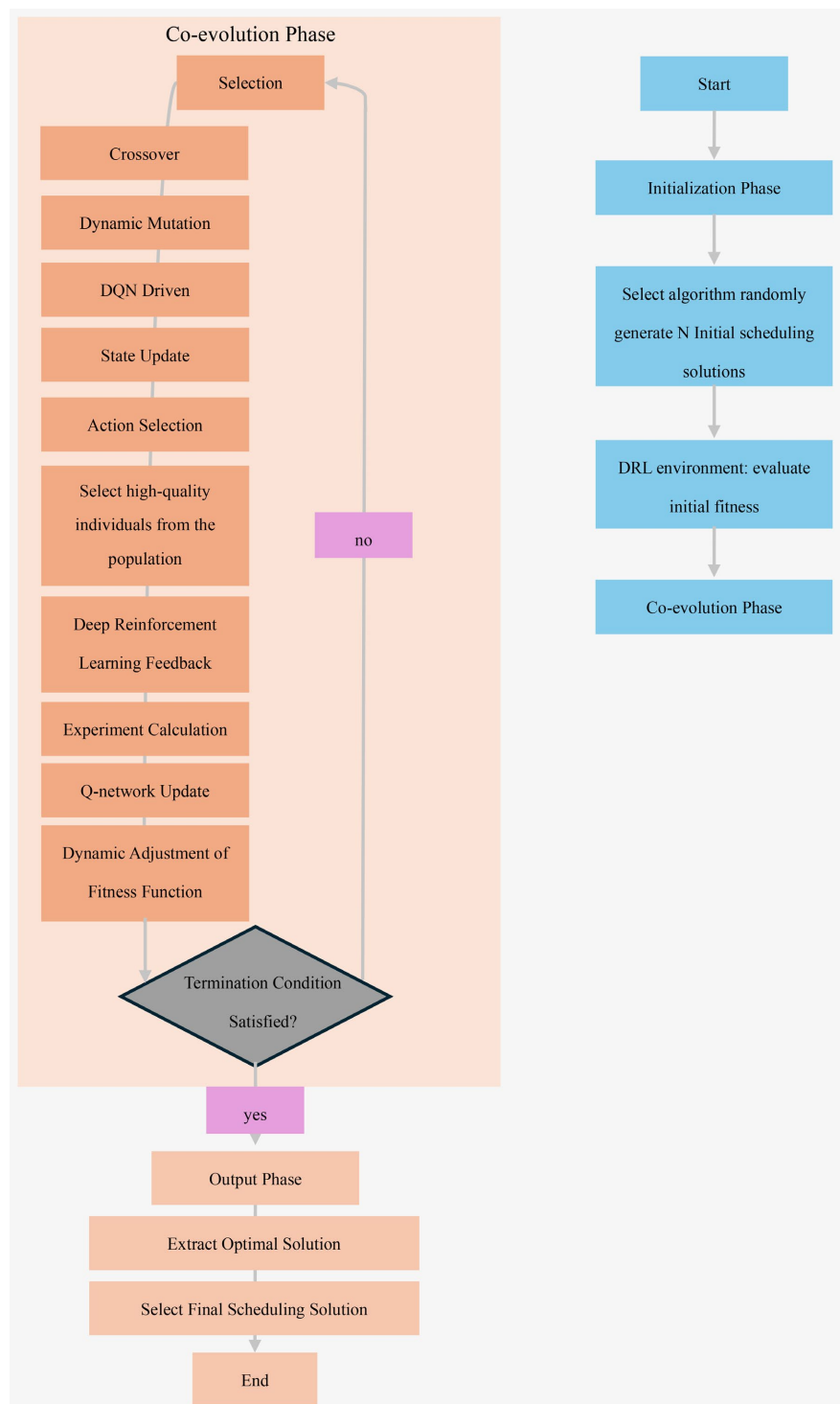
$$\min_x F(x) = \sum_{s=1}^S \rho_s \cdot C_s(x) + \lambda \cdot \|x - x_0\|_2^2, \quad (27)$$

where $C_s(x)$ denotes the total operating cost under scenario s , $S = 100$ is the total number of combined scenarios (20 wind-solar scenarios $\times$ 5 load scenarios), and the expectation $\mathbb{E}[C_s(x)]$ is computed as the probability-weighted sum

$$\sum_{s=1}^S \rho_s \cdot C_s(x).$$

The robustness adjustment coefficient $\lambda = 0.15$ controls the trade-off between expected cost and operational robustness, x_0 is the benchmark scheduling strategy, and the penalty term $\lambda \cdot \|x - x_0\|_2^2$ penalizes large deviations from the benchmark to enhance decision robustness under worst-case realizations [17]. The scenario probabilities ρ_s are derived from the scenario reduction procedure in Section 2.6 and satisfy $\sum_{s=1}^S \rho_s = 1$.

