frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 07 May 2024
DOI 10.3389/fenrg.2024.1390073

OPEN ACCESS

EDITED BY
Minglei Bao,
Zhejiang University, China

REVIEWED BY
Wenliang Yin,
The University of Sydney, Australia
Tianbai Deng,
Anhui University, China

*CORRESPONDENCE
Song Yang,
yangsong19971@163.com

RECEIVED 22 February 2024
ACCEPTED 08 April 2024
PUBLISHED 07 May 2024

CITATION
Yang S, Wang C, Sun S, Cheng Y and Yu P (2024), Cluster partition-based two-layer expansion planning of grid–resource–storage for distribution networks.
Front. Energy Res. 12:1390073.
doi: 10.3389/fenrg.2024.1390073

Cluster partition-based two-layer expansion planning of grid–resource–storage for distribution networks

Song Yang*, Chenglong Wang, Shumin Sun, Yan Cheng and Peng Yu

State Grid Shandong Electric Power Research Institute, Jinan, China

In order to realize the optimal planning of grid–resource–storage for distribution networks (DNs) with high penetrated distributed photovoltaics (PVs), a cluster partition-based two-layer expansion planning for DNs is proposed. First, a comprehensive cluster partition index-based cluster partition method is proposed, which involves the indexes such as electrical distance, power balance of the cluster, and cluster size. Second, a cluster partition-based two-layer expansion planning model is proposed. In the upper layer, a line planning model for clusters is established to carry out the planning of cluster connection lines. In the lower layer, a robust source-storage planning model is established with the uncertainty of PVs and loads, and then, the optimal location and capacity of PVs and energy storages (ESs) can be obtained. In addition, the uncertainty regulation parameter is utilized to control the range of uncertainty sets, which can reduce the conservatism of the optimization. Finally, the proposed method is carried out in a real DN in China, which can effectively improve the economy of DN planning.

KEYWORDS

cluster partition, distribution network, expansion planning, distributed photovoltaics, energy storages

1 Introduction

With the rapid growth of energy demand, photovoltaics (PVs) are developing rapidly in China. The large amount of distributed PVs has significantly changed the power flow of the distribution network (DN) (Li Z. et al., 2022), which poses new challenges to DN planning and operation (Li et al., 2022b). How to carry out optimal DN planning is the key to realize the economic operation of the DNs with large-scale distributed PVs.

Currently, the main idea for planning the DNs with distributed PVs is to build a centralized planning model (Liu et al., 2021) by strengthening or extending the lines of the DNs (Wu H. et al., 2022), which takes into account the investment of PVs (Koutsoukis et al., 2018) and the operating costs (Shen et al., 2018). The centralized planning model is suitable for the DN planning when the proportion of distributed PVs is low. However, when large-scale distributed PV is connected to the DNs, the dimensionality of the variables in the centralized planning model increases significantly (Wu L. et al., 2022), and the planning model becomes too complex to be solved (Zhang et al., 2021). To solve the challenges of centralized optimization, a cluster partition-based planning method provides a new way for DN planning. The cluster partition-based planning method can not only decompose the

centralized optimization problem into simple sub-problems of cluster optimization but also maximize the degree of power matching between PVs and the load within the cluster during the planning process, which can greatly increase the PV consumption (Hu et al., 2023).

