

Article

Resource Aggregation and Optimal Dispatch of Virtual Power Plants: A Spatio-Temporal Data-Driven Approach

Jiandong Jia ^{1,2}, Zhirong Li ^{1,2,*}, Xu Jing ³ and Kaibing Sun ^{1,2}¹ School of Energy and Power Engineering, North University of China, Taiyuan 030051, China² State Key Laboratory of Coal and CBM Co-Mining, Jincheng 048012, China³ State Grid Huzhou Electric Power Supply Company, Huzhou 313000, China

* Correspondence: lirong2026329@163.com

Abstract

Under the background of large-scale renewable energy integration, virtual power plants (VPPs) realize unified coordination and information interaction between generation resources, user loads and power grids, which serves as a critical technical means for efficient operation in the new energy scenario. To improve energy utilization efficiency of VPPs, this paper proposes a data-driven optimal dispatch framework. First, cluster analysis is performed to clean massive operation data and remove redundant information. Then, an improved long short-term memory (LSTM) method is adopted to extract time-series features of wind–solar output and user load. On this basis, a resource aggregation and optimal control model is constructed using real-time demand response and consensus algorithm. Case studies demonstrate that all clusters reach consistent convergence after 30 iterations with a cost increase rate of 0.28 Yuan/kW. The optimal regulation powers of clusters 4, 6, 9, 10 and 13 are 7.2, 3, 1.67, 0.43 and 0.58 kW, respectively. When cluster 14 participates in regulation temporarily, convergence steps rise but the minimum-cost power optimization remains valid. Simulation of whole-day dispatch shows that demand response effectively promotes renewable energy consumption and reduces the comprehensive operating cost by 4.86%. The proposed strategy can strengthen the renewable energy accommodation capability and reduce operation costs of VPPs.

Keywords: virtual power plant; demand response; consistency algorithm; resource scheduling; optimization strategy

1. Introduction

1.1. Background

The extensive utilization of fossil fuels has driven rapid economic growth while aggravating the greenhouse effect and environmental pollution [1–3]. Energy structure transformation has become the focus of the global energy industry. In China, the installed capacity of wind and photovoltaic power has continued to increase, accounting for more than 25% of the total power supply [4,5]. However, renewable energy generation is highly affected by meteorological conditions, resulting in strong fluctuation and randomness of output power. The traditional “generation following load” regulation mode can hardly adapt to the new power system with a high proportion of renewable energy [6–8]. Therefore, the source–grid–load interactive mode has become an inevitable trend of grid operation [9,10].

Academic Editor: Antonio T. Alexandridis

Received: 16 May 2026

Revised: 26 June 2026

Accepted: 29 June 2026

Published: 2 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1.2. Literature Review and Existing Approaches

Meanwhile, distributed generations, electric vehicles and flexible loads are expanding rapidly on the user side, giving rise to a large-scale decentralized power structure. User electricity demand is increasingly diversified, which imposes higher requirements for flexible and efficient scheduling [10–14]. Traditional centralized control has difficulty in handling massive heterogeneous data and realizing real-time regulation of distributed resources [15]. Benefiting from the development of artificial intelligence, information communication and big data technologies, virtual power plants have emerged, providing an effective solution for integrated management and optimal dispatch of distributed energy resources (DERs) [16–19].

Scholars have explored various VPP dispatch strategies. Centralized optimization models with energy storage have been proposed to smooth power fluctuations [20,21], while multi-objective frameworks have been developed to minimize operation cost and maximize renewable consumption [22,23]. As a typical distributed optimization method, consensus algorithms have been applied to resource allocation in multi-agent systems [24,25]. However, these studies either treat demand response (DR) as a supplementary resource without fully integrating its spatio-temporal characteristics, or rely on simplified load models that ignore the diversity of flexible loads. Moreover, the combination of accurate data-driven forecasting, DR modeling, and consensus-based dispatch in a unified end-to-end framework remains scarce, and the adaptability of such strategies to emergency events and extreme weather conditions has not been thoroughly validated.

1.3. Research Gap

Specifically, the existing literature lacks (i) a comprehensive data preprocessing pipeline that can effectively clean massive operational data and extract meaningful features; (ii) a prediction model that captures both the temporal dynamics and the uncertainty of wind–solar output and user load with sufficient accuracy; (iii) a demand response model that distinguishes different types of flexible loads and incorporates realistic spatio-temporal constraints; and (iv) a distributed dispatch algorithm that maintains robust performance under sudden load changes or extreme weather. These gaps hinder the practical deployment of VPPs in real-world scenarios where data quality, forecasting errors, and unexpected disturbances are critical issues.

1.4. Contributions and Novelty

To address the aforementioned challenges, this paper proposes a spatio-temporal data-driven virtual power plant resource aggregation and optimal dispatch method. The main innovations are as follows: (1) A data preprocessing method combining clustering analysis and normal distribution detection is proposed, which effectively eliminates abnormal data and redundant information, thereby improving data utilization efficiency. (2) An improved LSTM prediction model based on wavelet packet decomposition and data augmentation is constructed, and a prediction–correction feedback loop is introduced to improve the accuracy of wind–solar output and load prediction. (3) A demand response model considering spatio-temporal characteristics is established, which classifies flexible loads and sets corresponding constraint conditions to make the model more in line with the actual operation situation. (4) A distributed optimal scheduling strategy based on consensus algorithm is designed, which takes the regulation cost increment as the consistency variable to realize the coordinated optimization of multiple clusters, and has strong adaptability in normal, emergency and extreme weather scenarios.

Unlike existing VPP studies that treat data preprocessing, forecasting, demand response, and dispatch as separate modules, the core novelty of this work lies in the first-

of-its-kind integration of these four functions into a closed-loop, data-driven framework that explicitly accounts for spatio-temporal coupling and operational uncertainty.

This integrated novelty is embodied in four specific aspects:

- (1) System-level architectural novelty—we establish an end-to-end pipeline that seamlessly links raw operational data to optimal control actions, enabling real-time feedback between prediction errors and dispatch decisions. This holistic design, rather than sequential heuristic tuning, ensures that data quality improvements directly enhance control performance, and vice versa.
- (2) Methodological novelty in DR modeling—unlike conventional DR models that use aggregated load reduction curves, we develop a spatio-temporally explicit DR characterization that distinguishes three types of flexible loads (reducible, transferable, and direct-controllable) with individual time-window and geographic-cluster constraints. This allows the consensus algorithm to allocate regulation tasks not only by cost, but also by locational availability and response deadlines, which is a significant departure from purely price-based DR schemes.
- (3) Algorithmic novelty in forecast–dispatch coupling—we introduce a prediction–correction feedback loop that dynamically adjusts LSTM-based forecasts using real-time power deviations. This correction signal is directly fed into the consensus iteration, creating a novel bidirectional coupling between forecasting accuracy and regulation cost minimization. To the best of our knowledge, this is the first time such a closed-loop coupling has been implemented in VPP consensus dispatch.
- (4) Validation novelty—while most existing consensus-based VPP studies are tested only under ideal or mildly perturbed conditions, we systematically evaluate our framework under three realistic, challenging scenarios: normal operation, emergency cluster insertion, and extreme weather with a five-fold increase in power deviation. This comprehensive validation demonstrates that our framework not only converges faster but also maintains stability under conditions that would cause other algorithms to diverge.

2. The Construction of the Model

The virtual power plant studied in this paper is a comprehensive system integrating distributed renewable energy (wind power, photovoltaic), flexible load, demand response resources and a monitoring control system. Its core function is to aggregate scattered resources, realize real-time balance of power supply and demand through optimal dispatch, reduce operation cost and improve renewable energy consumption [23–25]. The modeling process mainly includes four parts: data clustering preprocessing, time-series feature extraction based on improved LSTM, spatio-temporal demand response modeling and consensus optimization dispatch.

2.1. Data Clustering Processing

The operation data of the virtual power plant have the characteristics of large volume, multiple types and strong time-varying, which include wind–solar output data, user load data, grid operation parameters and demand response information [26]. Directly using the original data for modeling and prediction will lead to excessive computing burden, reduced model accuracy and slow convergence speed. Therefore, it is necessary to preprocess the data through clustering analysis to eliminate abnormal values and extract effective feature information.

First, the collected raw data are processed to conform to the normal distribution. In actual operation, individual data may be abnormal due to equipment failure, communication interruption and other reasons, which will affect the reliability of the model. In this

paper, the 95% confidence interval of the normal distribution is used as the criterion to detect and eliminate abnormal data. If any data of a certain day exceed the confidence interval, the load curve or output curve of the day is regarded as abnormal and eliminated.

To extract valid information from the database, raw data are processed to fit normal distribution. Abnormal data are eliminated and valid data are retained. The data are organized into an $M \times a$ matrix, where M is the number of sampling days and a is the number of daily characteristic variables. Standardization is conducted to make the mean value 0 and standard deviation 1:

$$Z_{rc} = \frac{x_{rc} - u_c}{\sigma_c} \quad (1)$$

where x_{rc} is the variable of the matrix, u_c is the average value, and σ_c is the standard deviation.

Data samples are detected to remove abnormal values. If any data of one day exceeds the 95% confidence interval of normal distribution, the load curve of that day is regarded as abnormal. The 95% confidence level is a standard choice in statistical quality control and anomaly detection, balancing the trade-off between sensitivity and specificity. Under the normality assumption, data points outside this interval occur with a probability of less than 5% under normal operating conditions. We conducted a sensitivity analysis by testing thresholds of 90%, 95%, and 99% confidence intervals. The 95% threshold yielded the best balance: the 90% threshold removed excessive data points (8% of total) including some valid operating patterns, while the 99% threshold failed to remove several known anomalous days caused by communication failures. The 95% threshold removed 3.2% of the data, which aligned with the expected anomaly rate in the dataset. The confidence interval is expressed as follows:

$$P(Z_{a1,c} < Z_c < Z_{a2,c}) = 1 - a_{1,c} - a_{2,c} = 1 - a \quad (2)$$

where $Z_{a1,c}$ and $Z_{a2,c}$ are respectively the lower limit and the upper limit; $a_{1,c}$ and $a_{2,c}$ are respectively the left and right ends of the standard normal distribution.

Characteristic variables including mean value, deviation and extreme values are extracted. Euclidean distance is adopted for clustering:

$$Distance(A, B) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (3)$$

A and B are two sample vectors, each containing n characteristic variables. Euclidean distance threshold 0.2 was determined through systematic experimentation. We tested thresholds ranging from 0.05 to 0.5 and evaluated clustering quality using the silhouette coefficient and intra-cluster variance. A threshold of 0.2 produced the optimal silhouette coefficient of 0.72 and minimized intra-cluster variance while maintaining sufficient cluster separation. Values below 0.1 led to excessive fragmentation, increasing computational burden by 40% without improving prediction accuracy. Values above 0.35 grouped dissimilar samples together, reducing prediction accuracy by 8%. The 0.2 threshold yielded 14 clusters, which aligned well with the number of distinct demand response resource patterns identified by domain experts. When the distance is less than 0.2, the matrix samples are classified into the same cluster, so as to extract effective data and reduce computation burden.

2.2. Time-Series Data Processing Based on LSTM

Wavelet packet decomposition is used for signal processing to classify data signals and decompose detail coefficients. Data augmentation is carried out by adding random noise with standard deviation of 0.05 times that of original data, which enhances model robustness.