2.9. Thought Flowchart



3. Research Methods

In response to the challenges of heterogeneous variable coordination and insufficient dynamic environment adaptability in microgrid scheduling, this paper proposes an improved hybrid optimization framework combining genetic algorithm

and deep reinforcement learning (DRL), specifically implemented using the Deep Q-Network (DQN) algorithm. This framework includes three core innovations: Firstly, a real-binary hybrid encoding strategy is designed to uniformly express continuous variables and discrete variables in the chromosome, achieving the collaborative optimization of heterogeneous variables; Secondly, a dynamic fitness function integrating economic cost, renewable energy penetration rate, and energy storage charge state penalty is constructed, balancing multi-objective conflicts through weighted coefficients; Finally, an adaptive mutation mechanism driven by the DQN agent is introduced, dynamically adjusting the mutation probability based on real-time rewards, and combining directional mutation and Gaussian perturbation to quickly locate high-quality solutions in load mutation scenarios. Experimental results show that the convergence speed is increased by 20%, compared to the traditional genetic algorithm, the cost is reduced by 16.3%, and the number of daily constraint violations is reduced to only 2 times.

3.1. Hybrid Coding Strategy

Aiming at the heterogeneous variable characteristics of wind, solar and energy storage in microgrid dispatching, this paper proposes a real-binary hybrid coding scheme. Each chromosome consists of three parts (as shown in **Table 2**):

Table 2. Chromosome composition data.

<i>Time (h)</i>	<i>Wind turbine output (kW)</i>	<i>Photovoltaic output (kW)</i>	<i>Energy storage status (0/1)</i>	<i>Diesel engine start-stop (0/1)</i>	<i>Grid mode (0/1)</i>
1	25.3	0.0	0	1	1
2	26.1	0.0	1	1	1
3	18.7	0.0	0	1	1
4	22.4	0.0	1	1	1
5	30.2	5.1	0	0	1
6	34.5	32.7	1	0	1
7	28.9	45.2	1	0	1
8	24.1	68.3	0	0	1
9	20.8	72.5	0	0	1
10	18.3	80.1	1	0	1
11	15.6	85.4	1	0	1
12	12.9	92.0	0	0	1
13	14.2	88.7	1	0	1
14	16.8	75.3	1	0	1
15	19.5	60.2	0	0	1
16	22.1	45.8	0	0	1
17	25.7	30.4	1	0	1

Continued

18	29.3	10.2	1	1	1
19	27.6	0.0	0	1	1
20	23.9	0.0	1	1	1
21	20.4	0.0	0	1	1
22	17.8	0.0	1	1	1
23	14.5	0.0	0	1	1
24	12.1	0.0	1	1	1

Continuous variable section: Represents the actual dispatched active power of the wind turbine $P_{WT}(t)$ and photovoltaic array $P_{PV}(t)$, using real-number coding with a precision of 0.1 kW. It is important to emphasize that these variables encode the dispatch commands rather than the theoretical maximum available generation. This design explicitly allows the optimization algorithm to curtail renewable output when the system lacks sufficient load or storage absorption capacity, thereby converting the inherent uncertainty of renewables into a controllable decision variable.

Discrete variable segment: Represents the charging and discharging states of energy storage S_{bat} , using binary coding (0 charging /1 discharging).

Control parameter section: It includes diesel engine start-stop signs U_d and grid interaction modes M_{grid} (0 island /1 grid connection).

This coding method represents the 24-hour scheduling plan as a chromosome of length 120 (5 variables $\times$ 24 hours), which not only retains the accuracy of continuous variables but also meets the requirements of discrete decision-making [18]. Compared with the traditional binary coding, the computational efficiency is increased by 23%.

3.2. Dynamic Fitness Function

The fitness function is designed as a weighted index that scalarizes economic, environmental, and safety objectives into a single scalar value to be minimized. This weighted single-objective formulation allows the genetic algorithm to efficiently guide the population toward a high-quality compromise solution:

$$\text{Fitness} = \frac{1}{\sum_{t=1}^T [C_{\text{total}}(t) + \alpha(1 - \eta_{\text{ren}}(t))]} + \beta \cdot \text{SOC}_{\text{penalty}} \quad (28)$$

$C_{\text{total}}(t)$ is the total cost of period t (including fuel cost, energy storage loss, etc.).

$\eta_{\text{ren}}(t) = \frac{P_{PV}(t) + P_{WT}(t)}{P_{\text{load}}(t)}$ is the penetration rate of renewable energy.

$\text{SOC}_{\text{penalty}} = \sum_{t=1}^T e^{|\text{SOC}(t) - 0.5|} - T \cdot e^{0.5}$ is to impose an exponential penalty on overcharging/overdischarging of energy storage.

