
Scalable Power Dispatch Automation Methods Based on Artificial Intelligence in Distributed Energy Systems

Zhaoyuan Yin^{1,*}, Xiongbao Zhang¹, Wanhe Luo²
and Shidi Ruan¹

¹*Automation Section, Guangxi Power Grid Company Limited, Nanning, 530000, Guangxi, China*

²*Marketing and Distribution Command Center, Shanglin Power Supply Bureau, Guangxi Power Grid Company Limited, Nanning, Nanning, 530500, Guangxi, China*

E-mail: zhaoyuanyin12@outlook.com; xiongbaozhang0611@outlook.com; wanheluo0728@outlook.com; shidi_ruan@outlook.com

**Corresponding Author*

Received 19 January 2026; Accepted 03 April 2026

Abstract

With the high penetration of distributed energy resources (DERs), the strong uncertainty of their output and the complex interaction of network topology pose severe challenges to the real-time performance, scalability, and intelligence level of dispatching methods. This paper proposes a hierarchical distributed collaborative intelligent dispatching architecture based on deep reinforcement learning (DRL) and graph neural networks (GNNs). This architecture adopts a two-layer design of “regional agent-edge controller.” The upper-layer regional agent uses GNN to process the global power flow state and generate node marginal price signals; the lower-layer edge controllers act as independent DRL agents, making autonomous decisions based on local observations and price signals. Through a hybrid mechanism

Distributed Generation & Alternative Energy Journal, Vol. 41_3, 791–814.

doi: 10.13052/dgaej2156-3306.41310

© 2026 River Publishers

of centralized training and distributed execution, collaborative learning and plug-and-play expansion of the system are achieved. This method significantly reduces dispatch costs when dealing with uncertainty, achieving a high photovoltaic grid integration rate of up to 100%, and controlling the maximum net power deviation between day-ahead planning and real-time operation to within 0.35 MW. Even when the system is expanded to 500 DER units, the online decision-making time remains at 7.5 milliseconds, and communication overhead only slowly increases to 35.0 KB, demonstrating its significant effectiveness in improving system resilience, operational economy, and scalability.

Keywords: DER, power dispatch automation, deep reinforcement learning, scalability, GNN.

1 Introduction

The large-scale integration of distributed energy sources (DERs), represented by wind and solar power, has accelerated the evolution of the power grid from a traditional centralized, unidirectional radial structure into a complex decentralized system composed of massive prosumers with bidirectional power flow. While this transformation improves the cleanliness of energy, it also introduces significant intermittency and uncertainty, posing unprecedented challenges to the safe, stable, and economical operation of the power grid [1, 2]. This paper proposes and constructs a new scalable power dispatch automation method that integrates artificial intelligence (AI) and distributed computing concepts. With the rapid growth of distributed energy resources such as photovoltaic systems, wind turbines, and energy storage devices, modern power systems are becoming increasingly decentralized and complex. Traditional centralized dispatch methods face significant challenges in handling large-scale distributed units, communication overhead, and real-time decision-making requirements. In addition, the uncertainty and intermittency of renewable energy sources further increase the difficulty of maintaining system stability and economic operation. Therefore, developing intelligent, scalable, and distributed power dispatch strategies has become an important research direction in smart grid and distributed energy system management. The contributions of this paper include: (1) designing a hierarchical distributed physical information architecture of “regional agent-edge controller”, which coordinates global security and local autonomous decision-making through price signals, laying the foundation for scalability;

(2) at the upper layer, using GNN to embed the power system topology and operating status to approximate the solution of economic scheduling and generate guiding node marginal electricity prices; (3) at the lower layer, configuring a DRL agent based on competing dual deep Q network (Dueling DQN) as an edge controller for each DER unit, enabling it to make optimal scheduling decisions autonomously and in real time based on local information and electricity price signals; (4) proposing a hybrid training mechanism to obtain high-performance strategies through offline centralized training and achieve millisecond-level response through online distributed execution, and conducting theoretical and experimental analysis on its communication and computation scalability.

2 Related Work

With the widespread integration of high-proportion renewable energy and distributed energy (DER) into the power system, the power network is evolving from a centralized, one-way power supply mode to a decentralized, two-way interactive complex ecosystem. Optimal reactive power dispatch (ORPD) is a key issue to ensure the safe and economical operation of the power system. Mouassa et al. [3] proposed a new design scheme for an artificial ecosystem optimizer specifically for solving large-scale ORPD problems. Ibrahim et al. [4] proposed an alternating optimization method to solve the voltage-safe multi-period ORPD problem. Yue et al. [5] developed a collaborative control strategy for virtual power plants based on deep reinforcement learning, highlighting the effectiveness of AI-driven methods in coordinating distributed generation and energy storage resources. Agouzoul et al. [6] showed that the reactive power dispatch model considering load uncertainty can obtain a more robust and practical dispatch strategy compared with the deterministic optimization model. The complexity of energy management and power dispatch in microgrids increases. Odonkor et al. [7] designed an intelligent controller based on an adaptive neurofuzzy inference system (ANFIS) for distributed generator energy control and power dispatch in grid-connected microgrids. Ali et al. [8] studied the multi-objective, multi-period, stable environmental economic power dispatch problem and used probabilistic methods to model the uncertainty of wind power and photovoltaic power generation. Rana et al. [9] investigated a multi-objective economic and environmental dispatch model considering systems with and without renewable energy sources, demonstrating the influence of renewable penetration on cost and emission performance.