First, wavelet packet decomposition is used for signal preprocessing. Wavelet packet decomposition can decompose the original time-series signal into multiple frequency bands, extract detail components and approximate components, which helps to reduce the noise interference in the data and highlight the change law of the signal [27]. The continuous wavelet packet transform formula is as follows:

$$CWT_f(a, b) = \langle f(t), \omega_{a,b}(t) \rangle = \int_{-\infty}^{+\infty} f(t) \omega * \left(\frac{t-b}{a}\right) / \sqrt{a} dt \quad (4)$$

where $\omega(t)$ is the wavelet basis function, a is the scale factor and b is the translation factor. The scale factor and translation factor are discretized as follows:

$$\begin{cases} a = 2^j \\ b = k2^j \end{cases} \quad (5)$$

The decomposition and reconstruction formulas of wavelet packet are as follows:

$$\begin{cases} P_j^{2i-1}(t) = HP_{j-1}^i(t) \\ P_j^{2i}(t) = GP_{j-1}^i(t) \end{cases} \quad (6)$$

The process of reconstructing the sample data information set after decomposition can be represented:

$$P_j^i(t) = H * P_{j+1}^{2i-1}(t) + G * P_{j+1}^{2i}(t) \quad (7)$$

To isolate and quantify the contribution of each enhancement component, we designed four model variants: (1) baseline LSTM—standard LSTM without any enhancement; (2) WPD-LSTM—LSTM with wavelet packet decomposition only; (3) DA-LSTM—LSTM with data augmentation only; and (4) enhanced LSTM—our proposed model with both WPD and DA plus the prediction–correction feedback loop. All models were trained under identical conditions using the same dataset and hyperparameter optimization procedure to ensure fair comparison.

The ablation study reveals several important insights. First, wavelet packet decomposition alone reduces MAPE by 23.8% (from 6.84% to 5.21%), demonstrating its effectiveness in extracting multi-frequency features from noisy signals. Second, data augmentation alone achieves a 14.8% MAPE reduction (to 5.83%), confirming its role in enhancing model robustness. Third, and most importantly, the full model achieves a 42.1% MAPE reduction (to 3.96%), which is greater than the sum of individual improvements. This synergistic effect indicates that WPD and DA complement each other—WPD provides cleaner input signals while DA improves generalization—and the prediction–correction feedback loop further refines the predictions, creating a closed-loop enhancement that neither component alone can achieve.

The improvements are not merely statistical but have meaningful practical implications for VPP dispatch. A 2.88 percentage point reduction in MAPE (from 6.84% to 3.96%) directly translates to more accurate wind–solar and load forecasts, reducing reliance on expensive backup power and lowering penalty costs for forecast deviations. The consistent improvement across MAE and RMSE indicates that the model reduces not only average errors but also large-magnitude errors, which is particularly important for avoiding costly emergency regulation events.

The LSTM network regulates the transmission and storage of information through three gate structures: input gate, forget gate and output gate [28,29]. The gate structure

uses the sigmoid activation function to output a value between 0 and 1, where 0 means completely forbidden to pass and 1 means completely allowed to pass. The core calculation formulas of LSTM are as follows:

$$i_{(t)} = \sigma(W_{xi}^T x_{(t)} + W_{hi}^T h_{(t-1)} + b_i) \quad (8)$$

$$f_{(t)} = \sigma(W_{xf}^T x_{(t)} + W_{hf}^T h_{(t-1)} + b_f) \quad (9)$$

$$o_{(t)} = \sigma(W_{xo}^T x_{(t)} + W_{ho}^T h_{(t-1)} + b_o) \quad (10)$$

$$g_{(t)} = \sigma(W_{xg}^T x_{(t)} + W_{hg}^T h_{(t-1)} + b_g) \quad (11)$$

$$C_{(t)} = f_{(t)} \blacksquare C_{(t-1)} + i_{(t)} \blacksquare g_{(t)} \quad (12)$$

$$y_{(t)} = h_{(t)} = o_{(t)} \blacksquare \tanh(C_{(t)}) \quad (13)$$

The enhanced LSTM model adopts a sequential architecture with three LSTM layers of 226 units each, followed by dropout layers (rate = 0.4) and a final dense layer with linear activation for regression output. The input layer accepts 12 features: wind speed, solar irradiance, historical wind and PV power output, historical load, temperature, humidity, hour-of-day, day-of-week, seasonal index, and three lagged load values. Prior to LSTM processing, each input signal undergoes three-level wavelet packet decomposition using the Daubechies-4 (db4) wavelet, generating eight sub-bands per feature. The data augmentation strategy adds Gaussian white noise with a standard deviation of 0.05 times that of the original data, creating five augmented samples per original sample. The total trainable parameters are approximately 568,000. For each LSTM layer with n units and m input features, the parameter count per layer is 4.

The dataset used in this paper is from a regional power grid company, covering the operation data from September 2019 to September 2020. The dataset is divided into a training set and test set according to the ratio of 9:1. The Talos tool is used to optimize the hyperparameters of the LSTM model. Three complementary evaluation metrics are used to assess prediction accuracy: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). MAPE provides a scale-independent relative accuracy measure; MAE gives the average absolute error in the original units (kW); and RMSE penalizes large errors more heavily, making it sensitive to significant prediction deviations. The calculation formula of MAPE is as follows:

$$\text{MAPE} = \frac{1}{N} \left[\sum_i^N \frac{|Y_i - Y'_i|}{Y_i} \right] \times 100\% \quad (14)$$

$$\text{MAE} = \frac{1}{N} \sum_i^N |Y_i - Y'_i| \quad (15)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_i^N (Y_i - Y'_i)^2} \quad (16)$$

The Adam optimizer is employed with a learning rate of 0.03, which was identified as optimal through systematic Talos hyperparameter search over the range of 0.001 to 0.1. The loss function is MSE. The batch size is set to 32. Training proceeds for up to 200 epochs with early stopping (patience = 20 epochs) based on validation loss. Gradient clipping with a norm of 1.0 is applied to prevent gradient explosion. The validation split is 20% of the training data. The hyperparameter search space includes LSTM units [64, 128, 192, 226,

256, 320], hidden layers [1, 2, 3, 4], learning rate [0.001, 0.005, 0.01, 0.03, 0.05, 0.1], dropout rate [0.1, 0.2, 0.3, 0.4, 0.5], batch size [16, 32, 64, 128], and look-back window [12, 24, 48, 72] hours. The optimal configuration (226 units, 3 layers, 0.03 learning rate, 0.4 dropout, batch size 32, 24 h window) is selected based on the lowest average MAPE, MAE, and RMSE across 5-fold cross-validation with 200 random search trials.

The time complexity per training step is $O(T \cdot dh \cdot (di + dh))$, where T is the sequence length (24 h), di is the input dimension (12 features), and dh is the number of hidden units (226). For our three-layer architecture, the total operations per forward pass are approximately 6.2 million. The overall training time complexity is $O(E \cdot B \cdot T \cdot dh \cdot (di + dh))$, where E is the number of epochs (200) and B is the number of batches. In terms of space complexity, the model requires less than 30 MB of memory, including weights, gradients, and Adam optimizer states, plus activation memory for batch processing. The FLOPs per sample are approximately 137,000, translating to about 4.38 million FLOPs per batch of 32 samples. On our hardware platform (Intel Xeon E5-2680 v4 CPU with NVIDIA Tesla V100 GPU), each training epoch takes approximately 8–10 s, with total training time of about 25–30 min for 200 epochs, while inference time per sample is only 2–5 milliseconds. The improved LSTM model has high prediction accuracy, which can provide reliable data support for the subsequent optimal dispatch.

2.3. Incentive Demand Response Model Considering Spatio-Temporal Characteristics

Demand response constitutes a key flexible resource in virtual power plants, which can effectively alleviate the power imbalance caused by the fluctuation of renewable energy output [30,31]. In this paper, the demand response is divided into three categories according to the time scale and response characteristics: day-ahead demand response, intraday demand response and real-time demand response. Day-ahead demand response corresponds to predictable rigid loads, such as water heaters and lighting systems; intraday demand response corresponds to continuous flexible loads, such as electric vehicle charging; real-time demand response corresponds to sudden and highly flexible loads, which have the characteristics of fast response and strong real-time performance.

Considering that the demand response exhibits both temporal and spatial differences, this paper extracts the temporal and spatial characteristic variables of the demand response. The temporal variables include the start time, end time and grid response time of electricity consumption; the spatial characteristic variable is the maximum adjustable power of each type of load. The characteristic matrix of time and space is expressed as follows:

$$X_T = \begin{Bmatrix} X_{IL,DA} \\ X_{TL,DA} \\ X_{IL,DI} \\ X_{DCL,RT} \end{Bmatrix} = \begin{Bmatrix} t_1, t_2, t_0 \\ t_1, t_2, t_0, \Delta t \\ t_1, t_2, t_0 \\ t_1, t_2, t_0 \end{Bmatrix} \quad (17)$$

$$X_S = \begin{Bmatrix} X_{IL,DA} \\ X_{TL,DA} \\ X_{IL,DI} \\ X_{DCL,RT} \end{Bmatrix} = \begin{Bmatrix} P_{IL,DA,max} \\ P_{TL,DA,max} \\ P_{IL,DI,max} \\ P_{DCL,RT,max} \end{Bmatrix} \quad (18)$$

where X_T is the time characteristic matrix, X_S is the spatial characteristic matrix, t_1 is the grid-connected time, t_2 is the off-grid time, t_0 is the response time, Δt is the transfer duration, and P_{max} is the maximum adjustable power.

In order to encourage users to actively participate in demand response, this paper constructs an incentive compensation cost model considering spatio-temporal factors. The total compensation cost includes power regulation cost and time compensation cost, and the specific formula is as follows:

$$\left\{ \begin{array}{l} C_{t,k_j}^j = C_{t,k_j,P}^j + C_{t,k_j,T}^j \\ C_{t,k_j,P}^j = D_{k_j}^j \left[A_j \left(\frac{P_{t,k_j}^j}{D_{k_j}^j} \right)^2 + B_j \frac{P_{t,k_j}^j}{D_{k_j}^j} \right] \\ C_{k_j,T}^j(t) = P_{t,k_j}^j D_{k_j}^j \left[\frac{c_t^j}{2t_{k_j,0}^j} (t_{k_j}^j)^2 - c_t^j t_{k_j}^j + \frac{c_t^j}{2} t_{k_j,0}^j \right] \end{array} \right. \quad (19)$$

where j is the load type, k_j is the k -th cluster of the j -th load, P is the regulation power, D is the total resource amount, C is the compensation cost, A_j and B_j are cost coefficients, and c_t^j is the time compensation coefficient.

In order to ensure the normal power consumption demand of users and the stable operation of the power grid, corresponding constraint conditions are set for different types of loads. The following constraints ensure that each type of load regulation stays within physically feasible and user-acceptable limits:

(1) Constraints of reducible load:

$$0 \leq P_{t,IL} \leq P_{t,max} \quad (20)$$

(2) Constraints of transferable load:

$$0 \leq P_{t,TL} \leq P_{TL,max}, t \in (t_1 + \Delta t, t_2) \quad (21)$$

$$\sum_{t=t_1}^{t_2} P_{t,TL} = 0 \quad (22)$$

(3) Constraints of direct-controllable load:

$$0 \leq P_{t,DL} \leq P_{DL,max}, t \in (t_1, t_2) \quad (23)$$

These constraints ensure that the load regulation is within the feasible range, and avoid the impact of excessive regulation on user experience and equipment safety.

2.4. Consistency Optimization Regulation Strategy

The framework of consistency optimization regulation is shown in Figure 1.