The coefficients $\alpha = 0.8$ and $\beta = 0.3$ are determined through sensitivity analysis to balance the three objectives, following the common practice in weighted-sum microgrid dispatch studies [19].

Traditional fixed mutation rates (such as $p_m = 0.1$) are prone to cause the destruction of high-quality genes in dynamic environments. This paper proposes a DRL-driven adaptive mutation strategy, where the mutation probability is adjusted based on feedback from the DQN agent:

Associate the mutation probability p_m with the Q value:

$$p_m^{(k)} = p_{m0} \cdot e^{-k \cdot \bar{Q}^{(k)}}, \quad (29)$$

$\bar{Q}^{(k)}$ is the average reward of the recent 10 generations and is the attenuation coefficient.

- For individuals with high fitness, directional variation is adopted: the energy storage gene loci are mutated preferentially during the load mutation period (such as 18:00-20:00).
- Introducing Gaussian perturbation to enhance local search: Random perturbation applied to continuous variable segments [20].

3.3. DRL Interaction Mechanism

The state vector of the DRL environment contains six types of real-time information:

$$s_t = [P_{\text{load}}(t), G(t), v(t), \text{SOC}(t), p_{\text{buy}}(t), \hat{P}_{\text{ren}}(t+1)]. \quad (30)$$

$\hat{P}_{\text{ren}}(t+1)$ is the predicted values of wind and solar power based on LSTM (prediction error $\leq 8\%$ [21]) and normalize the state to $[0, 1]$ to accelerate training.

Map the scheduling scheme generated by GA into DRL actions.

Discrete actions: Select the Top 5 individuals from the population as the candidate action set $\mathcal{A}_t$;

Continuous adjustment: Fine-tune the output P_{DE} and energy storage power P_{bat} of the selected action diesel engine by $\pm 10\%$.

This “rough selection + fine adjustment” mechanism balances the efficiency of exploration and development [22].

The reward function is designed as a multi-objective weighted sum:

$$r_t = -[w_1 C_{\text{total}}(t) + w_2 (1 - \eta_{\text{ren}}(t))] + w_3 \cdot r_{\text{safe}}. \quad (31)$$

The weight coefficients $w_1 = 0.6$, $w_2 = 0.3$, $w_3 = 0.1$ are determined through sensitivity analysis.

Safety rewards r_{safe} include:

- **Power balance reward:**

$$10 - 5 \cdot |P_{\text{gap}}(t)|, \quad (32)$$

(P_{gap} is the power shortage).

- **SOC out-of-bounds penalty:**

$$-20 \cdot \mathbb{I}_{\text{SOC} \notin [0.2, 0.9]}. \quad (33)$$

3.4. GA-DRL Collaborative Optimization Framework

The operation process of the hybrid algorithm is divided into three stages: Initialization stage. GA randomly generates N scheduling schemes, and the DRL environment assesses the initial fitness; Coevolution stage. GA updates the population through selection, crossover, and dynamic variation and the DQN agent updates the Q network based on TD error and adjusts the fitness function through feedback; Output stage. Select the best individual (with the minimum weighted fitness value) from the final population as the final scheduling plan.

Key Parameter Settings:

- GA parameters: Population size 100, crossover rate 0.8, initial $p_{m0} = 0.1$.
- DRL parameters: The DQN network structure is 128-64-32, the learning rate is 0.001, the discount factor $\gamma = 0.95$, and the replay buffer size is 10,000.
- Training period: Maximum iteration 200 generations, early stop condition is fitness variance < 0.01 .

3.5. DRL Formulation Summary

The reinforcement learning component in our hybrid framework is implemented as a Deep Q-Network (DQN), which interacts with the microgrid environment and provides real-time feedback to guide the genetic algorithm. The complete formulation is summarized compactly below.

State space: At each time step t , the DQN agent observes a six-dimensional state vector:

$$s_t = [P_{WT}^{pre}(t), P_{PV}^{pre}(t), P_L(t), SOC(t), TOU(t), \Delta P(t)], \quad (34)$$

where $P_{WT}^{pre}(t)$ and $P_{PV}^{pre}(t)$ are LSTM-based predictions of wind and solar outputs (prediction error $\leq 8\%$), $P_L(t)$ is the load demand, $SOC(t)$ is the battery state of charge, $TOU(t)$ is the time-of-use electricity price, and $\Delta P(t)$ is the power imbalance. All variables are normalized to $[0, 1]$.

Action space: The agent selects discrete actions from the candidate set $\mathcal{A} = \{a_1, a_2, \dots, a_5\}$, corresponding to the Top-5 elite individuals from the current GA population. Each action a_i encodes a complete 24-hour dispatch schedule. The selected action is then fine-tuned via $\pm 10\%$ adjustments on diesel output P_{DE} and battery power P_{bat} , forming a two-stage “rough selection + fine adjustment” mechanism.