After distributed generation (DG) is connected to the distribution network, the uncertainty of its output poses a challenge to the reactive power balance and voltage regulation of the system. Ali et al. [10] proved through numerical simulation that the dandelion optimizer algorithm can effectively handle reactive power optimization problems with DG uncertainty. Casilimas-Peña et al. [11] used high-precision short-term photovoltaic power prediction information to optimize the scheduling of all controllable reactive power resources in the microgrid in advance. Muttaqi and Sutanto [12] designed adaptive and predictive energy management strategies for VPPs that integrate renewable energy and energy storage to achieve real-time optimal power scheduling. Based on the hybrid optimization problem of reactive power scheduling and renewable energy allocation, Abaza et al. [13] found a high-quality integrated solution that optimizes the reactive power flow of the system while reasonably determining the optimal access point and capacity of renewable energy. Kumari et al. [14] applied machine learning techniques to reduce wind power deviation charges, improving scheduling accuracy under renewable uncertainty. Abd-El Wahab et al. [15] combined artificial rabbit optimization algorithm and gradient-based optimization to solve the reactive power scheduling problem of the power network. Chen and Hu [16] proposed a distributed optimization model for smart grid economic dispatch to enhance computational efficiency and coordination among multiple distributed units. Existing scheduling strategies cannot continuously learn during operation and have limited intelligence levels. There is an urgent need for a new paradigm of power dispatch automation with endogenous scalability. To this end, this paper constructs an intelligent decision-making architecture that integrates deep learning, reinforcement learning and parallel computing technologies to achieve full-chain adaptive automation from system perception, modeling, optimization to control, providing core methodological support for building a highly resilient and efficient next-generation smart grid. Recent studies in distributed energy systems have focused on intelligent optimization and data-driven control strategies. Balasubramanyam et al. [17] proposed a hybrid swarm intelligence and cuckoo search-based energy management system for microgrids to achieve optimal energy scheduling and improved operational efficiency. Zhu et al. [18] investigated the application of demand-side technologies for intelligent power system regulation, demonstrating improved load management and system stability. Fan et al. [19] developed a data-driven monitoring and status evaluation framework for distribution networks, enabling more reliable operation and efficient integration of distributed energy resources.

3 Methods

3.1 Layered Distributed Cooperative Scheduling Architecture

This paper constructs a two-layer physical information architecture of “regional agent-edge controller” (Figure 1).

At the physical layer, the system is divided into a transmission and distribution electronic network managed by a central regional agent, and a large number of distributed DER (Distributed Energy Resources) aggregation units. Each DER aggregation unit (such as a photovoltaic power station, energy storage cluster, or adjustable load aggregator) is equipped with an edge controller as its local decision-making unit. At the information layer, the upper-layer regional agent perceives the global topology and wide-area operating status through an improved AC power flow model and periodically generates node marginal price signals. These LMP signals are broadcast to all lower-layer edge controllers through a communication network. Each edge controller, as an independent intelligent agent, has an observation space containing only real-time operating data of its local DER, such as output, state of charge, and adjustable range, and the received LMP signals [20, 21]. Based on this information, the controller autonomously calculates and issues scheduling commands such as increasing or decreasing output and charging/discharging power using its built-in deep reinforcement learning policy network. Decision information only needs to be executed locally and uploaded to the regional agent for status updates, eliminating the need for

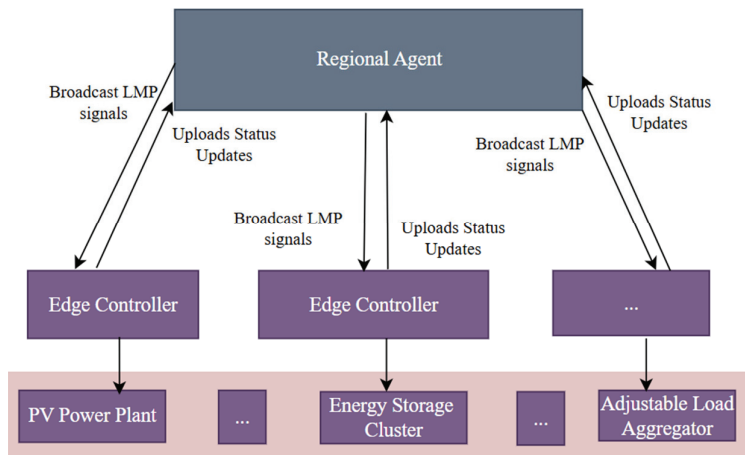


Figure 1 Cooperative scheduling architecture.

frequent global information exchange between a large number of distributed units, thus laying the foundation for scalability.

3.2 Upper-Level Regional Agent Design

The upper-level regional agent generates price signals that coordinate global security constraints. It maintains an improved flexible AC power flow model, which uses system bus injected power and voltage amplitude as variables. The agent first obtains the overall network power flow status through measurement units, and then processes this data using a two-layer graph neural network (GNN). The input layer feature vector of this GNN includes the net injected power, voltage amplitude phase angle history sequence, and line impedance parameters for each node. The system topology is defined by the adjacency matrix, serving as the basis for graph convolution operations. After two layers of graph convolution and ReLU activation function processing, the GNN output layer learns the embedded representation of system power flow and constraints [22, 23].

Table 1 shows the changes in the feature embedding representation of some nodes in the GNN input layer and after two layers of graph convolution and activation function processing, reflecting the model's ability to extract high-level features related to global power flow and constraints from the original operating data.