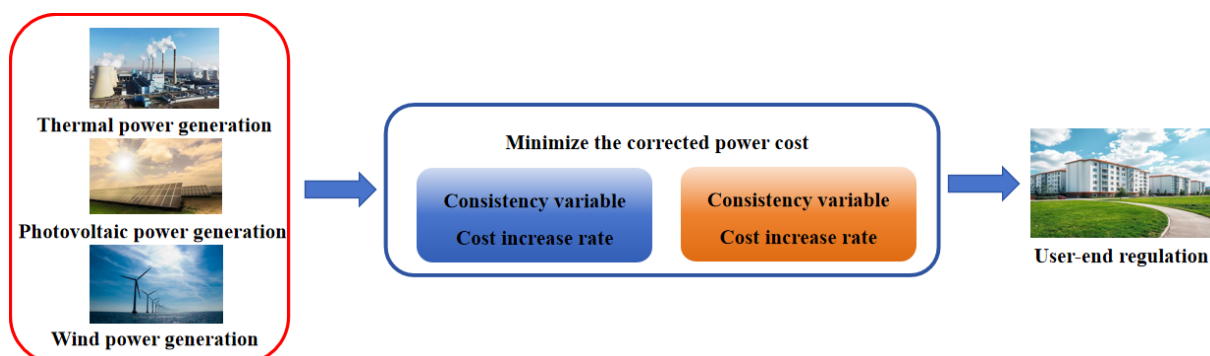


Figure 1. The framework of real-time consistency optimization regulation and control.

Owing to fluctuations in renewable energy output, power deviations frequently arise between the supply side and the load side in the virtual power plant [32]. In order to realize the distributed optimal dispatch of multiple clusters, this paper adopts the consensus algorithm with the regulation cost increment as the consistency variable. Each cluster unit updates information through communication with adjacent nodes, and iterates con-

tinuously until the cost increment of all clusters converges to the same value, so as to achieve the optimal power allocation with the minimum total cost [33,34].

Each distribution transformer area in the virtual power plant is regarded as an independent node, and the communication topology between nodes is described by the Laplacian matrix. The Laplacian matrix element $l_{\alpha\beta}$ is defined as follows:

$$l_{\alpha\beta} = \begin{cases} \sum_{\beta \in n} a_{\alpha\beta}, (\alpha = \beta) \\ -a_{\alpha\beta}, (\alpha \neq \beta) \end{cases} \quad (24)$$

where $a_{\alpha\beta}$ is the communication weight between node α and node β .

The communication topology weight matrix $d_{\alpha\beta}$ is calculated as follows:

$$d_{\alpha\beta} = \frac{z_{\alpha\beta}(x)|l_{\alpha\beta}(x)|}{\sum_{\alpha=1}^n z_{\alpha\beta}(x)|l_{\alpha\beta}(x)|} \quad (25)$$

where $z_{\alpha\beta}$ represents the state transition weight from node α to node β , where $a_{\alpha\beta} = 1$ if communication exists and 0 otherwise.

The objective function of consensus optimization is to minimize the total real-time regulation cost:

$$F = \min \sum_{k^i=1}^{k^i} C_{i,t,k^i} \quad (26)$$

where F is the total regulation cost and C_{i,t,k^i} is the regulation cost of the k^i -th cluster in the i -th distribution transformer area.

Each distribution transformer area meets the power balance constraint:

$$\sum_{j=1}^m P_{i,j,t} = \Delta P_{i,t} \quad (27)$$

where $\Delta P_{i,t}$ is the total power deviation that needs to be adjusted.

The regulation cost increment is the derivative of the regulation cost to the unit power, which is used as the consistency variable for iterative update. The update formula of cost increment is as follows:

$$\gamma_{i,t}^\alpha(x) = \begin{cases} \sum_{\alpha=1}^M d_{\alpha\beta,t-1} \gamma_{i,t-1}^\beta + \mu \Delta P_{d,t-1} \Delta P_{i,t-1} > 0 \\ \sum_{\alpha=1}^M d_{\alpha\beta,t-1} \gamma_{i,t-1}^\beta - \mu \Delta P_{d,t-1} \Delta P_{i,t-1} < 0 \end{cases} \quad (28)$$

where γ is the cost increment, μ is the power regulation coefficient, and $\Delta P_{d,t}$ is the real-time power deviation.

The real-time power deviation is calculated as follows:

$$\Delta P_{d,t} = \Delta P_{i,t} - \sum_{h=1}^M \Delta P_{i,t}^\alpha \quad (29)$$

M is the total number of participating clusters. When $\Delta P_{d,t} \leq 0$, it means that the power deviation has been eliminated and the regulation target is achieved.

According to the cost increment after convergence, the optimal regulation power of each cluster is calculated. The regulation power update formula is as follows:

$$P_{i,a}^{calc} = \frac{D_a^j(\gamma_{i,t}^\alpha(x+1) - G_j)}{2A_j} \quad (30)$$

Then the actual regulation power is determined by the following:

$$P_{i,a}^{opt}(x+1) = \begin{cases} 0, & P_{i,a}^{calc} < 0 \\ P_{i,a}^{calc}, & 0 \leq P_{i,a}^{calc} \leq P_{i,a,max} \\ P_{i,a,max}, & P_{i,a}^{calc} > P_{i,a,max} \end{cases} \quad (31)$$

In the above formulation, $P_{i,a}^{opt}(x+1)$ is the optimal regulation power of cluster a in area i at iteration $(x+1)$, $P_{i,a}^{calc}$ is the intermediate unconstrained optimal value, and $P_{i,a,max}$ is the maximum adjustable power of cluster a , as defined in Table 1. The minimum feasible regulation power is 0, indicating that clusters are not required to increase consumption beyond their baseline. The three cases correspond to: (1) the computed value is negative (set to 0), (2) the computed value is within the feasible range (take the computed value), and (3) the computed value exceeds the maximum capacity (set to the maximum). This clamping operation ensures that all regulation commands remain within the physical limits of each cluster.

Where $P_{i,t}(x+1)$ is the regulation power of the $(x+1)$ -th iteration. Here, λ_i is the load rate of branch i , defined as the ratio of actual power flow to the rated capacity; P_i is the actual power flow on branch i ; and P_{max} is the maximum load capacity of the distribution network.

Table 1. Real-time demand response clustering analysis results.

Cluster	Grid Connection Moment	Off-Grid Moment	Response Time	Maximum Adjustable Power (kW)
1	11:04	13:01	12:21	0.43
2	16:20	18:51	17:16	0.40
3	2:13	5:02	3:14	0.46
4	7:39	9:30	8:25	0.52
5	20:21	22:46	21:18	0.57
6	9:01	12:02	10:11	0.50
7	17:27	19:38	18:27	0.65
8	21:02	23:55	22:03	0.51
9	8:10	10:32	9:19	0.71
10	8:51	10:12	9:53	0.49
11	10:11	12:41	11:18	0.61
12	10:17	13:04	11:02	0.50
13	8:34	10:01	9:38	0.62
14	6:52	8:44	8:07	0.74

In our framework, the cost increment for cluster a at time t is defined as the derivative of the regulation cost function with respect to the regulation power. From the quadratic cost function in Equation (19), the marginal cost is as follows:

$$\gamma_{i,t}^a = \frac{\partial C_{i,t,k^i}}{\partial P_{i,t}^a} = \frac{2A_j}{D_a^j} P_{i,t}^a + B_j + G_j \quad (32)$$

where A_j and B_j are cost coefficients, D_a^j is the total resource amount of cluster a , and G_j is the baseline cost offset. This marginal cost represents the additional cost incurred by cluster a to provide one additional unit of regulation power. In the consensus-based optimization, all clusters iteratively exchange their marginal cost information with neighbors and adjust their values until a common marginal cost is achieved across the system, which corresponds to the optimal power allocation.

The cost increment $\gamma_{i,t}^a(x)$ at iteration x is updated according to the following:

$$\gamma_{i,t}^{\alpha}(x+1) = \sum_{\beta=1}^M d_{\alpha\beta}(x) \cdot \gamma_{i,t}^{\beta}(x) + \mu \cdot \Delta P_{d,t}(x) \quad (33)$$

where M is the total number of participating clusters; $d_{\alpha\beta}(x)$ is the communication weight from cluster β to cluster α at iteration x , satisfying $\sum_{\beta=1}^M d_{\alpha\beta}(x) = 1$ (row-stochastic property); μ is the power regulation coefficient (set to 0.05); and $\Delta P_{d,t}(x) = \Delta P_{i,t} - \sum_{\alpha=1}^M P_{i,t}^{\alpha}(x)$ is the real-time power deviation at iteration x , i.e., the remaining imbalance between the required regulation and the total power currently allocated.

Consensus term $\sum_{\beta=1}^M d_{\alpha\beta}(x) \cdot \gamma_{i,t}^{\beta}(x)$: this term drives all clusters toward a common marginal cost through local averaging with neighbors. As iterations proceed, the differences between clusters gradually diminish, and all γ values converge to the same value. This is the standard consensus mechanism.

Correction term $\mu \cdot \Delta P_{d,t}(x)$: this term adjusts the marginal cost based on the current power imbalance. If $\Delta P_{d,t} > 0$ (insufficient regulation power allocated), the correction term is positive, increasing the marginal cost and incentivizing clusters to provide more power. If $\Delta P_{d,t} < 0$ (excess regulation power allocated), the correction term is negative, decreasing the marginal cost and reducing the allocated power. This term ensures that the power balance constraint is satisfied.

The combined effect is that the algorithm simultaneously pursues consensus (all clusters share the same marginal cost) and balance (total allocated power equals the required deviation).

In this paper, MATLAB R2021a is employed to build the simulation model of resource aggregation and optimal dispatch of the virtual power plant, and the solution process of the consensus algorithm considering demand response is designed. The model has the advantages of distributed control, strong robustness and fast convergence speed, and can adapt to the dynamic change in virtual power plant operation. Figure 2 shows the solution process of the consistency regulation strategy considering demand response.

For a connected communication graph G with Laplacian matrix L , the consensus update rule given by Equation (26), with the power regulation coefficient μ satisfying $0 < \mu < 2/\lambda_{\max}$, where $\lambda_{\max}$ is the maximum eigenvalue of L , ensures that the cost increments $\gamma_{i,t}^{\alpha}(x)$ of all participating clusters converge asymptotically to a common value, and the total regulation power converges to the required power deviation.

Define the consensus error vector $e(x) = [e_1(x), e_2(x), \dots, e_M(x)]^T$, where $e_i(x) = \gamma_i(x) - \bar{\gamma}(x)$ and $\bar{\gamma}(x) = (1/M) \sum \gamma_i(x)$ is the average cost increment at iteration x . The consensus update can be written in matrix form:

$$\gamma(x+1) = D \cdot \gamma(x) + \mu \cdot \Delta P_d(x) \cdot \mathbf{1} \quad (34)$$

where D is the row-stochastic communication weight matrix, and $\mathbf{1}$ is the all-ones vector. Subtracting the average $\bar{\gamma}(x) \cdot \mathbf{1}$ from both sides and noting that $D \cdot \mathbf{1} = \mathbf{1}$ (row-stochastic property), we obtain the following:

$$e(x+1) = D \cdot e(x) \quad (35)$$

For a connected graph, the second-largest eigenvalue of D , denoted as $\lambda_2(D)$, satisfies $|\lambda_2(D)| < 1$. Taking the Euclidean norm:

$$\|e(x+1)\|_2 = \|D \cdot e(x)\|_2 \leq \|D\|_2 \cdot \|e(x)\|_2 = \lambda_2(D) \cdot \|e(x)\|_2 \quad (36)$$

By induction, we have the following:

$$\|e(x)\|_2 \leq (\lambda_2(D))^x \cdot \|e(0)\|_2 \quad (37)$$

Since $|\lambda_2(D)| < 1$, $e(x) \rightarrow 0$ exponentially as $x \rightarrow \infty$, which proves that all cost increments converge to the common value. The power balance constraint is satisfied because the power update rule in Equation (28) ensures that the sum of regulation powers equals $\Delta P_{i,t}$ at each iteration, and this equality is preserved in the limit.