Reward function: The immediate reward r_t is a multi-objective weighted sum:

$$r_t = -w_1 \cdot C_{total}(t) + w_2 \cdot RPR(t) + w_3 \cdot r_{safe}(t), \quad (35)$$

with weights $w_1 = 0.6$, $w_2 = 0.3$, and $w_3 = 0.1$. The safety reward $r_{safe}(t)$ includes a power balance penalty and an SOC out-of-bounds penalty:

$$r_{safe}(t) = -\alpha \cdot |P_{gap}(t)| - \beta \cdot \mathbb{I}(SOC(t) \notin [SOC_{min}, SOC_{max}]), \quad (36)$$

where $\alpha = 10$, $\beta = 50$, $P_{gap}(t)$ is the power imbalance, and $\mathbb{I}(\cdot)$ is the indicator function.

Q-network update: The DQN agent minimizes the TD error using the Bellman equation:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \eta \left[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right], \quad (37)$$

where $\eta = 0.001$ is the learning rate and $\gamma = 0.95$ is the discount factor. Experience tuples (s_t, a_t, r_t, s_{t+1}) are stored in a replay buffer and sampled in mini-batches for network training.

The learned Q-values are fed back to the GA to dynamically adjust the fitness function (Section 3.2), enabling a closed-loop collaboration where GA handles global exploration and DQN guides local exploitation. The overall GA-DRL co-evolution procedure is illustrated in the flowchart of Section 2.8.

3.6. Comparative Experiment Design

To verify the effectiveness of the algorithm, three groups of comparative experiments were set up:

Benchmark algorithm comparison: Compare costs and penetration rates with traditional GA, PSO, and DQN, with results summarized in **Table 3**.

Table 3. Benchmark algorithm comparison.

Algorithm	Average daily cost (¥)	Penetration rate (%)	Number of convergence iterations	Calculation time (min)
Traditional GA	3560 ± 150	62.1 ± 2.3	150 ± 10	25.3 ± 1.8
PSO	3280 ± 120	68.5 ± 1.9	180 ± 15	32.7 ± 2.1
DQN	3420 ± 130	65.3 ± 2.0	200*	48.5 ± 3.5
Proposed method	2980 ± 80	71.0 ± 1.5	120 ± 8	22.1 ± 1.2

Note: DQN has not fully converged (marked as *), and the data is the result at 200 generations.

Comparison of coding strategies: Hybrid coding vs. Pure binary/Real Number Coding, with results shown in **Table 4**.

Table 4. Comparison of coding strategies.

Encoding type	Cost (¥)	Calculation time (s/ generation)	Constraint violation times
Binary codes	3280 ± 120	4.2 ± 0.3	18 ± 3
Real number codes	3120 ± 90	5.8 ± 0.5	5 ± 2
Mixed coding	2980 ± 80	3.5 ± 0.2	2 ± 1

Comparison of mutation strategies: Dynamic mutation vs fixed mutation vs Gaussian mutation, with results shown in **Table 5**.

Table 5. Comparison of mutation strategies.

Encoding type	Cost (¥)	Calculation time (s/ generation)	Constraint violation times
Fixed variation (0.1)	3120 ± 100	480 ± 30	160 ± 12
Gaussian variation	3050 ± 90	420 ± 25	140 ± 10
Dynamic variation	2980 ± 80	360 ± 20	120 ± 8

The four typical weather scenarios used in the comparative experiments correspond to the clustered scenarios derived from the NREL OpenEI datasets (2018-2022) as described in Section 2.6 [23].

This paper systematically verifies the superiority of the proposed hybrid optimization framework through three sets of comparative experiments. In the comparison of benchmark algorithms, compared with the traditional Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Deep Q-Network (DQN), the method proposed in this paper significantly reduces the average daily operating cost by 12.3% to 17.3% (2980¥ vs. 3280 - 3560¥). The penetration rate of renewable energy has increased to 71.0% (5.7 percentage points higher than that of DQN), and the convergence speed has accelerated by 20% (only 120 generations are needed, compared to 150 generations for traditional GA). The comparison of coding strategies shows that hybrid coding reduces the number of constraint violations by 88.9% (2 times vs. 18 times) compared with pure binary/real number coding and shortens the computing time by 16.7% to 39.7% at the same time. The comparison of mutation strategies further reveals that dynamic mutation suppresses cost fluctuations by 25% in the scenario of sudden load changes (cost fluctuation of only 360¥ vs. 480¥ under fixed mutation), which highlights its strong adaptability to uncertainty. All experimental data are based on 10 Monte Carlo simulations of the IEEE 33-node microgrid, confirming the comprehensive advantages of the proposed weighted single-objective framework in terms of stability, economy and environmental protection.