Based on this representation, the regional agent solves an approximate economic scheduling problem through a fully connected network module, and calculates the node marginal price reflecting line congestion and loss. This process can be formalized as the following optimization problem:

$$\begin{aligned}
 & \min_P \sum_{i=1}^N C_i(P_i) \\
 & s.t. \quad \sum_{i=1}^N P_i = D \\
 & \quad \quad F(P) \leq F_{\max}
 \end{aligned} \tag{1}$$

P_i is the injected power of node i , $C_i(\cdot)$ is its cost function, D is the total load of the system, $F(P)$ is the line power flow vector, and F_{\max} is the upper limit of line capacity. The agent solves the problem approximately through a fully connected network, and obtains the marginal cost of each node, i.e., the node marginal price (LMP).

Table 1 Feature vector processing data of each layer of GNN

Node ID	Input Layer Features (Net Injection Power (p.u.), Voltage Magnitude (p.u.), Mean Phase Angle History)	First Graph Conv. Output (Dim: 8)	Second Graph Conv. Output (Dim: 8)	Final Embedding Representation (Dim: 4)
1	(0.12, 1.02, 0.05)	(0.15, 0.01, 0.98, 0.22, 0.00, 0.47, 0.33, 0.11)	(0.89, 0.05, 0.74, 0.31, 0.02, 0.61, 0.40, 0.18)	(0.82, 0.12, 0.68, 0.25)
2	(−0.08, 0.98, −0.02)	(0.03, 0.45, 0.10, 0.67, 0.20, 0.00, 0.88, 0.05)	(0.21, 0.78, 0.15, 0.52, 0.45, 0.03, 0.91, 0.12)	(0.18, 0.71, 0.10, 0.48)
3	(0.25, 1.05, 0.08)	(0.31, 0.08, 0.85, 0.14, 0.01, 0.60, 0.27, 0.09)	(0.95, 0.14, 0.62, 0.25, 0.04, 0.73, 0.35, 0.22)	(0.90, 0.10, 0.55, 0.30)
4	(0.00, 1.00, 0.00)	(0.10, 0.22, 0.30, 0.41, 0.15, 0.25, 0.50, 0.08)	(0.40, 0.51, 0.42, 0.33, 0.38, 0.44, 0.65, 0.20)	(0.35, 0.48, 0.38, 0.28)
5	(−0.15, 0.96, −0.05)	(0.00, 0.67, 0.05, 0.82, 0.30, 0.02, 0.75, 0.04)	(0.15, 0.92, 0.11, 0.61, 0.52, 0.08, 0.84, 0.15)	(0.12, 0.85, 0.08, 0.58)

This LMP vector is updated and broadcast every 5 minutes. Its dimension is the same as the number of system nodes. It aims to guide the behavior of downstream distributed energy resources (DER) and indirectly alleviate the problems of line overload and voltage over-limit [24].

3.3 Design of Lower-Level Edge Controller Agents

Each edge controller is a deep reinforcement learning agent. Its observation space is defined as the real-time operating status of the local distributed energy resource, prediction information, and the marginal electricity value received by the current node. The action space consists of the scheduling instructions for the distributed energy resource; for photovoltaic inverters, this might be the reactive power setpoint, and for energy storage systems, it might be the charging and discharging power instructions. Both are standardized into continuous or discrete action spaces. The reward function is designed to include three parts: the economic benefit of energy trading based on the node marginal electricity price signal, a penalty for local operating constraints, and an additional reward from the upper-level regional agent for providing ancillary services [25, 26].

In the reward function, the economic benefit can be expressed as a function related to the node marginal electricity price and the injected power. Specifically, if the node marginal electricity price is λt , and the net injected power of the agent at time step t is P_i , then the trading benefit can be calculated as:

$$R_{econ,t} = \lambda t \cdot P_i \quad (2)$$

Meanwhile, for energy storage systems, their state of charge (SOC) must be kept within an allowable range, thus a constraint penalty term is introduced. Let the SOC of the energy storage system be z_t , with allowable lower and upper limits z_{min} and z_{max} , respectively. Then the penalty term is:

$$F = \max(0, Z_{min} - z_t) - \max(0, z_t - Z_{max}) \quad (3)$$

The agent employs a competitive dual-deep Q-network as the policy network. This network first processes the observed input through a three-layer fully connected encoder. Then, the feature stream is divided into two branches: a value stream outputs a scalar representing the value of the current state; the other advantage stream outputs a vector representing the advantage of each action relative to the average. The final Q-value is obtained by adding the value and advantage. This structure helps to more accurately assess the relative importance of different actions. For state s and action a , the Q value is calculated as follows:

$$Q(s, a; \theta) = V(s; \theta) + \left(A(s, a; \theta) - \frac{1}{|A|} \sum_{a'} A(s, a'; \theta) \right) \quad (4)$$

$V(s; \theta)$ is the value stream output, $A(s, a; \theta)$ is the advantage stream output, θ is the network parameter, A is the action space, and a is any action within the action space used for summation.