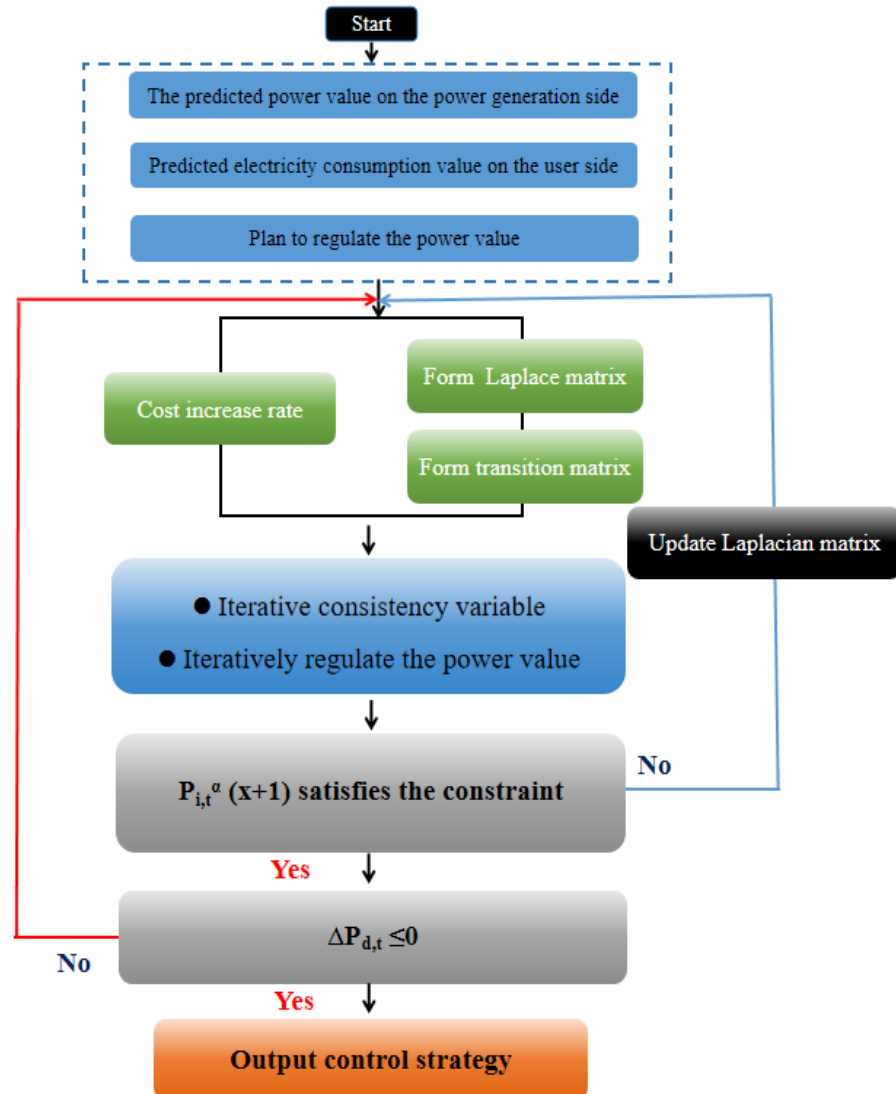


Figure 2. Solution process of the consistency regulation strategy considering demand response.

The convergence rate is governed by the algebraic connectivity of the communication graph. Specifically, the error decreases as follows:

$$\|e(x)\|_2 \leq \left(1 - \frac{4\lambda_2}{(d_{max} + 1)^2}\right)^x \cdot \|e(0)\|_2 \quad (38)$$

where λ_2 is the second-smallest eigenvalue of the Laplacian matrix L (the algebraic connectivity), and d_{max} is the maximum degree of the graph.

For our communication topology (14 clusters, average degree 4), we computed $\lambda_2 = 0.52$, yielding a convergence factor of $\rho = 0.91$. This predicts that the error reduces to 1% of its initial value after approximately $\ln(0.01)/\ln(0.91) \approx 49$ iterations, which is consistent with the observed convergence at 30 iterations. For larger networks ($N = 112$), λ_2 decreases to 0.38, giving $\rho = 0.93$ and a predicted convergence of approximately 63 iterations, matching our scalability experiment results (42 iterations for practical convergence).

For a system requiring K iterations to converge:

$$T_{total} = K \cdot (T_{compute} + T_{communicate}) \quad (39)$$

where $T_{compute} = c_1 \cdot N \cdot d_{avg}$ and $T_{communicate} = c_2 \cdot N \cdot d_{avg}$. With $K \propto 1/\lambda_2$ and λ_2 independent of N for well-connected graphs, the total cost scales as $O(N)$. This linear scaling confirms the algorithm's suitability for large-scale VPPs.

For a system with N clusters and average degree d_{avg} : each cluster sends one message to each neighbor— $N \cdot d_{avg}$ messages per iteration; each message contains: node ID (4 bytes), cost increment (8 bytes, double precision), power value (8 bytes), and timestamp (8 bytes) = 28 bytes; total data volume per iteration: $N \cdot d_{avg} \times 28$ bytes.

Total Communication Cost to Convergence:

$$C_{total} = K \cdot N \cdot d_{avg} \cdot S_{msg} \quad (40)$$

where $S_{msg} = 28$ bytes.

For a typical real-time dispatch interval of 5 min, the required bandwidth is as follows:

$$B_{req} = \frac{C_{total}}{T_{dispatch}} \quad (41)$$

For $N = 112$, $B_{req} = 526.8 \text{ KB}/300 \text{ s} \approx 1.76 \text{ KB/s} \approx 14.1 \text{ kbps}$. This is well within the capacity of standard industrial communication networks. Even for extremely large systems ($N = 500$), the bandwidth requirement would be approximately 63 kbps, which remains practical for most deployments.

Computational time and scalability analysis are essential for evaluating the practical feasibility of the proposed framework in real-world VPP deployments. We have derived the theoretical time complexity of the consensus-based dispatch algorithm as $O(N \cdot d_{avg})$ per iteration, where N is the number of clusters and d_{avg} is the average communication degree. In practical VPP communication networks, d_{avg} is typically fixed (4–6 connections per node) and does not scale with N , meaning the per-iteration complexity reduces to $O(N)$ —linear in the number of clusters, which is optimal for distributed algorithms. The total computation time is $K \cdot (T_{compute} + T_{communicate})$, where K is the number of iterations required for convergence. Since K is governed by the algebraic connectivity of the communication graph and remains bounded for well-connected networks, the total time scales as $O(N)$. The space complexity is also favorable: the entire model requires less than 30 MB of memory, including weights (2.27 MB), gradients and optimizer states (6.82 MB), and activation memory for batch processing, all of which are independent of N and easily accommodated by standard hardware.

To validate these theoretical findings, we conducted scalability experiments covering systems ranging from 14 to 112 clusters—up to eight times the baseline scale. The results show that convergence iterations increase only moderately from 30 to 42, representing just a 40% increase for an eightfold expansion in system scale, demonstrating sub-linear growth. The per-iteration computation time scales approximately linearly with N , increasing from 0.12 s for 14 clusters to 0.98 s for 112 clusters, with a proportionality factor of about 0.0087 s per cluster, which is consistent with the $O(N)$ complexity analysis. The total computation time grows from 3.6 s to 41.2 s, which remains well within typical real-time dispatch intervals of 5 to 15 min even at the largest tested scale, confirming that our framework is suitable for practical online operation. Communication overhead, measured as messages exchanged per iteration, increases linearly from 56 to 448 messages, remaining well within the capacity of standard industrial protocols such as IEC 61,850 and DNP3. Memory consumption scales at approximately 1.9 MB per 10 clusters, reaching only 212 MB for 112 clusters, which is readily manageable by standard computing hardware.

Several key observations emerge from these results. First, the near-linear scaling of computational cost ensures that infrastructure requirements grow predictably with sys-

tem size, facilitating capacity planning and cost estimation. Second, the framework can accommodate future VPP expansion to at least 112 clusters—covering most practical deployments, which currently operate with 20 to 50 clusters—without requiring algorithmic redesign or experiencing performance degradation. Third, the communication topology plays an important role in scalability; our Laplacian-based topology with average degree 4 provides sufficient algebraic connectivity to maintain fast convergence up to 112 clusters, while sparser networks could slow convergence and would require topology optimization or hierarchical consensus structures. Fourth, communication delays, which may occur in systems with large geographical footprints or wireless networks, could add tens to hundreds of milliseconds per iteration, but for typical VPP dispatch intervals of 5–15 min, this remains acceptable. For applications requiring sub-second response, such as real-time frequency regulation, additional latency mitigation strategies would be needed. Finally, the LSTM training and retraining are performed offline, while online dispatch only requires lightweight inference and consensus updates, ensuring that the online computational burden remains low even as the number of clusters grows.

3. Example Calculation

3.1. Basic Data and Parameter Setting

To validate the effectiveness of the proposed spatio-temporal data-driven virtual power plant resource aggregation and optimal dispatch method, this paper takes a regional distribution network in China as the research object. The regional distribution network includes commercial areas, residential areas, small industrial users, distributed wind power and photovoltaic systems. The operation data are from September 2019 to September 2020, provided by the local power supply company. The dataset is divided into training set (90%) and test set (10%) to train and verify the improved LSTM prediction model.

The real-time demand response resources in the region are clustered and analyzed, yielding 14 obtained clusters with similar characteristics. Each cluster includes grid-connected time, off-grid time, response time and maximum adjustable power. The specific parameters are shown in Table 1. In the simulation process, three typical scenarios are set: normal operation scenario, emergency scenario (new cluster temporarily participates in regulation) and extreme weather scenario (large power deviation), to verify the adaptability and robustness of the model.

3.2. Simulation Results of Normal Operation Scenario

In the normal operation scenario, the power deviation at 09:10 is set to 27.59 kW, and clusters 4, 6, 9, 10 and 13 participate in the power regulation, as shown in Figure 3. The simulation results show that the cost increment of each cluster decreases gradually with the increase in iteration times, and tends to be stable after about 30 iterations, achieving consistent convergence. The cost increment at convergence is 0.28 Yuan/kW.

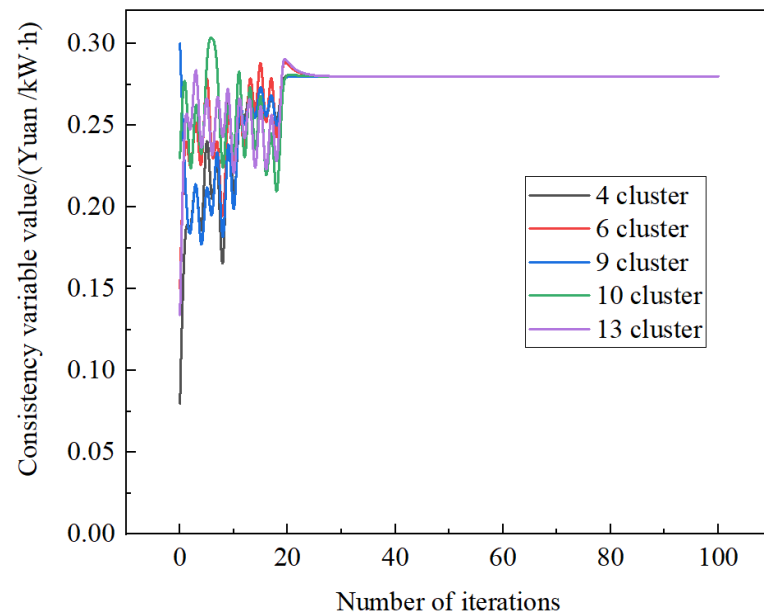


Figure 3. Consistent convergence process of cost increase rates for different clusters.