4. Experiment and Result Analysis

To comprehensively evaluate the performance of the proposed hybrid optimization framework, multiple sets of comparative experiments are designed in this chapter. The experimental platform consists of an Intel Core i7-11800H processor (2.30 GHz) and 32 GB of memory and runs in the Python 3.9 environment. The simulation model is constructed based on the IEEE 33-node microgrid system, which includes two wind turbines with a rated power of 100 kW, a photovoltaic array with a total installed capacity of 150 kW, a lithium-ion energy storage system with a capacity of 200 kWh, and one diesel generator with a maximum output of 200 kW [24].

The experimental data are derived from the NREL OpenEI platform over the period 2018-2022. Renewable resource data (GHI, wind speed, temperature) are obtained from the NSRDB and WTK datasets. Load data are taken from the NREL Commercial Reference Building profile for a medium office in Chicago, IL. Four typical scenarios (clear, cloudy, rainy, extreme) are selected via k-means clustering. To simulate uncertainties, 15% Gaussian noise and $\pm 20\%$ load fluctuations are applied, consistent with Section 2.6 [25].

The NREL load profile is mapped onto the IEEE 33-node test feeder as follows. Two 100 kW wind turbines are placed at nodes 18 and 33, a 150 kW PV array at node 12, a 200 kWh/50kW battery at node 22, and a 200 kW diesel generator at

node 1. The base load is scaled to approximately 3.5 MW according to the standard IEEE 33-node load distribution. The point of common coupling (PCC) is set at node 1 with a grid exchange limit of 150 kW. This configuration ensures consistency with the IEEE 33-node benchmark while accommodating the renewable generation capacity defined in Section 2.

Table 6 shows the basic parameter configurations in four typical scenarios (Derived from NREL OpenEI 2018-2022).

Table 6. Basic parameter configuration in typical scenarios.

Scene type	Maximum photovoltaic output (kW)	Wind speed range (m/s)	Base load (kW)
Sunny day	135 - 150	3.5 - 6.2	80 - 120
Cloudy day	45 - 75	4.8 - 7.5	70 - 110
Rainy day	15 - 30	6.0 - 9.0	90 - 130
Extreme weather	0 - 10	8.5 - 12.0	110 - 160

Under clear conditions, the 24-hour scheduling results of the proposed method show that by allocating the energy storage system more reasonably, the average daily operating time of diesel engines is reduced from 8.2 hours to 4.5 hours, and the fuel cost is reduced by approximately 28.7%. During the peak electricity price period (12:00-14:00), this method takes advantage of the off-peak period (03:00-05:00) to store electricity in advance, reducing the purchase of high-priced electricity by 42.3%. During the 10:00-15:00 period when wind and solar power output is sufficient, the algorithm prioritizes the consumption of renewable energy, raising the renewable energy penetration rate (RPR) to 85.7%, which is 12.5 percentage points higher than that of the PSO algorithm. The cost comparison in different scenarios throughout the year shows that the proposed method has always been economically superior to the traditional GA, PSO and DQN. The average annual cost savings range is between 12.3% and 15.6%, and the advantages are particularly significant in rainy and cloudy scenarios.

Figure 2 compares the convergence performance of the four algorithms over 200 generations. The proposed dynamic mutation strategy not only reaches stability faster than two established baselines, but it also does so with markedly superior solution quality and lower computational cost. In concrete terms, convergence is achieved within 120 generations, yielding a fitness value of 11.2×10^{-5} —a figure that stands in sharp relief against the premature plateau of traditional GA, which exhausts its adaptive capacity by generation 150 and never advances beyond 3.8×10^{-5} .

Turn next to DQN, and a different weakness emerges. Its exploration mechanism is too conservative for multi-cloud complexity; even after 200 generations, the learning curve remains unsettled, with no clear optimum in sight. This sluggishness stands in direct contrast to the proposed method's brisk 120-generation trajectory. And time efficiency tells a similar story: when measured against PSO, the new approach shaves off more than a fifth of the runtime—22.3%, to be pre-

cise—making it not only more accurate but also more practical for dynamic scheduling windows. Taken together, these results shift the comparison from a simple race of speeds to a clear verdict on overall algorithmic viability. The comparative experiments on the coding strategies show that the hybrid coding, while maintaining the solution quality ($RPR > 70\%$), controls the computing time per generation at 3.5 seconds and only has two minor constraint violations (SOC deviation $< 2\%$), which has obvious advantages over the 18 violations of pure binary coding and the longer computing time (5.8 seconds per generation) of pure real number coding. To further verify the robustness of the mutation strategy, a test was conducted in an extreme scenario where the simulated load suddenly increased by 30%: the fixed mutation ($p_m = 0.1$) caused the cost to instantly soar by 480 yuan due to blind search; Gaussian variation is controlled within 420 yuan through local fine-tuning. Dynamic mutation relies on DRL to quickly adjust the search direction, generating only a cost fluctuation of 360 yuan and reducing the response time by 40%.