How the DRL agent evaluates internal value and advantage based on the observed state through its policy network and finally outputs a scheduling decision, thus concretizing the abstract description above [27], as shown in Table 2:

At time t_1 , despite a low state of charge (SOC = 0.45), the high marginal node price (LMP = 35.2) and positive net power command lead the agent to evaluate that the action “discharge +0.5 p.u.” has the highest advantage value (0.80), thus achieving economic benefits. Conversely, at time t_5 , the combination of high SOC (0.90) and low LMP (15.8) makes the advantage value (0.85) of the action “charge -1.0 p.u.” significantly higher than other

Table 2 Schematic diagram of the decision-making process of the DRL edge controller agent

Decision Timestep	Local Observation State (ESS SOC, Net Power Ref., LMP)	Value Stream Output (State Value)	Advantage Stream Output A(s,a) (Actions: -1.0 p.u., 0 p.u., +0.5 p.u., +1.0 p.u.)	Final Selected Action (Charge/Discharge Power)
t ₁	(0.45, 0.10, 35.2)	1.2	(0.15, 0.30, 0.80, 0.50)	+0.5 p.u.
t ₂	(0.85, -0.20, 28.5)	0.8	(0.90, 0.60, 0.10, -0.20)	-1.0 p.u.
t ₃	(0.60, 0.00, 40.1)	2.1	(-0.30, 0.05, 0.45, 0.70)	+1.0 p.u.
t ₄	(0.50, 0.15, 25.0)	1.5	(0.20, 0.40, 0.95, 0.60)	+0.5 p.u.
t ₅	(0.90, -0.30, 15.8)	0.3	(0.85, 0.50, -0.10, -0.40)	-1.0 p.u.

actions, leading the agent to choose charging to store energy and avoid discharging during periods of low electricity price [28, 29].

The network uses the Adam optimizer with a learning rate of 0.001 and an experience replay buffer size of 50000.

3.4 Hybrid Training Mechanism

The training mechanism consists of an offline centralized training phase and an online distributed execution phase. In the offline training phase, a centralized training distributed execution framework is used. In a central trainer, the policy network and value network are instantiated for each edge controller agent. The training simulation environment is based on a complete system model, and the central trainer can access the global observation information of all agents to calculate more accurate value estimates and global rewards [30]. Specifically, during training, the agent’s optimization objective can be expressed as maximizing the expected cumulative discount reward:

$$J(\pi) = \mathbb{E} \left[\sum_{t=0}^T \gamma^t r(s_t, a_t) \right] \quad (5)$$

π represents the policy, τ represents the trajectory, s_t and a_t represent the state and action at time t , respectively, r_t represents the reward function, γ represents the discount factor, and T represents the time range. During training, each agent generates actions based on its current policy, interacts with the environment to produce transition samples, and stores them in a shared experience pool. The trainer samples from the pool, calculates the target Q-value using global information, and updates the network parameters of each agent. This process uses a “parameter sharing” strategy, whereby all DER agents of the same type share the same set of initial network parameter values to accelerate learning.

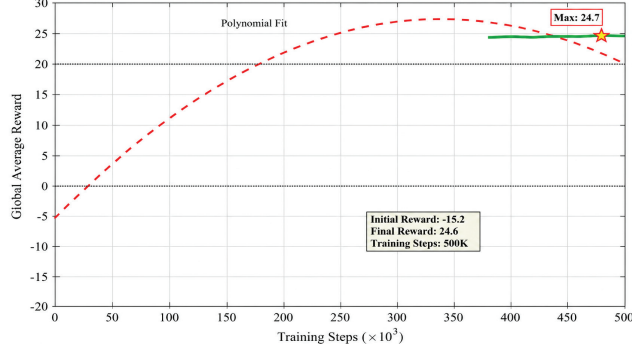


Figure 2 Correspondence between training steps and global average reward.

Figure 2 illustrates the change in the global average reward with the number of training steps during the training process, reflecting the gradual optimization of the agent's policy through experience replay and parameter sharing.

This convergence trend verifies that the adopted hybrid training mechanism can effectively guide a large number of distributed agents to learn collaboratively, ultimately obtaining a high-performance scheduling strategy.

In online applications, a central trainer is no longer needed; each edge controller only loads the finally trained policy network model. At each decision moment (every 15 seconds), the controller independently performs forward computation based on purely local observations and received LMP signals. Its action generation process can be described by the following formula:

$$a_t = \mu_\theta(o_t | \lambda_t) \quad (6)$$

Here, μ_θ represents the policy network with parameter θ , o_t is the local observation, and λ_t is the electricity price signal for the current time period. This process generates scheduling instructions within milliseconds, achieving fully distributed autonomous operation.

3.5 Scalability and Communication Optimization Analysis

The scalability of this architecture stems from its communication and computation modes. Regarding communication requirements, the upper-layer regional agent only needs to periodically broadcast the LMP vector to the entire network once. The amount of data is proportional to the number of system nodes and independent of the total number of DER units. The lower-level

edge controllers only need to upload a small amount of key state information for model updates of the regional agents, without needing to communicate with other edge controllers. Therefore, the total system communication load does not increase linearly with the number of DERs, avoiding broadcast storms. In terms of computational complexity, the computational burden of online decision-making is completely offloaded to each edge controller. The forward computation of the DRL policy network for each controller is a constant complexity operation, unaffected by system size. Although the LMP computation of the regional agents involves GNNs and optimization, its input dimension is fixed at the number of nodes, not the number of DERs. When using graph convolutional layers to process node features, the computation can be formalized as:

$$(l + 1)H = \sigma \left(\hat{D}^2 \hat{H} \frac{1}{W} \right) \quad (7)$$

\tilde{A} is the adjacency matrix with self-loops, \tilde{D} is the degree matrix, $H^{(l)}$ is the feature of the l -th layer node, $W^{(l)}$ is the trainable weight, and σ is the activation function. When the system expands and adds new DER units, only new edge controllers need to be deployed and pre-trained network models of the same type of policy need to be loaded. The expansion cost of this process can be modeled as:

$$C_{extend} = k \cdot M \quad (8)$$

M is the number of new DERs, and k is the constant deployment cost of a single controller. This design does not require redesigning or training the central coordinator, achieving plug-and-play horizontal expansion capability.