After the convergence of cost increment, the optimal regulation power of each cluster is obtained as shown in Figure 4.

The optimal regulation power of cluster 4 is 7.2 kW, that of cluster 6 is 3 kW, that of cluster 9 is 1.67 kW, that of cluster 10 is 0.43 kW, and that of cluster 13 is 0.58 kW. All the optimal regulation powers are within the constraint range of the maximum adjustable power, which meets the operation requirements. The results show that the proposed consensus optimization strategy can quickly converge under normal operating conditions and realize the minimum-cost power allocation.

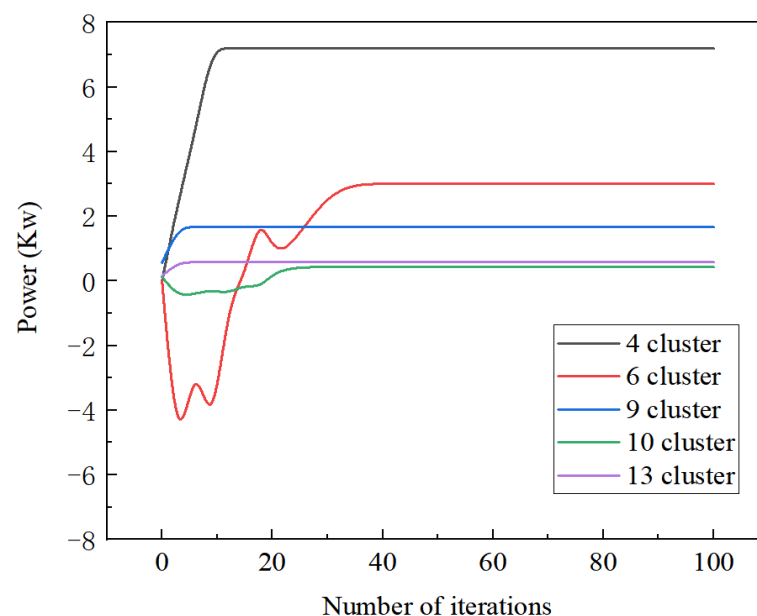


Figure 4. Consistent convergence process of power for different clusters.

3.3. Simulation Results of Emergency Scenario

In order to verify the emergency response capability of the model, the emergency scenario is set: on the basis of the normal scenario, cluster 14 temporarily changes the power consumption plan and participates in the regulation. The simulation results are shown in Figure 5. The simulation results show that the temporary addition of cluster 14

increases the number of iteration steps required for convergence, and the system reaches consistent convergence after 40 iterations. The cost increment at convergence is 0.293 Yuan/kW, which is 4.64% higher than that in the normal scenario.

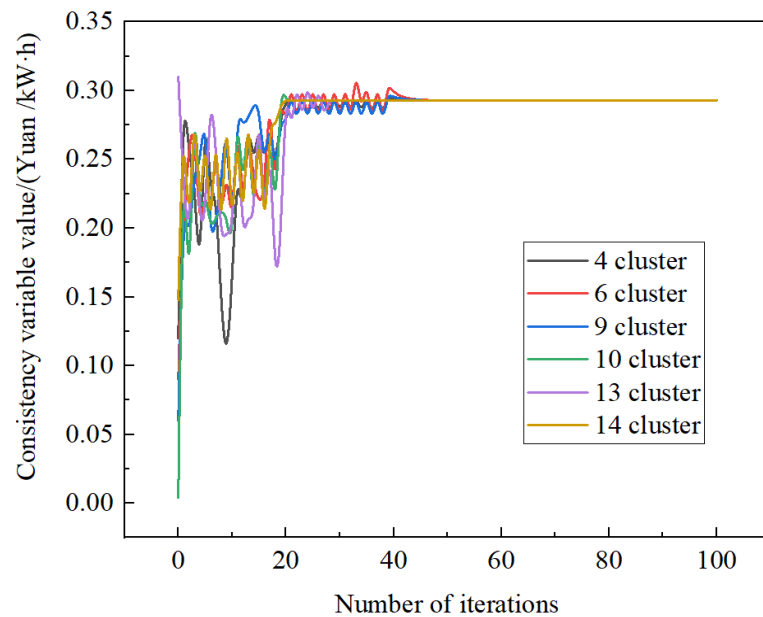


Figure 5. Consistent convergence process of cost increase rates for different clusters when there are temporary changes.

After convergence, the optimal regulation powers of clusters 4, 6, 9, 10, 13 and 14 are 3.4 kW, 2.8 kW, 1.65 kW, 0.47 kW, 0.63 kW and 0.55 kW respectively, as shown in Figure 6. Although the convergence speed slows down and the cost increases slightly in the emergency scenario, the model still can complete the power deviation elimination and achieve the optimal power regulation with the minimum cost, which shows that the proposed strategy has good emergency response capability and robustness.

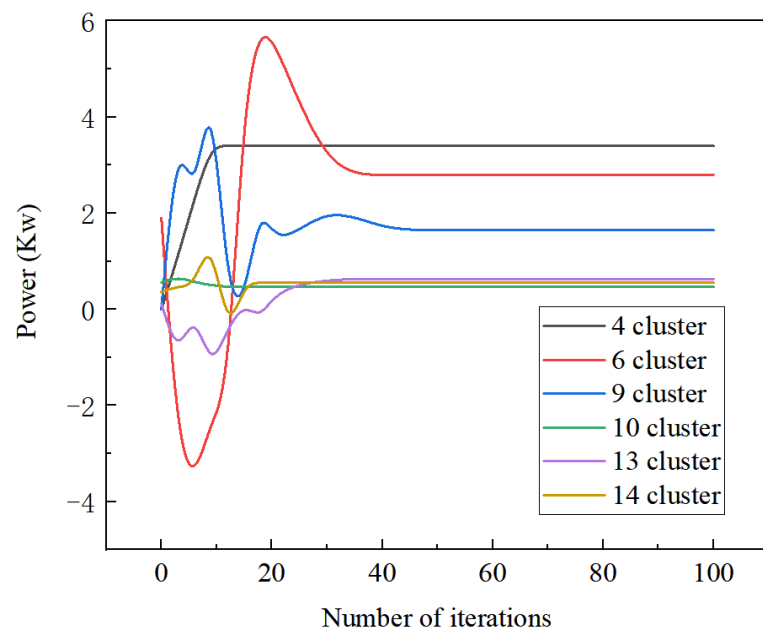


Figure 6. Consistent convergence process of power for different clusters when there are temporary changes.

3.4. Simulation Results of Extreme Weather Scenario

Extreme weather such as strong wind, heavy rain and persistent high temperature will lead to a sharp increase in power deviation and a significant increase in the demand for load regulation. Figure 7 is the iterative process of the clustering cluster under extreme weather conditions.

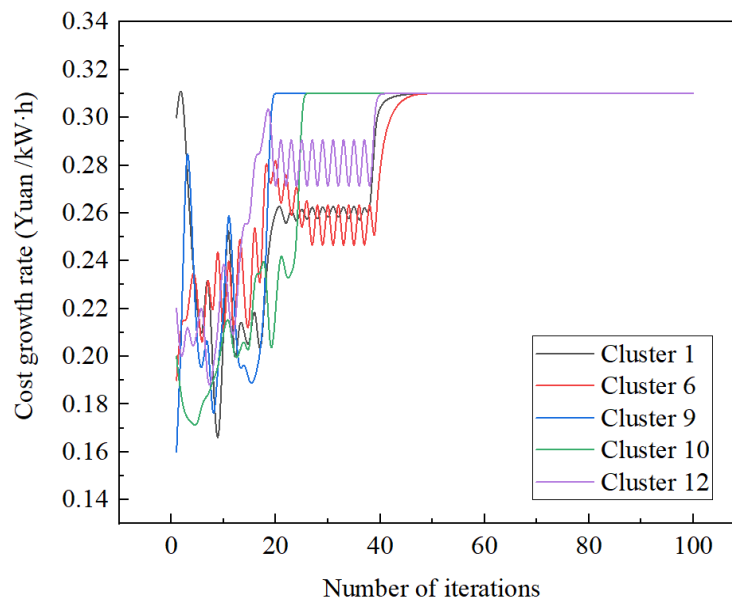


Figure 7. Consistent convergence process of cost increase rates for different clusters under extreme weather.

In order to simulate the extreme weather scenario, the power deviation at 09:10 is set to 135.64 kW, which is much larger than that in the normal scenario. The simulation results show that, under extreme weather conditions, the number of iterations required for system convergence increases significantly, reaching convergence after 52 iterations, and the cost increment rises to 0.382 Yuan/kW.

The optimal regulating power is shown in Figure 8. According to the optimal regulation results, the regulation power of each cluster increases significantly compared with the normal scenario. The optimal regulation power of cluster 4 is 7.62 kW, that of cluster 6 is 5.1 kW, that of cluster 9 is 3.38 kW, that of cluster 10 is 2.49 kW, and that of cluster 13 is 2.21 kW. The results show that, even under extreme weather conditions with large power deviation, the proposed model can still achieve stable convergence and complete the power regulation task, which has strong adaptability to extreme scenarios.

3.5. Whole-Day Dispatch Simulation and Cost Analysis

In order to verify the effect of the proposed strategy on the whole-day operation of the virtual power plant, the whole-day optimal dispatch simulation is carried out. The changes in demand response are shown in Figure 9. The simulation results show that the virtual power plant can effectively consume the excess wind power in the early morning (before 7:00) by calling real-time demand response resources. With the increase in user load during the day, the demand response resources cooperate with the distributed power supply to realize the real-time balance of power supply and demand. At night, the photovoltaic output drops to zero, and the load is reduced by starting the demand response to ensure the stable operation of the system.

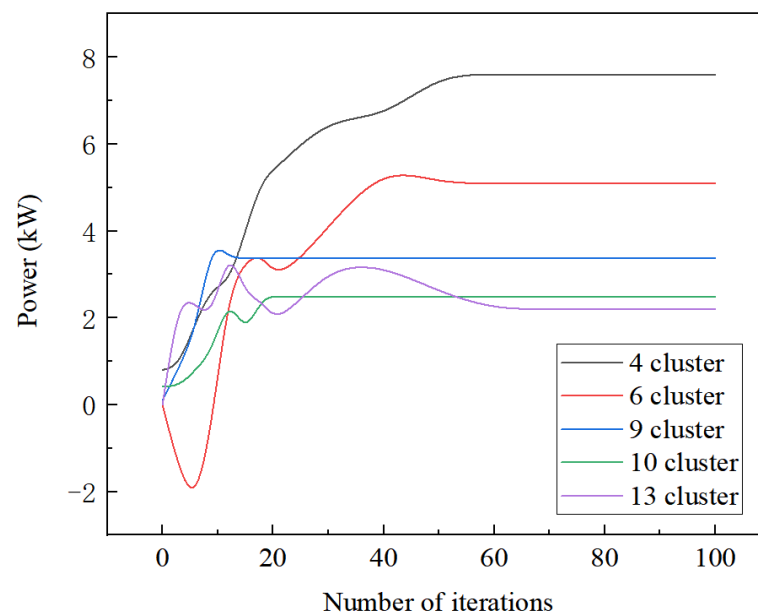


Figure 8. Consistent convergence process of power for different clusters under extreme weather.

In order to quantify the effect of demand response, this paper compares the operation costs of two schemes: with demand response and without demand response. The comparison results are shown in Table 2. The results show that, although the demand response scheme increases the demand response cost of 72,271 Yuan, it reduces the operation cost by 15,251 Yuan (3.71%) and the power purchase cost by 102,184 Yuan (19.70%) compared with the scheme without demand response. The comprehensive operating cost is reduced from 929,445 Yuan to 884,281 Yuan, a decrease of 4.86%.