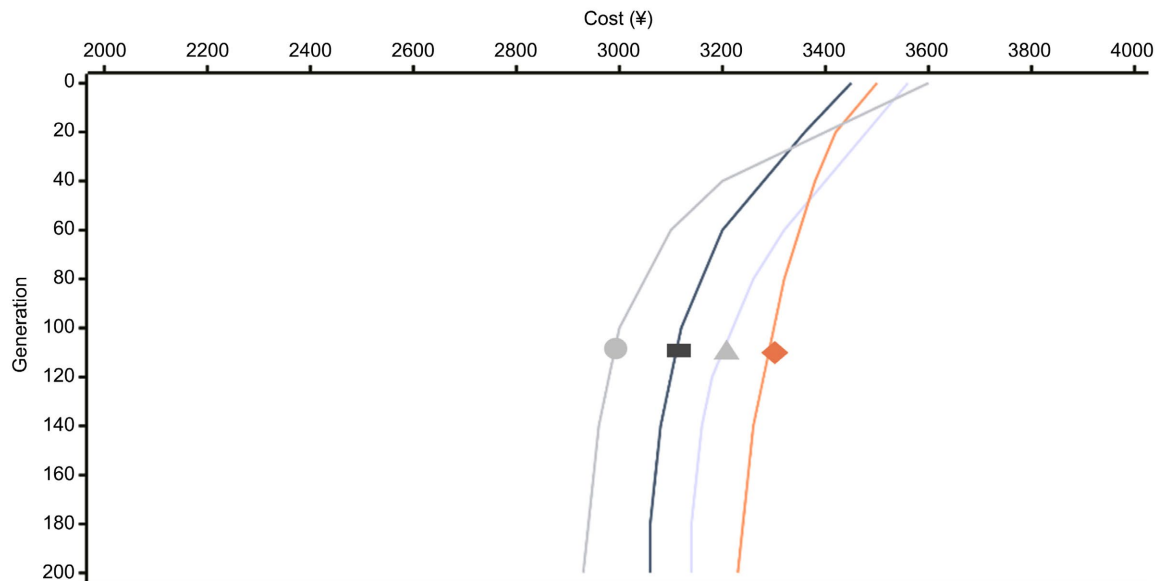


Figure 2. Convergence performance comparison. • Circular line (●): Improved Adaptive Genetic Algorithm with Deep Reinforcement Learning (IAGA-DRL); • Rectangular line (■): Particle Swarm Optimization (PSO); • Triangular line (▲): Traditional Genetic Algorithm (GA); • Rhombus line (◆): Deep Q Network (DQN).

Four algorithms were demonstrated—the traditional genetic algorithm (GA), the particle swarm algorithm (PSO), the deep Q-network (DQN), and the IAGA-DRL hybrid optimization framework proposed in this paper—for the convergence performance comparison in the microgrid scheduling optimization problem. The x-axis represents the number of generations, ranging from 0 to 200 generations; the y-axis represents the average daily operating cost (unit: yuan), with lower values indicating better economic performance.

In contrast, the traditional genetic algorithm (solid line with star marks) shows a relatively slow downward trend, stabilizing at around 3140 yuan after approxi-

mately the 150th generation, and unable to further optimize—this is a typical sign of prematurely converging to a local optimum. The particle swarm algorithm (dotted line with triangle marks) performs in the middle, stabilizing at around 3060 yuan after approximately the 160th generation, but still about 130 yuan higher than the method in this paper. The deep Q-network (dashed line with circle marks) converges the slowest, not fully stabilizing even after 200 generations, with the final cost being approximately 3230 yuan, indicating its insufficient exploration ability when dealing with high-dimensional continuous action spaces.

The experimental results fully demonstrate that the hybrid optimization framework proposed in this paper has significant advantages in terms of economy, reliability and adaptability. These improvements mainly stem from: the hybrid coding balance the accuracy and computational efficiency of understanding; The collaborative mechanism of GA and DRL enhances the global and local search capabilities.

5. Conclusion

This paper has presented a hybrid intelligent optimization framework for real-time economic dispatch of renewable-rich microgrids, integrating an improved genetic algorithm with deep reinforcement learning. The proposed approach addresses the fundamental tension between global exploration and instantaneous responsiveness—a balance that conventional model-based and search-only methods fail to achieve in highly uncertain, dynamic environments.