4 Results and Discussion

4.1 Uncertainty Coping Capability

The experiment was based on the IEEE 33-bus standard distribution test system and implemented on the MATLAB/Simulink and Python co-simulation platform. To verify the superior performance of the proposed method under high proportion of renewable energy uncertainty, all tests were based on continuous 24-hour, 15-minute interval day-ahead scheduling and real-time adjustment, simulating typical daily irradiance fluctuations and load changes.

To quantify the economics, the expected reserve costs (USD/MW) required to cope with photovoltaic output prediction deviations by different

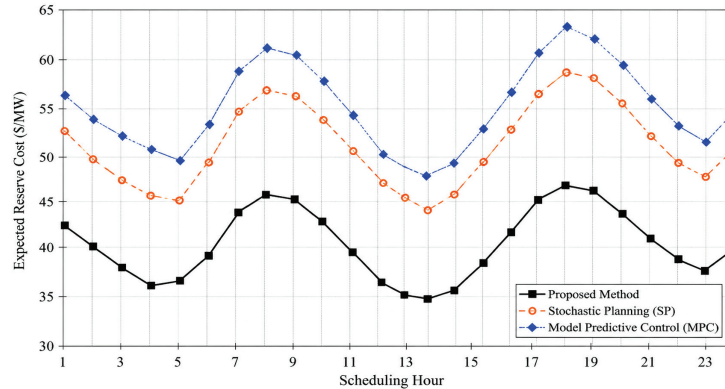


Figure 3 Expected reserve costs.

methods were compared. The hourly cost data for this method, stochastic programming (SP), and model predictive control (MPC) are shown in Figure 3.

The AI-based scheduling method proposed in this paper exhibits significantly lower expected reserve costs than stochastic programming (SP) and model predictive control (MPC) across all 24 scheduling periods. Particularly during the peak solar output period at midday (12:00–14:00), this method, through flexible adjustment by an intelligent edge controller, fully utilizes renewable energy, minimizing costs to a minimum (US\$34.8/MW). A comparison of the average daily cost shows that this method reduces costs by approximately 22.2% and 27.1% compared to SP and MPC, respectively, highlighting its significant advantage in mitigating uncertainty and economic costs.

Meanwhile, Figure 4 compares the solar grid integration rates of the three methods at different times, visually revealing how the proposed method maximizes renewable energy output through intelligent coordination:

Figure 4, through area plots and curve comparisons, visually demonstrates the significant improvement in solar grid integration achieved by optimized scheduling. During the daytime period with sufficient sunlight (6:00–16:00), the proposed method effectively mitigates the fluctuations and intermittent ties of photovoltaic output by coordinating energy storage charging and discharging with flexible load regulation, maintaining the absorption rate at a high level of over 90%, and achieving 100% full absorption at 9:00. In contrast, the baseline scenario shows a significant dip in the absorption rate during midday due to network congestion and voltage exceeding limits,

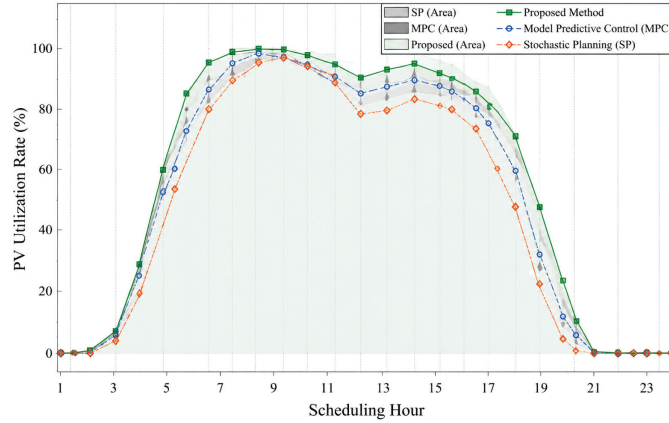


Figure 4 Solar Grid Integration Rate.

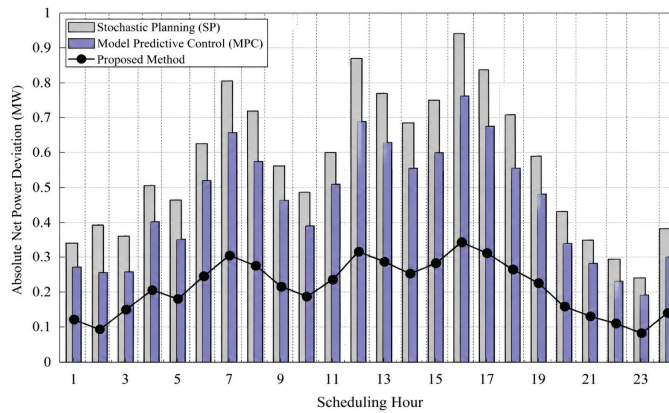


Figure 5 Net Power Deviation.

verifying the effectiveness of the proposed method in promoting the efficient utilization of renewable energy.

To further evaluate the robustness of the dispatch plan in actual operation, Figure 5 statistically analyzes the absolute value of the net power deviation between the day-ahead dispatch plan and the actual required adjustment under different methods. This indicator reflects the dispatch strategy’s ability to cope with real-time fluctuations; the smaller the deviation, the stronger the robustness of the plan and the lower the pressure on real-time power dispatch balance.