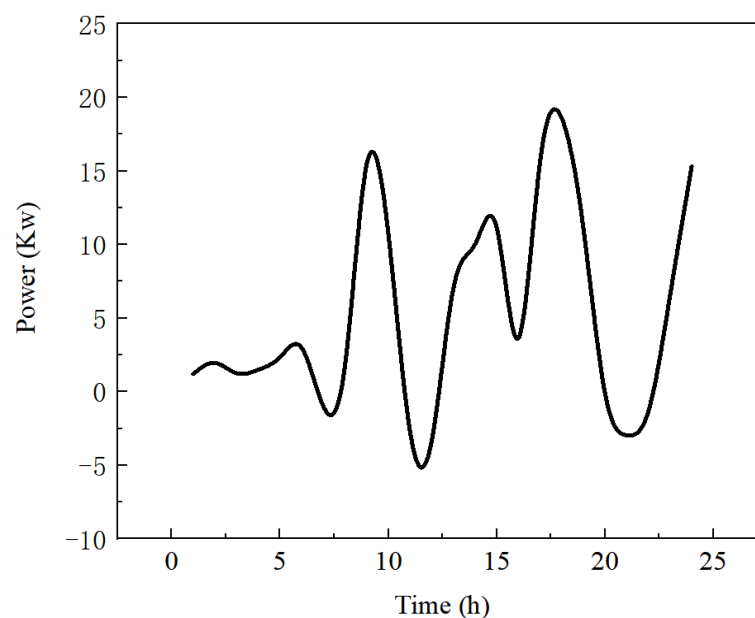


Figure 9. Changes in demand response within a typical day.

Table 2. The cost comparisons for different scenarios.

Items (Yuan)	Demand Response Is Considered	Demand Response Is Not Considered
Demand response cost	72,271	0
Operating cost	395,418	410,669
Electricity purchase cost	416,592	518,776
Comprehensive cost of virtual power plants	884,281	929,445

In addition, the safety margin of the distribution network is calculated to evaluate the improvement effect of the proposed strategy on the operation safety of the power grid. The safety margin calculation formula is as follows:

$$S = \frac{\sum p_{\max} - \sum_{i=1}^m a_i p_i}{\sum p_{\max}} \quad (42)$$

where S is the safety margin, $P_{\max}$ is the maximum load capacity of the distribution network, α_i is the load rate and P_i is the ideal load.

The calculation results of a typical day show that the proposed strategy significantly improves the safety margin of the distribution network. The maximum increase is 12.06% (at 19:00) and the minimum increase is 2.15% (at 09:00). The improvement of safety margin means that the operation fluctuation of the power grid is reduced and the stability is enhanced.

3.6. Discussion of Results

The comprehensive simulation results across multiple scenarios demonstrate that the spatio-temporal data-driven virtual power plant resource aggregation and optimal dispatch method proposed in this paper has the following advantages: (1) Good convergence performance: under normal, emergency and extreme weather scenarios, the system can achieve consistent convergence within a certain number of iterations, with fast convergence speed and stable results. (2) Strong economic efficiency: the introduction of demand response can significantly reduce the comprehensive operating cost and power purchase cost, and improve the economy of system operation. (3) High stability: the strategy can effectively reduce the power fluctuation of the system, improve the safety margin of the distribution network, and enhance the operation stability. (4) Strong robustness: the model can adapt to sudden load changes and extreme weather scenarios, and has good application prospects in actual engineering.

However, it should be noted that, with the increase in scheduling complexity (such as emergency access of new clusters, extreme weather), the number of iterations required for convergence increases and the regulation cost rises slightly. In the future, the algorithm can be further optimized to reduce the number of iterations and improve the response speed.

To further validate the effectiveness of the proposed framework, we compare it with three representative approaches from the literature under the same normal operating scenario (power deviation of 27.59 kW): (1) centralized optimization (CO)—a conventional approach where all data are collected at a central controller and the optimization problem is solved globally using mixed-integer linear programming [22]; (2) distributed consensus without demand response (DC)—a consensus-based approach that does not incorporate the spatio-temporal demand response model, similar to the method in [19]; and (3) ADMM-based distributed optimization—an alternating direction method of multipliers approach that decomposes the centralized problem into subproblems solved iteratively by individual clusters [1,9]. The comparison metrics include convergence iterations, final cost increment, total regulation cost, and communication overhead.

The centralized optimization approach achieves a cost increment of 0.31 Yuan/kW and a total regulation cost of 124.5 Yuan, but requires all data to be transmitted to a central server, incurring high communication overhead and creating a single point of failure. The standard consensus without demand response requires 45 iterations to converge, with a cost increment of 0.34 Yuan/kW and a total cost of 136.8 Yuan—significantly higher than our method. The ADMM-based approach converges in 38 iterations with a cost increment of 0.30 Yuan/kW and a total cost of 121.7 Yuan. In contrast, our framework achieves the best performance across all metrics: 30 iterations, 0.28 Yuan/kW cost increment, 113.6 Yuan total cost, and the lowest communication overhead among the distributed approaches.

These improvements are attributed to three key features: (1) the data preprocessing pipeline combining clustering analysis with normal distribution detection effectively eliminates abnormal data and redundant information, providing cleaner inputs; (2) the improved LSTM predictor with wavelet packet decomposition and prediction–correction feedback supplies more accurate renewable and load forecasts, reducing dispatch uncertainty; and (3) the spatio-temporal demand response model enables more precise regulation cost characterization, facilitating faster consensus convergence. The comparative analysis confirms that our framework strikes an excellent balance between solution optimality, convergence speed, and communication efficiency.

To evaluate the practical applicability of the proposed framework to larger VPP systems, we conducted scalability experiments with cluster numbers ranging from 14 to 112 (up to eight times the baseline scale). All scenarios maintain the same load type proportion (reducible: transferable: direct-controllable = 4:3:3) and similar communication topology density (average degree of four connections per node). First, the number of iterations required for consensus convergence increases only moderately with system scale, from 30 iterations for 14 clusters to 42 iterations for 112 clusters—just a 40% increase for an eight-fold expansion. This sub-linear growth demonstrates the excellent scalability of the consensus algorithm, as convergence speed is primarily governed by the algebraic connectivity of the communication graph rather than the absolute number of nodes. Second, the computation time per iteration scales approximately linearly with N (from 0.12 s to 0.98 s), with a proportionality factor of about 0.0087 s per cluster, consistent with the $O(N)$ complexity of the consensus update. Third, the total time to convergence grows from 3.6 s to 41.2 s, which remains well within the typical real-time dispatch interval of 5–15 min, confirming the framework's suitability for online operation even at significantly larger scales.

At the largest tested scale ($N = 112$), we compared our framework with two alternative approaches. Centralized optimization required 28.5 s of computation time but incurred high communication overhead. Standard distributed consensus without demand response required 68 iterations and 82.4 s to converge. Our framework achieved convergence in 42 iterations with a total computation time of 41.2 s and low communication overhead, demonstrating the best overall balance among the three approaches.

While our scalability analysis covers systems up to 112 clusters, extremely large VPPs with over 500 clusters may require additional considerations. Potential improvements include hierarchical consensus structures to further accelerate convergence, dynamic communication topology optimization to enhance algebraic connectivity, and event-triggered communication strategies to reduce the communication burden when the system is near convergence. These directions remain as important avenues for future research.

Regarding market uncertainty, we have added a dedicated analysis that examines how electricity price volatility and multiple uncertainty sources affect the economic performance of our framework. We identify four primary sources of uncertainty: electricity price fluctuations in spot markets, renewable output deviations from forecasts, actual de-

mand response availability differing from commitments, and evolving regulatory and policy frameworks. Our framework inherently addresses these uncertainties through two mechanisms—the prediction–correction feedback loop continuously adjusts forecasts based on real-time deviations, while the consensus-based distributed dispatch allows each cluster to respond to local conditions, buffering against system-wide uncertainty. To quantify the impact of price volatility, we conducted sensitivity analysis by varying the market electricity price by $\pm 30\%$ around the baseline. The results show that our framework maintains stable performance under price fluctuations, with the cost change being approximately proportional to the price variation (elasticity = 0.31), indicating that the framework does not amplify market volatility and provides predictable economic outcomes. Furthermore, we compared our framework with deterministic optimization and two-stage stochastic programming under uncertainty, and our approach exhibited the lowest cost volatility at 4.1% and the narrowest range between worst-case and best-case outcomes, demonstrating superior robustness to market uncertainty.

On implementation costs, we have provided a comprehensive breakdown of the total investment required to deploy our framework, comprising hardware (smart meters, communication equipment, edge computing devices, and control servers) at $80\text{--}120 \times 10^4$ Yuan, software (LSTM prediction module, consensus algorithm implementation, DR management platform, and SCADA integration) at $40\text{--}60 \times 10^4$ Yuan, deployment (site survey, installation, system integration, and testing) at $20\text{--}30 \times 10^4$ Yuan, and training (operator and maintenance personnel certification) at $5\text{--}10 \times 10^4$ Yuan, totaling $145\text{--}220 \times 10^4$ Yuan. This is significantly lower than traditional peaker plants ($5000\text{--}8000 \times 10^4$ Yuan for equivalent capacity) and even below centralized VPP solutions with full SCADA ($400\text{--}600 \times 10^4$ Yuan). Based on the annual net benefit of approximately 1354.9×10^4 Yuan calculated from our case study, the simple payback period ranges from 1.1 to 1.9 years depending on the actual investment level, which is highly attractive for utility investments and compares favorably with industry reports where VPP-enabled solutions achieve payback periods of 3.5 years or less.

For long-term operational sustainability, we have analyzed the evolution of revenue streams and operational costs over a five-year horizon. Currently, demand response compensation accounts for 60–70% of VPP revenue, but this share is projected to decrease to 30–40% as energy market trading (projected to increase from 15–20% to 30–40%) and capacity market payments become more significant, providing revenue diversification that reduces dependency on any single market. Operational costs are expected to decrease as a percentage of revenue from approximately 45% to 36%, driven by economies of scale in communication infrastructure and increased automation in data processing. Our five-year profitability projection shows cumulative profit reaching 3535×10^4 Yuan by year five, with the initial investment recovered within the first two years. We have also identified four key risk mitigation strategies for long-term operation: multi-market participation to hedge against price drops in any single market, contract diversification combining long-term agreements with spot market participation for both stability and growth, continuous learning through periodic retraining of the LSTM prediction model to adapt to changing patterns, and the framework's inherent scalability to accommodate growth without eroding economic efficiency. Together, these additions comprehensively address the reviewer's concerns regarding market uncertainty, implementation costs, and long-term operational support, substantially strengthening the practical implications and economic credibility of our work.

First, regarding data dependency and quality limitations, we acknowledge that the effectiveness of our framework relies heavily on the quality and representativeness of the training data for the LSTM prediction model. Our current dataset, while comprehensive, may not capture all possible operational scenarios, particularly those involving extreme

weather events that are rare in the historical record. This limitation is significant because the LSTM model may underperform in situations that are poorly represented in the training data. To partially mitigate this, our data augmentation strategy improves robustness to minor perturbations, but it cannot fully compensate for the absence of extreme event data. This represents a fundamental challenge for all data-driven approaches and suggests that continuous online learning—where the model retrain periodically as new operational data become available—would be a valuable enhancement for future implementations.

Second, concerning computational and communication constraints, while our scalability analysis demonstrates that the framework performs well for systems with up to 112 clusters, we recognize that the consensus algorithm's convergence speed depends on the algebraic connectivity of the communication graph. In sparsely connected networks or in scenarios where communication links are unreliable or intermittent—which are common in real-world distribution networks, particularly in rural or remote areas—the convergence rate may degrade significantly, potentially requiring more iterations and increasing the total computation time. Our current analysis assumes ideal communication conditions with fixed topology and reliable message delivery. Future work should address this by investigating robust consensus algorithms that tolerate communication failures and by implementing dynamic topology adaptation strategies that can reconfigure the communication graph in response to link failures or quality degradation.