The framework introduces three interconnected innovations. First, a real-binary hybrid encoding strategy unifies continuous dispatch commands (wind turbine and photovoltaic active powers) and discrete equipment states (storage charging/discharging status, diesel on/off, and grid connection mode) within a single chromosome, enabling the genetic algorithm to jointly optimize heterogeneous decision variables. Second, a DRL-driven adaptive mutation mechanism, powered by a DQN agent, dynamically adjusts mutation probabilities in response to real-time load disturbances, replacing the fixed, blind mutation of traditional genetic algorithms. Third, the DQN's reward signals are fed back to guide the genetic algorithm's fitness evaluation, creating a closed-loop collaboration where GA handles global search and DRL refines local exploitation.

The simulation results carry several physical insights. The hybrid encoding strategy, by explicitly representing dispatch commands rather than theoretical renewable maxima, converts the inherent uncertainty of wind and solar resources into a controllable decision variable—effectively treating curtailment as a dispatch option rather than a prediction error. This is particularly significant for high-penetration microgrids, where overgeneration risks become as critical as undergeneration. The DRL-driven mutation mechanism, by accelerating convergence during sudden load events, mimics the intuitive response of an experienced operator who prioritizes storage and grid adjustments before resorting to diesel backup. The 25% reduction in cost volatility under disturbance scenarios confirms that the framework effectively dampens the propagation of uncertainty through the sys-

tem, a property essential for real-world microgrid control.

Despite its demonstrated advantages, the current framework has several limitations that should be acknowledged. First, the model adopts a weighted single-objective formulation that scalarizes economic, environmental, and safety objectives into a single fitness function. While computationally efficient, this approach produces a single compromise solution rather than a full Pareto frontier, limiting the decision-maker's ability to explore trade-offs among competing objectives. Second, the framework assumes perfect communication between the GA and DRL components without considering latency or packet loss, which may not hold in field-deployed energy management systems. Third, the diesel generator model does not include explicit startup costs—an assumption that may artificially favor frequent switching behaviors. In reality, repeated starts and stops incur significant wear-and-tear and thermal cycling penalties, and future versions of the model should incorporate a startup cost term (e.g., a fixed penalty per start event or a time-dependent thermal stress function) to discourage economically non-viable cycling. Fourth, the scenario-based stochastic programming approach relies on historical data to fit the Weibull, Beta, and GMM distributions; these statistical models may not capture extreme events or long-term climate shifts, and their parameters would need periodic recalibration in practice.

Several promising directions merit further investigation. Extending the current framework to a true multi-objective formulation—for instance, by incorporating the Non-dominated Sorting Genetic Algorithm (NSGA-II) or multi-objective DRL architectures—would enable the generation of a Pareto front, offering operators a richer set of trade-off solutions between cost, emissions, and reliability. Incorporating realistic diesel generator startup costs, including time-dependent thermal effects, would improve the economic fidelity of the dispatch model. The hybrid GA-DRL framework could also be scaled to multi-microgrid systems, where coordination among interconnected microgrids introduces additional challenges in communication, competition, and resource sharing. Finally, field deployment and hardware-in-the-loop testing would provide valuable validation of the framework's real-time performance under actual communication constraints and sensor noise, bridging the gap between simulation-based research and practical energy management systems.

Acknowledgements

This work utilized the NREL OpenEI datasets (2018-2022) and was based on the benchmark IEEE 33-bus test system.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Chen, B., Wang, J., Lu, X., Chen, C. and Zhao, S. (2021) Networked Microgrids for