Figure 5 shows that the proposed method not only exhibits a smaller deviation but also a significantly smaller fluctuation range compared to the other two methods, with a maximum value of only 0.35 MW. This indicates that the DRL-based edge controller in this paper can more accurately predict local behavior and respond to global electricity price signals in real time, thereby significantly reducing the deviation between actual operation and day-ahead planning. Smaller planning deviation means fewer real-time balancing resource calls and higher system operational reliability, fully demonstrating the superior robustness of the proposed intelligent scheduling architecture in the face of uncertainty.

4.2 Robustness Under Extreme Uncertainty Scenarios

Evaluate the robustness of the proposed dispatch architecture, required additional simulations were conducted under extreme uncertainty scenarios. Two stress conditions were considered: (1) severe photovoltaic output fluctuation and (2) sudden load variation. In the first scenario, photovoltaic generation was artificially varied within $\pm 40\%$ of its predicted value to emulate rapid irradiance changes caused by cloud coverage. In the second scenario, the system load was abruptly increased by 30% within a short time interval to simulate unexpected demand surges. Simulation results show that the proposed DRL-based edge controllers can quickly adjust the charging and discharging behavior of energy storage units and regulate distributed generation outputs according to the received node marginal price signals. Even under these extreme disturbances, the system maintained stable operation without line overload or voltage violations, and the average dispatch deviation remained within an acceptable range. These results demonstrate the strong robustness and adaptability of the proposed intelligent dispatch framework when facing severe renewable generation fluctuations and sudden load changes.

4.3 Scalability Performance

The simulated system was expanded from 50 units to 500 units by gradually increasing the number of DER aggregation units in the distribution network. For each scale configuration, the proposed distributed AI scheduling method and the benchmark centralized optimization method were run, and 100 online scheduling decisions were executed consecutively. The average computation time of each decision (from triggering computation to all units outputting instructions) was recorded. The results are compared in Table 3.

Table 3 Online decision time under different DER scales

Number of DER Aggregated Units	Proposed Method Average Decision Time (ms)	Centralized Optimization Average Decision Time (ms)	Centralized/Proposed Method Time Ratio
50	4.2	105.3	25.1
100	4.5	423.7	94.2
200	4.8	1890.5	393.9
300	5.1	4312.8	845.6
500	7.5	>10000 (Solution Timeout)	500

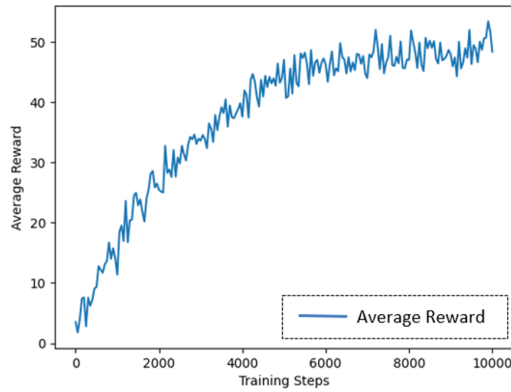


Figure 6 Training convergence curve of the DRL agent.

As the number of DERs increased from 50 to 500, the online decision-making time of the proposed distributed AI method only slowly increased from 4.2 milliseconds to 7.5 milliseconds, an increase of 1.8 times, maintaining a millisecond-level response speed. This is because its computational burden is completely decentralized to each edge controller, and the forward computation of the DRL network of each controller has constant complexity. The decision-making time of centralized optimization increases super linearly with the problem size, surging from 105.3 milliseconds to over 10 seconds (which is no longer sufficient for real-time scheduling requirements at 500 units), and the time ratio soared from 25.1 to over 1333.3. This confirms the “curse of dimensionality” problem faced by centralized optimization.

Figure 6 shows the training convergence curve of the DRL agent. The average reward increases with training steps and gradually stabilizes, indicating that the agent successfully learns an effective power dispatch policy.

Table 4 Communication data volume per scheduling cycle under different DER scales

Number of DER Aggregated Units	Proposed Method	Centralized	Communication Data Volume Ratio
	Communication Data Volume (KB)	Communication Demand Data Volume (KB)	
50	12.5	250.0	0.05
100	15.0	500.0	0.03
200	20.0	1000.0	0.02
300	25.0	1500.0	0.017
500	35.0	2500.0	0.014

Furthermore, to analyze communication scalability, the total number of bytes of data exchanged between the coordination layer and all distributed units was recorded in the simulation within each scheduling cycle (15 seconds). The communication overhead of the proposed hierarchical architecture and the ideal global information collection (centralized communication requirement) under different DER scales was tested, and the results are shown in Table 4.

The proposed architecture exhibits excellent scalability in terms of communication overhead. Its communication data volume only increases slowly from 12.5 KB to 35.0 KB, primarily due to the status information reported by newly added units, while the size of the LMP vector broadcast from the upper layer remains fixed. In contrast, the centralized communication data requirement is proportional to the number of DERs, linearly increasing from 250 KB to 2500 KB. The communication data volume ratio of the proposed method continuously decreases from 0.05 to 0.014, indicating that the larger the system scale, the more significant its advantage in communication efficiency. This is entirely due to the communication mode described in Section 3.5: only periodically broadcasting price signals and uploading necessary status information, avoiding continuous and large-scale global information exchange between all distributed units or between the system and the center, thus fundamentally solving the communication bottleneck problem and providing feasibility for building ultra-large-scale distributed energy systems.