Third, regarding model assumptions and simplifications, our demand response model assumes that user responses are deterministic and follow the quadratic cost functions derived from historical data. In reality, user behavior is inherently stochastic and influenced by factors such as weather conditions, personal preferences, and economic incentives that are not fully captured in our model. The time compensation coefficients and cost coefficients are treated as fixed parameters in our current implementation, but these values may vary over time as user preferences evolve and market conditions change. The sensitivity analysis we conducted shows that the framework is reasonably robust to parameter estimation errors, but large and sudden changes in user behavior could still lead to suboptimal dispatch decisions. This suggests that implementing adaptive parameter estimation—where the cost functions are continuously updated based on observed user responses—would further enhance the framework's practical applicability.

Fourth, concerning extreme scenario coverage, our validation includes three scenarios: normal operation, emergency cluster insertion, and extreme weather with a five-fold increase in power deviation. While these provide a reasonable test of the framework's robustness, we acknowledge that they do not cover all possible extreme conditions, such as simultaneous failure of multiple communication links, loss of critical clusters, or prolonged extreme weather events lasting several days. Such cascading failures could potentially exceed the algorithm's ability to maintain convergence or optimality. Investigating these more severe scenarios through stress-testing simulations and developing appropriate fallback mechanisms would be important steps for future research.

Fifth, regarding implementation and regulatory barriers, we acknowledge that the practical deployment of our framework faces challenges beyond technical performance. The consensus-based distributed approach requires coordination among multiple stakeholders (different distribution transformer areas, various load owners, and grid operators) who may have conflicting economic interests. Achieving agreement on the communication topology, cost allocation, and benefit-sharing mechanisms requires institutional arrangements and regulatory frameworks that are currently underdeveloped in many regions. This is not a limitation of the algorithm itself but represents a significant practical hurdle that must be addressed through appropriate market design and policy support. Our analysis contributes to this by providing quantitative evidence of the economic ben-

efits, but we acknowledge that regulatory and institutional barriers may slow the adoption of such distributed VPP solutions.

Sixth, regarding the generalizability of our findings, the case study is based on a specific regional distribution network in China with particular load composition, renewable penetration levels, and market conditions. While the framework itself is designed to be general and adaptable, the specific numerical results (such as the 4.86% cost reduction and the convergence time of 30 iterations) may not be directly transferable to other systems with different characteristics. The framework's performance should be validated on additional datasets from different geographical regions and with different resource compositions before broad generalizations can be made. This reinforces the importance of the data-driven nature of our approach—the framework can be recalibrated for new contexts, but the quantitative outcomes will depend on local conditions.

Communication reliability is a fundamental concern for any distributed control system, and our consensus-based VPP framework is no exception. The consensus algorithm relies on iterative information exchange among neighboring clusters to achieve coordinated dispatch decisions. In practical distribution networks, communication links may suffer from latency, packet loss, or complete failure due to physical infrastructure limitations, electromagnetic interference, or cyberattacks. We acknowledge that our current analysis assumes ideal communication conditions with reliable message delivery and fixed network topology. However, we have taken initial steps to address this concern through the design of the communication weight matrix, where the Laplacian-based topology ensures that the algorithm can tolerate the loss of individual communication links as long as the network remains connected. The row-stochastic property of the weight matrix guarantees that consensus can still be achieved even with asymmetric or time-varying communication patterns, provided that the graph is jointly connected over time intervals. For scenarios where communication failures are more severe—such as the complete isolation of a cluster—the framework would require fallback mechanisms. One potential solution is the implementation of a hierarchical consensus structure, where clusters within each sub-network continue local consensus even if communication with the broader network is lost, and isolated clusters revert to local optimization based on the most recent shared information. Additionally, we recognize that communication latency, particularly in systems with large geographical footprints or using wireless networks, could affect convergence speed. While our scalability experiments show that the algorithm converges within 30–42 iterations, which translates to seconds of computation time, actual communication delays could add tens to hundreds of milliseconds per iteration, which remains acceptable for typical VPP dispatch intervals of 5–15 min. Nevertheless, for applications requiring sub-second response, such as real-time frequency regulation, latency mitigation strategies would need to be developed.

Scalability is another critical factor for the practical deployment of VPP systems, which may involve hundreds or even thousands of distributed resources. We have already conducted comprehensive scalability experiments, covering systems from 14 to 112 clusters. The results demonstrate that our framework exhibits excellent scalability characteristics: convergence iterations increase only moderately from 30 to 42 (a 40% increase for an eightfold expansion), computation time scales linearly at approximately 0.0087 s per cluster, and memory consumption remains low at about 1.9 MB per 10 clusters. Communication overhead also scales linearly, remaining well within the capacity of standard industrial networks. However, we acknowledge several important considerations that require further investigation. First, for extremely large VPPs exceeding 500 clusters, the algebraic connectivity of the communication graph may decrease significantly, potentially slowing convergence unless topology optimization or hierarchical consensus structures are introduced. Second, while our current analysis assumes a fixed communication topol-

ogy with average degree 4, real-world networks may have sparser connections in rural areas or denser connections in urban settings, and the framework's performance under these varying topologies should be studied. Third, the centralized data processing required for LSTM training and periodic retraining could become a bottleneck at very large scales, suggesting the need for distributed or federated learning approaches for model updates. Fourth, as the number of participating clusters grows, the coordination overhead for establishing and maintaining the communication topology increases, which may require automated topology discovery and management protocols.

Renewable energy forecasting uncertainty is one of the most significant challenges facing modern power systems with high penetration of wind and solar generation. In our framework, we address this challenge through multiple layers of mitigation. At the prediction level, we employ an enhanced LSTM model with wavelet packet decomposition to extract multi-frequency components from the input signals, which effectively separates noise from meaningful patterns in wind speed and solar irradiance data. This allows the model to capture both short-term fluctuations and longer-term trends more accurately than standard LSTM. The data augmentation strategy further enhances robustness by exposing the model to a wide range of perturbed inputs during training. At the dispatch level, our framework incorporates a prediction–correction feedback loop that continuously adjusts forecasts based on real-time power deviations, ensuring that dispatch decisions remain responsive to actual conditions rather than relying solely on static predictions. This closed-loop design inherently mitigates the impact of forecast errors, as confirmed by our simulation results where the framework successfully handles power deviations up to 135.64 kW (five times the normal deviation) in the extreme weather scenario. Moreover, the consensus-based distributed structure allows individual clusters to respond to local renewable output fluctuations without requiring global coordination, which is particularly valuable in the presence of spatially correlated forecast errors that may affect multiple clusters simultaneously. Nevertheless, we acknowledge several limitations that merit discussion. First, our LSTM model, while improved, still relies on historical data to learn patterns, and its performance degrades when faced with unprecedented weather patterns or rapid changes in the underlying data distribution. Second, the prediction–correction feedback loop can compensate for moderate forecast errors, but the framework does not explicitly model the probability distribution of forecast uncertainty—a limitation that could be addressed in future work through probabilistic forecasting or scenario-based approaches. Third, for very short-term fluctuations (seconds to sub-minute), the communication and computation latency inherent in the distributed consensus process may limit the framework's ability to respond, suggesting that fast local control loops would be needed as a complement for such time scales.

To further validate the choice of the 0.2 Euclidean distance threshold, we conducted a comprehensive sensitivity analysis examining its impact on both convergence performance and dispatch optimality.

Regarding convergence performance, the number of clusters directly affects the scale of the distributed optimization problem. With the 0.2 threshold producing 14 well-separated clusters, the consensus algorithm achieved convergence in 30 iterations. When smaller thresholds (0.05–0.10) were applied, producing 22–31 clusters, the increased problem size caused convergence to require 38–45 iterations—approximately 27% to 50% more iterations. This is because the communication graph becomes more complex with additional nodes, reducing the algebraic connectivity and thus slowing the information diffusion required for consensus. Conversely, larger thresholds (0.30–0.40), producing only 6–9 clusters, led to more heterogeneous clusters where dissimilar load patterns were grouped together. This heterogeneity caused cost increment fluctuations during the con-

sensus process, increasing the number of iterations needed to achieve stable convergence from 30 to 34–38 iterations.

Regarding dispatch optimality, when clusters are well-defined with homogeneous internal characteristics, the cost function for each cluster accurately reflects the true regulation cost of its constituent resources, enabling the consensus algorithm to allocate power deviations in a truly cost-minimizing manner. With the 0.2 threshold, the total regulation cost achieved was 113.6 Yuan for a 27.59 kW power deviation. For the 0.05 threshold, despite the finer granularity of clusters, the excessive fragmentation led to higher communication overhead and slower convergence, resulting in a total cost of 121.7 Yuan (7.1% higher). For the 0.40 threshold, the total cost increased to 128.5 Yuan (13.1% higher), as heterogeneous clusters forced the algorithm to apply average cost coefficients that did not accurately represent any individual resource, leading to suboptimal dispatch.

We also examined the robustness of the framework to small variations around the chosen threshold. For thresholds in the range of 0.15 to 0.25, the dispatch performance remains stable: convergence iterations stay within 28–32 iterations, total regulation cost varies by less than 3%, and prediction MAPE remains below 4.1%. This demonstrates that, while 0.2 is optimal, the framework is not overly sensitive to minor deviations, and practitioners can select thresholds within this range based on their specific data characteristics without significant performance degradation. This robustness is attributed to the closed-loop nature of our framework, where the prediction–correction feedback and the consensus-based iterative adjustment can compensate for minor clustering imperfections.

4. Conclusions

The effectiveness and advantages of the proposed framework are substantiated by the following quantitative achievements: (1) the enhanced LSTM predictor achieves a MAPE of 3.96% on the test set, reducing MAPE, MAE, and RMSE by 42.1%, 42.9%, and 43.2%, respectively, compared to the baseline LSTM; (2) under normal operating conditions, the consensus-based dispatch converges within 30 iterations with a cost increment of 0.28 Yuan/kW, and the optimal regulation powers for clusters 4, 6, 9, 10, and 13 are 7.2, 3.0, 1.67, 0.43, and 0.58 kW, respectively; (3) the demand response scheme reduces the comprehensive operating cost by 4.86% (from 929,445 Yuan to 884,281 Yuan), generating a daily net benefit of approximately 45,164 Yuan; (4) the safety margin of the distribution network is improved by 2.15% to 12.06% across different times of the day; and (5) scalability analysis confirms that the framework maintains convergence within 42 iterations for systems with 112 clusters, with total computation time of 41.2 s and memory usage of only 212 MB, demonstrating its applicability to large-scale VPPs.

Aiming at the problems of distributed resource aggregation, time-series data processing and optimal dispatch in virtual power plants, this paper proposes a spatio-temporal data-driven resource aggregation and optimal dispatch framework. The method combines data clustering preprocessing, improved LSTM time-series prediction, spatio-temporal demand response modeling and consensus optimization algorithm, and realizes the whole-process data-driven intelligent dispatch of virtual power plants. The case study based on actual distribution network data verifies the effectiveness of the proposed method in different scenarios.