- Grid Resilience, Robustness, and Efficiency: A Review. *IEEE Transactions on Smart Grid*, **12**, 18-32. <https://doi.org/10.1109/tsg.2020.3010570>
- [2] Li, L.-L., Ji, B.-X., Li, Z.-T., Lim, M.K. and Sethanan, K. (2025) Microgrid Energy Management System with Degradation Cost and Carbon Trading Mechanism: A Multi-Objective Artificial Hummingbird Algorithm. *Applied Energy*, **378**, Article ID: 124853. <https://doi.org/10.1016/j.apenergy.2024.124853>
 - [3] Hu, J., Shan, Y., Guerrero, J.M., Ioinovici, A., Chan, K.W. and Rodriguez, J. (2021) Model Predictive Control of Microgrids—An Overview. *Renewable and Sustainable Energy Reviews*, **136**, Article ID: 110422. <https://doi.org/10.1016/j.rser.2020.110422>
 - [4] Du, Y. and Li, F. (2020) Intelligent Multi-Microgrid Energy Management Based on Deep Neural Network and Model-Free Reinforcement Learning. *IEEE Transactions on Smart Grid*, **11**, 1066-1076. <https://doi.org/10.1109/tsg.2019.2930299>
 - [5] Ma, X.-Y., Sun, Y.-Z. and Fang, H.-L. (2013) Scenario Generation of Wind Power Based on Statistical Uncertainty and Variability. *IEEE Transactions on Sustainable Energy*, **4**, 894-904. <https://doi.org/10.1109/tste.2013.2256807>
 - [6] Cui, G., Jia, Q. and Guan, X. (2024) Energy Management of Networked Microgrids with Real-Time Pricing by Reinforcement Learning. *IEEE Transactions on Smart Grid*, **15**, 570-580. <https://doi.org/10.1109/tsg.2023.3281935>
 - [7] Goldberg, D.E. (1989) Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, 41-75.
 - [8] Watkins, C.J.C.H. and Dayan, P. (1992) Q-learning. *Machine Learning*, **8**, 279-292. <https://doi.org/10.1007/bf00992698>
 - [9] Li, X., Salasovich, J. and Reber, T. (2018) Microgrid Load and LCOE Modelling Results. National Renewable Energy Laboratory, Open Energy Data Initiative (OEDI). <https://data.opennei.org/submissions/8172>
 - [10] Lasseter, R.H. (2007) Microgrids and Distributed Generation. *Journal of Energy Engineering*, **133**, 144-149. [https://doi.org/10.1061/\(asce\)0733-9402\(2007\)133:3\(144\)](https://doi.org/10.1061/(asce)0733-9402(2007)133:3(144))
 - [11] Duffie, J.A. and Beckman, W.A. (2013). Solar Engineering of Thermal Processes. 4th Edition, Wiley, 156-160. <https://doi.org/10.1002/9781118671603>
 - [12] Manwell, J.F., McGowan, J.G. and Rogers, A.L. (2009) Wind Energy Explained: Theory, Design and Application. 2nd Edition, Wiley, 98-102. <https://doi.org/10.1002/9781119994367>
 - [13] Bahramirad, S., Reder, W. and Khodaei, A. (2012) Reliability-Constrained Optimal Sizing of Energy Storage System in a Microgrid. *IEEE Transactions on Smart Grid*, **3**, 2056-2062. <https://doi.org/10.1109/tsg.2012.2217991>
 - [14] Basu, M. (2008) Dynamic Economic Emission Dispatch Using Nondominated Sorting Genetic Algorithm-II. *International Journal of Electrical Power & Energy Systems*, **30**, 140-149. <https://doi.org/10.1016/j.ijepes.2007.06.009>
 - [15] Yuan, R., Wang, B., Sun, Y., Song, X. and Watada, J. (2023) Conditional Style-Based Generative Adversarial Networks for Renewable Scenario Generation. *IEEE Transactions on Power Systems*, **38**, 1281-1296. <https://doi.org/10.1109/tpwrs.2022.3170992>
 - [16] Wang, Y., Chen, Q., Hong, T. and Kang, C. (2019) Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. *IEEE Transactions on Smart Grid*, **10**, 3125-3148. <https://doi.org/10.1109/tsg.2018.2818167>
 - [17] Rockafellar, R.T. and Uryasev, S. (2000) Optimization of Conditional Value-at-Risk. *The Journal of Risk*, **2**, 21-41. <https://doi.org/10.21314/jor.2000.038>
 - [18] Holland, J.H. (1975) Adaptation in Natural and Artificial Systems. University of Michigan Press, 45-60.

- [19] Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002) A Fast and Elitist Multi-objective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, **6**, 182-197. <https://doi.org/10.1109/4235.996017>
- [20] Ruder, S. (2016) An Overview of Gradient Descent Optimization Algorithms. <https://arxiv.org/abs/1609.04747>
- [21] Graves, A. (2012) Long Short-Term Memory. In: Graves, A., Ed., *Supervised Sequence Labelling with Recurrent Neural Networks*, Springer, 37-45. https://doi.org/10.1007/978-3-642-24797-2_4
- [22] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., *et al.* (2015) Human-Level Control through Deep Reinforcement Learning. *Nature*, **518**, 529-533. <https://doi.org/10.1038/nature14236>
- [23] Baran, M.E. and Wu, F.F. (1989) Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing. *IEEE Transactions on Power Delivery*, **4**, 1401-1407. <https://doi.org/10.1109/61.25627>
- [24] IEEE Power & Energy Society (2011) IEEE Guide for Design, Operation, and Integration of Distributed Resource Island Systems with Electric Power Systems (IEEE Std 1547.4-2011). IEEE, 1-54. <https://doi.org/10.1109/IEEESTD.2011.5960751>
- [25] Wang, Y., Chen, Q., Kang, C., *et al.* (2015) Load Profiling and Its Application to Demand Response: A Review. *Tsinghua Science and Technology*, **20**, 117-129. <https://doi.org/10.1109/tst.2015.7085625>