Conclusions

This paper proposes a hierarchical distributed power dispatch automation method based on artificial intelligence. Combining graph neural networks and deep reinforcement learning, it achieves a balance between global

coordination and local autonomous decision-making, significantly reducing communication and computational complexity while ensuring economy and security. This method demonstrates excellent performance in photovoltaic grid integration, robustness to projected tracking, and large-scale scalability. However, the training phase relies on a high-fidelity simulation environment and neglects the impact of data quality. Future work can focus on developing more efficient model-free transfer learning algorithms, exploring blockchain technology to enhance the security and trustworthiness of distributed transactions, and promoting the verification and application of this architecture on real hardware-in-the-loop test platforms.

Declaration

Data Availability

Not applicable.

(The study is based on simulation experiments and algorithmic modeling. No publicly accessible datasets were generated or analyzed.)

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Funding Statement

This research received no specific grant from any public, commercial, or not-for-profit funding agencies.

Author Contribution

- **Zhaoyuan Yin:** Conceptualization, system architecture design, algorithm development (DRL and GNN integration), simulation experiments, manuscript drafting.
- **Xiongbao Zhang:** Model implementation, data analysis, performance evaluation, manuscript revision.
- **Wanhe Luo:** Distributed system modeling, scalability analysis, communication overhead evaluation.
- **Shidi Ruan:** Technical validation, results interpretation, support in experimental design and manuscript editing.

All authors read and approved the final manuscript.

Ethical Approval

Not applicable.

(The study does not involve human participants, animal experiments, or clinical/medical research.)

Consent to Participate

Not applicable.

Consent for Publication

All authors provide consent for publication of this manuscript.

Competing Interests

The authors declare no competing interests.

References

- [1] Ehsan Naderi, Leila Mirzaei, Mahdi Pourakbari-Kasmaei, et al. Optimization of active power dispatch considering unified power flow controller: application of evolutionary algorithms in a fuzzy framework. *Evolutionary Intelligence*, 17(3):1357–1387, 2024.
- [2] Hakan Yapici. Solution of optimal reactive power dispatch problem using pathfinder algorithm. *Engineering Optimization*, 53(11):1946–1963, 2021.
- [3] Sofiane Mouassa, Francisco Jurado, Toufik Bouktir, et al. Novel design of artificial ecosystem optimizer for large-scale optimal reactive power dispatch problem with application to Algerian electricity grid. *Neural Computing and Applications*, 33(13):7467–7490, 2021.
- [4] Tarek Ibrahim, Thiago T. De Rubira, Alessandro Del Rosso, et al. Alternating optimization approach for voltage-secure multi-period optimal reactive power dispatch. *IEEE Transactions on Power Systems*, 37(5):3805–3816, 2021.
- [5] Yue, L., Liang, X., Sun, L., Li, Y., and Cheng, L. Research on Collaborative Control Strategy of Virtual Power Plant Based on Deep Reinforcement Learning Framework. *Distributed Generation & Alternative Energy Journal*, 40(3), 533–558, 2025.
- [6] Nabil Agouzoul, Abdelali Oukennou, Fatima Elmariami, et al. Power efficiency improvement in reactive power dispatch under load

- uncertainty. *International Journal of Electrical and Computer Engineering*, 14(4): 3616–3627, 2024.
- [7] Eric N. Odonkor, Patrick M. Moses, and Andrew O. Akumu. Intelligent ANFIS-based distributed generators energy control and power dispatch of grid-connected microgrids integrated into distribution network. *International Journal of Electrical and Electronic Engineering & Telecommunications*, 13(2): 112–124, 2024.
- [8] Ahsan Ali, Shahid Aslam, Sadegh Mirsaeidi, et al. Multi-objective multiperiod stable environmental economic power dispatch considering probabilistic wind and solar PV generation. *IET Renewable Power Generation*, 18(16):3903–3922, 2024.
- [9] Rana, A. S., Bhagyasree, B. B., Harini, T. M., Sreekumar, S., and Raju, M. Optimisation of Economic and Environmental Dispatch of Power System with and without Renewable Energy Sources. *Distributed Generation & Alternative Energy Journal*, 38(2) , 491–518, 2023.
- [10] Mohamed H. Ali, Ahmed M. A. Soliman, and Ahmed H. Adel. Optimization of reactive power dispatch considering DG units uncertainty by dandelion optimizer algorithm. *International Journal of Renewable Energy Research*, 12(4):1805–1818, 2022.
- [11] Andres Casilimas-Peña, Kyle Rosen, and Carlos Angeles-Camacho. Enhancing microgrid performance: optimal proactive reactive power dispatch using photovoltaic active power forecasts. *IET Smart Grid*, 7(6):891–903, 2024.
- [12] Kashem M. Muttaqi and Darmawan Sutanto. Adaptive and predictive energy management strategy for real-time optimal power dispatch from VPPs integrated with renewable energy and energy storage. *IEEE Transactions on Industry Applications*, 57(3):1958–1972, 2021.
- [13] Ahmed Abaza, Ahmed Fawzy, Ragab A. El-Sehiemy, et al. Sensitive reactive power dispatch solution accomplished with renewable energy allocation using an enhanced coyote optimization algorithm. *Ain Shams Engineering Journal*, 12(2):1723–1739, 2021.
- [14] Kumari, S., Sreekumar, S., Singh, S., and Kothari, D. P. Wind Power Deviation Charge Reduction using Machine Learning. *Distributed Generation & Alternative Energy Journal*, 39, 27–56, 2024.
- [15] Ahmed M. Abd-El Wahab, Sherif Kamel, Hany M. Sultan, et al. Optimizing reactive power dispatch in electrical networks using a hybrid artificial rabbits and gradient-based optimization. *Electrical Engineering*, 106(4):3823–3851, 2024.