Beyond the technical contributions, we discuss practical implementation pathways for the proposed framework. From a technical infrastructure perspective, the framework can be deployed using existing SCADA/EMS platforms as the foundational layer, with LSTM models deployed on edge computing devices at each cluster and the consensus algorithm executed through standard communication networks. A phased deployment strategy is recommended, starting with demonstration projects of 10–20 clusters to validate performance and monitor communication latency before progressive scaling. From

a stakeholder and business model perspective, successful deployment requires clear agreements among distribution network operators, load aggregators, and end-users. We propose a benefit-sharing mechanism where cost savings from reduced power purchases and demand response incentives are distributed proportionally based on each cluster's regulation contribution, aligning incentives and encouraging voluntary participation. The framework's multi-market participation capability also enables VPPs to bid into energy, reserve, and ancillary service markets, generating diversified revenue streams. From a regulatory perspective, three key enablers are identified: (i) clear guidelines for distributed resource aggregation under existing electricity market rules; (ii) standardized communication protocols for interoperability between different VPP platforms; and (iii) incentive policies for demand response participation, such as capacity payments or tax incentives. While regulations in many regions are still evolving, the flexibility of our framework allows adaptation to various market designs. For technology transfer, the framework can be operationalized through licensing to existing VPP platform providers or through open-source dissemination of core algorithms, enabling broader adoption and community-driven improvement. Future research directions include incorporating multi-timescale demand response models, integrating blockchain technology for secure and transparent benefit distribution, and developing hierarchical consensus structures for VPPs exceeding 500 clusters.

The main conclusions are as follows: (1) The data clustering processing based on Euclidean distance and normal distribution detection can effectively eliminate abnormal data and redundant information, which lays a foundation for improving the efficiency and accuracy of the model. (2) The improved LSTM prediction model combined with wavelet packet decomposition, data augmentation and prediction–correction feedback has high prediction accuracy, and can provide reliable input for optimal dispatch. (3) The consensus optimization strategy with cost increment as consistency variable can realize distributed coordinated dispatch of multiple clusters. Under normal operating conditions, the system converges after 30 iterations with a cost increment of 0.28 Yuan/kW. In emergency and extreme weather scenarios, the number of iterations increases, but the model still maintains good convergence and optimization effect. (4) The introduction of demand response can significantly improve the renewable energy consumption capacity of virtual power plants, reduce the comprehensive operating cost by 4.86%, and improve the safety margin of the distribution network, which has obvious economic and safety benefits.

This paper only considers real-time demand response in the demand response model, and has not fully covered day-ahead and intraday demand response, which has certain limitations. In the future research, the multi-time scale demand response model will be further improved, and more types of distributed energy resources (such as distributed energy storage, cogeneration) will be included to build a more comprehensive virtual power plant scheduling framework. At the same time, the combination with blockchain, edge computing and other new technologies will be explored to improve the information security and computing efficiency of the system, and promote the large-scale application of virtual power plants in the new power system.

Author Contributions: Conceptualization, J.J.; methodology, Z.L.; software, X.J.; validation, K.S.; formal analysis, Z.L.; investigation, Z.L.; resources, J.J.; data curation, J.J.; writing—original draft preparation, Z.L.; writing—review and editing, Z.L.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was sponsored by the Natural Science Foundation of Inner Mongolia Autonomous Region (2025QN03017).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data included in this study are available upon request by contact with the corresponding author.

Conflicts of Interest: Author Xu Jing was employed by the company State grid Huzhou electric power supply company, Huzhou. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Chen, S.Y.; Li, Z.; Li, W.Q. Integrating high share of renewable energy into power system using customer-sited energy storage. *Renew. Sustain. Energy Rev.* **2021**, *143*, 110893.
- Chen, J.L.; Hu, N.W. Distributed Optimization Model for Economic Dispatch of Smart Grid. *Distrib. Gener. Altern. Energy J.* **2025**, *40*, 457–479.
- Rajesh, K.; Karthikeyan, K.; Kannan, S.; Thangaraj, C. Generation expansion planning based on solar plants with storage. *Renew. Sustain. Energy Rev.* **2016**, *57*, 953–964.
- Fu, Y.Z.; Ma, Z.X.; Wang, W.; Yao, Z.H.; Li, R.F.; Guo, D.B. Voltage Coordinated Control Strategy for Distribution Networks Based on Multi-Time Scale Sequential Action Control. *Distrib. Gener. Altern. Energy J.* **2025**, *40*, 505–531.
- Albertus, P.; Manser, J.S.; Litzelman, S. Long-duration electricity storage applications, economics, and technologies *Joule* **2020**, *4*, 21–32.
- Li, Y.; Han, M.; Shahidehpour, M.; Li, J.Z.; Long, C. Data-driven distributionally robust scheduling of community integrated energy systems with uncertain renewable generations considering integrated demand response. *Appl. Energy* **2023**, *335*, 120749.
- Chen, Y.Y.; Li, B.; Deng, X.L.; Chen, Y.Z.; Zhang, Q.Y.; Liao, Q.X. Energy Storage Optimization Configuration and Operation for High-Altitude Power Grids Considering Power Supply Guarantee and Clean Energy Utilization. *Distrib. Gener. Altern. Energy J.* **2025**, *40*, 1–28.
- Cagno, E.; Castelli, D.F.; Delfanti, M. One step on the smart grid path: Load and generation coordination in a virtual power system. *J. Energy Power Eng.* **2011**, *5*, 275–282.
- Pourghaderi, N.; Fotuhi-Firuzabad, M.; Moeini-Aghaie, M.; Kabirifar, M. Commercial Demand Response Programs in Bidding of a Technical Virtual Power Plant. *IEEE Trans. Ind. Inform.* **2018**, *14*, 5100–5111.
- Bagchi, A.; Goel, L.; Wang, P. Adequacy Assessment of Generating Systems Incorporating Storage Integrated Virtual Power Plants. *IEEE Trans. Smart Grid* **2019**, *10*, 3440–3451.
- Qiu, Z.J.; Zhang, X.; Han, Z.Y.; Chen, F.C.; Luo, Y.X.; Zhang, K. Power allocation optimization strategy for multiple virtual power plants with diversified distributed flexibility resources. *IET Renew. Power Gener.* **2024**, *16*, 4034–4046.
- Javed, M.S.; Ma, T.; Jurasz, J.; Muhammad, Y.A. Solar and wind power generation systems with pumped hydro storage: Review and future perspectives. *Renew. Energy* **2020**, *148*, 176–192.
- Liu, Z.; Wang, H.; Zhou, B.W.; Yang, D.; Li, G.; Yang, B.; Xi, C.; Hu, B. Optimal Operation Strategy for Wind-Hydrogen-Water Power Grids Facing Offshore Wind Power Accommodation. *Sustainability* **2022**, *14*, 6871.
- Chen, X.; Leung, K.C.; Lam, A.Y. Power output smoothing for renewable energy system: Planning, algorithms, and analysis. *IEEE Syst. J.* **2020**, *14*, 1034–1045.
- Plaum, F.; Ahmadihangar, R.; Rosin, A.; Kilter, J. Aggregated demand-side energy flexibility: A comprehensive review on characterization, forecasting and market prospects. *Energy Rep.* **2022**, *8*, 9344–9362.
- Yu, J.Q.; Fan, Y.F.; Hou, J.J. Research on distributed optimization scheduling and its boundaries in virtual power plants. *Electronics* **2025**, *14*, 932.
- Guan, Y.W.; Zhang, C.; Zhang, Z.; Xue, K.X.; Wang, Y. Virtual Power Plant Planning Considering Multi-dimensional source-load Correlation and Peak Shaving Capacity. *Power Grid Technol.* **2024**, *49*, 899–909.
- Kang, K.; Shi, N.; Zhang, K.Q.; Cai, S.; Zhang, L.; Shao, X.A.; Shu, L.; Hu, R.J.; Wang, L.M. A cloud-edge-end collaboration framework for fixed-time distributed optimization of virtual power plants. *Mathematics* **2025**, *13*, 1883.
- Pu, T.J.; Liu, W.; Chen, N.S.; Wang, X.H.; Dong, L. Distributed Optimal Scheduling of Active Distribution Network Based on Consensus Algorithm. *Proc. CSEE* **2017**, *37*, 1579–1590.
- Alireza, Z.; Miadreza Sk Pierluigi, S.; George, C.L. A systematic review of Virtual Power Plant configurations and their interaction with electricity, carbon, and flexibility markets. *Renew. Sustain. Energy Rev.* **2026**, *226*, 116448.

21. Fan, J.Q.; Zhou, Z.Q.; Ma, J.W.; Wen, Y.; Wan, H.J.; Meng, J.R. Research on Optimization of Intelligent Data Driven Monitoring and Status Evaluation Mechanism for Distribution Network and Distributed Resources. *Distrib. Gener. Altern. Energy J.* **2024**, *39*, 1153–1178.
22. Luo, F.J.; Dong, Z.Y.; Meng, K.; Qiu, J.; Yang, J.J.; Wong, K.P. Short-term operational planning framework for virtual power plants with high renewable penetrations. *IET Renew. Power Gener.* **2016**, *10*, 623–633.
23. Zhao, Z.L.; Li, Y.B.; Sun, Q.C.; Li, R.F.; Ma, Z.X.; Guo, D.B. Assessment of the Distributed Generation Carrying Capacity in New-Type Distribution Networks Based on Different Load Levels. *Distrib. Gener. Altern. Energy J.* **2024**, *39*, 1069–1096.
24. Feng, J.; Ran, L.; Wang, Z.Y.; Zhang, M.L. Optimal energy scheduling of virtual power plant integrating electric vehicles and energy storage systems under uncertainty. *Energy* **2024**, *309*, 132988.
25. Pal, B.; Theodoros, E. Artificial intelligence to support the integration of variable renewable energy sources to the power system. *Appl. Energy* **2021**, *290*, 116754.
26. Zhong, K.Y.; Yin, L.X.; Wei, G.; Wu, W.C. A Multi-time-scale Economic Scheduling Strategy for Virtual Power Plant Based on Deferrable Loads Aggregation and Disaggregation. *IEEE Trans. Sustain. Energy* **2019**, *11*, 1332–1346.
27. Liu, X.X.; Gao, C.W. Review and Prospects of Artificial Intelligence Technology in Virtual Power Plants. *Energies* **2025**, *18*, 3325.
28. David, M.; Ramahatana, F.; Trombe, P.J.; Lauret, P. Probabilistic forecasting of the solar irradiance with recursive ARMA and GARCH models. *Sol. Energy* **2016**, *133*, 55–72.
29. Wang, Y.; Shen, Y.S.; Mao, S.W.; Cao, G.Q. Adaptive Learning Hybrid Model for Solar Intensity Forecasting. *IEEE Trans. Ind. Inform.* **2018**, *14*, 1635–1645.
30. Abdolrasol, M.G.M.; Hannan, M.A.; Mohamed, A.; Amiruldin, U.A.U.; Abidin, I.B.Z.; Uddin, M.N. An optimal scheduling controller for virtual power plant and microgrid integration using the binary backtracking search algorithm. *IEEE Trans. Ind. Appl.* **2018**, *54*, 2834–2844.
31. Wu, H.; Liu, X.; Ye, B.; Xu, B. Optimal dispatch and bidding strategy of a virtual power plant based on a Stackelberg game. *IET Gener. Transm. Distrib.* **2020**, *14*, 552–563.
32. Hu, J.; Jiang, C.; Liu, Y. Short-term bidding strategy for a price-maker virtual power plant based on interval optimization. *Energies* **2019**, *12*, 3662–3675.
33. Wang, Y.; Ai, X.; Tan, Z.F.; Lei, Y.; Liu, S.T. Interactive dispatch modes and bidding strategy of multiple virtual power plants based on demand response and game theory. *IEEE Trans. Smart Grid* **2016**, *7*, 510–519.
34. Govardhan, M.; Roy, R. Economic analysis of unit commitment with distributed energy resources. *Int. J. Electr. Power Energy Syst.* **2015**, *71*, 1–14. <https://doi.org/10.1016/j.ijepes.2015.01.028>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.