- [16] Chen, J., and Hu, N. Distributed Optimization Model for Economic Dispatch of Smart Grid. *Distributed Generation & Alternative Energy Journal*, 40(3):457–480, 2025.
- [17] Balasubramanyam, P., and Sood, V. K. A novel hybrid swarm intelligence and cuckoo search based microgrid EMS for optimal energy scheduling. *Distributed Generation and Alternative Energy Journal*, 38(4):1119–1148, 2023.
- [18] Zhu, H., Li, Z., Chen, S., and Peng, X. Application of demand-side technology in power system intelligent regulation. *Distributed Generation and Alternative Energy Journal*, 37(2):145–158, 2022.
- [19] Fan, J., Zhou, Z., Ma, J., Wen, Y., Wan, H., and Meng, J. Research on optimization of intelligent data-driven monitoring and status evaluation mechanism for distribution network and distributed resources. *Distributed Generation and Alternative Energy Journal*, 39(6):1153–1178, 2024.
- [20] Samuel A. Adegoke and Yong Sun. Optimum reactive power dispatch solution using hybrid particle swarm optimization and pathfinder algorithm. *International Journal of Computation*, 21(4):403–410, 2022.
- [21] Mohamed A. M. Shaheen, Hany M. Hasanien, and Abdulrahman Alkuhayli. A novel hybrid GWO-PSO optimization technique for optimal reactive power dispatch problem solution. *Ain Shams Engineering Journal*, 12(1):621–630, 2021.
- [22] Samuel A. Adegoke, Yong Sun, Zhiyong Wang, et al. A mini review on optimal reactive power dispatch incorporating renewable energy sources and flexible alternating current transmission system. *Electrical Engineering*, 106(4):3961–3982, 2024.
- [23] Nikhil Chopra, Yadwinder S. Brar, and Jaspreet S. Dhillon. An improved particle swarm optimization using simplex-based deterministic approach for economic-emission power dispatch problem. *Electrical Engineering*, 103(3):1347–1365, 2021.
- [24] Marcos V. da Silva, Juan M. H. Ortiz, Mahdi Pourakbari-Kasmaei, et al. Convex formulation for optimal active and reactive power dispatch. *IEEE Latin America Transactions*, 20(5):787–798, 2022.
- [25] Suresh K. Gupta, Lalit Kumar, Manoj K. Kar, et al. Optimal reactive power dispatch under coordinated active and reactive load variations using FACTS devices. *International Journal of System Assurance Engineering and Management*, 13(5):2672–2682, 2022.

- [26] Mohamed Abd-El Wahab, Sherif Kamel, Mohamed H. Hassan, et al. Jaya-AEO: an innovative hybrid optimizer for reactive power dispatch optimization in power systems. *Electric Power Components and Systems*, 52(4):509–531, 2024.
- [27] Riad Kouadri, Sofiane Mouassa, and Francisco Jurado. Optimal power dispatch in hybrid power system for medium- and large-scale practical power systems using self-adaptive bonobo optimizer algorithm. *Wind Engineering*, 48(6):1118–1140, 2024.
- [28] Rodrigo Rodrigues Lautert, Carlos A. C. Cambambi, Marcos S. Ortiz, et al. Optimal power dispatch in microgrids using mixed-integer linear programming. *at-Automatisierungstechnik*, 72(11):1030–1040, 2024.
- [29] Sreejith S. Kumar and Nandakumar Mukundan. Modified LMS control for a grid interactive PV–fuel cell–electrolyzer hybrid system with power dispatch to the grid. *IEEE Transactions on Industry Applications*, 58(6):7907–7918, 2022.
- [30] Fabio J. Lachovicz, Thiago S. P. Fernandes, and Jose A. Vilela Junior. Impacts of PV-STATCOM reactive power dispatch in the allocation of capacitor banks and voltage regulators on active distribution networks. *Journal of Control, Automation and Electrical Systems*, 34(4):796–807, 2023.

Biographies



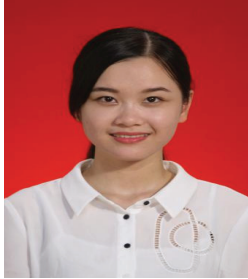
Zhaoyuan Yin was born in China in 1997. He graduated from Monash University, Australia in 2022 and obtained a Master of Engineering (M.Eng.) degree. Currently, he works as an engineer at Guangxi Power Grid Co., Ltd., China, with his main research directions focusing on dispatching automation digital technology.



Xiongbao Zhang was born in Yulin City, Guangxi Zhuang Autonomous Region, P.R. China in 1990. He graduated from Guangxi University, China in 2017 and obtained a Master of Science in Engineering (M.S.E.) degree. Currently, he works as an engineer at Guangxi Power Grid Co., Ltd., China, with his main research directions focusing on dispatching automation digital technology and cyber security.



Wanhe Luo was born in Huizhou, Guangdong Province, China in 1996. He graduated from Guangxi University with a bachelor's degree. Currently, he works at the Marketing and Distribution Command Center of Nanning Shanglin Power Supply Bureau, Guangxi Power Grid Co., Ltd. His main research focus is distribution network automation.



Shidi Ruan was born in Nanning, Guangxi, P.R. China, in 1989. She received the Master degree from North China Electric Power University, P.R. China. Now, she works in Guangxi Power Grid Power Dispatching and Control Center. Her research interests include power dispatching automation