Cluster partition-based DN planning mainly includes two aspects of cluster partition and cluster planning. In terms of cluster partition, existing research mainly establishes cluster partition indexes based on the grid structure (Xiao et al., 2017) and the power balance in the clusters (Kong et al., 2022). Cluster partition is optimized by particle swarm algorithms (Li and Yang, 2022), clustering algorithms (Wang et al., 2021), and community detection algorithms (Yang et al., 2017). By improving the community algorithm, the division of reactive and active clusters considering the power balance and node coupling degree is realized by Ge et al. (2024). The gray clustering method based on the improved whitening weight function is used to partition the distribution network by Xu L. et al. (2021), and the index weight is obtained by comprehensively applying the analytic hierarchy process and the entropy weight method. The modular index based on the electrical distance and the active power balance index are used as comprehensive division indexes by Li et al. (2022c), and the distributed photovoltaic generation in the distribution network is divided into clusters by using genetic algorithms. Based on the theory of the modularity function model in complex networks, a voltage coordination control method of partitioning the aggregated domain of reactive voltage sensitivity weights and active network loss-voltage sensitivity weights of power systems is proposed by Wang Z. et al. (2023). The nodes with a strong coupling relationship are merged to determine the initial number of partitioning by Ji et al. (2023), and then, the final partitioning result is determined according to the affiliation between each load node and each reactive power source. Combining the K-means clustering algorithm and optimized PSO algorithm for voltage regulation within the cluster ensures that the voltage crossing problem is solved by Su et al. (2023). A cluster partition index system considering the structural and functional properties is proposed by Pan et al. (2021), and the modular index that takes into account the characteristics of electric and heat networks is used on the structural property to describe the connection strength between different network nodes. However, the existing cluster partition indexes only concern the active power balance in each cluster, ignoring the impact of reactive power. Meanwhile, the existing cluster partition indexes ignore the influence of cluster size on the planning results, which can easily lead to large differences in the cluster size, even leading to isolated nodes (Li et al., 2023). In addition, the existing cluster partition methods have insufficient computational accuracy, and for a complex cluster partitioning index, the optimization results tend to fall into local optimum solutions.

The current active distribution network (ADN) planning strategy usually includes the reinforcement or expansion of distribution networks and DG integration under the active management of DG outputs (Mukherjee and Sossan, 2023). A two-level robust optimal feeder routing model for the planning of radial distribution networks is proposed by Zdraveski et al. (2023), where power demand is uncertain. The robust model is solved by implementing the column and constraint generation strategy. A method based on calculating the probability of electric vehicles (EVs) entering each parking lot is proposed by Haji-Aghajani et al. (2023) for the long-term planning of EV parking lots. An integrated power and gas systems of IPGS considering cascading effects for enhancing resilience is proposed by Wang Y. et al. (2023), and the two-phase framework containing phases of "demand reachability evaluation" and "integrated planning" is proposed. A framework for the optimal planning of battery swapping stations (BSSs) in centralized charging mode is proposed by Shaker et al. (2023), and in this mode, the batteries are charged at a central charging station. Possible equipment measures are classified into several categories by Sasaki et al. (2023), formulating the "low-voltage system configuration determination problem;" in addition, a solution algorithm based on the practical priorities of classified measures is proposed. The resilience-oriented distribution network planning problem utilizing a novel three-stage hybrid framework is proposed by Faramarzi et al. (2023), and the decision-making on the line hardening and DG placement is carried out in the first stage. In the second stage, emergency and normal operation optimization is conducted. A collaborative stochastic expansion planning model of a cyber–physical system with resilience constraints is proposed by Zhang et al. (2023), and the model can reduce the coupling risk and enhance the resilience under extreme scenarios. An appropriate probabilistic wind power capacity expansion planning method for a bundled wind–thermal generation system with retrofitted coal-fired units is reformulated as a mixed-integer second-order cone programming problem by Lei et al. (2023). In terms of cluster planning, the existing research mainly considers deterministic scenarios as the research background (Bi et al., 2019), ignoring the impact of source-load uncertainty (Cai et al., 2022). Aiming at solving the problems of resource waste caused by the large-scale access of distributed generators to distribution networks and improving the economy of energy storage systems, a cluster energy-storage control strategy for prompting the distributed generation accommodation and improving the economy of energy storage systems is proposed by Li et al. (2021). In considering the optimization of load distribution among units and introducing consumption costs, a grid evaluation index system including the coordination index of the power transmission and distribution network is constructed by Xu X. et al. (2021). A novel cluster-based distributed generation planning approach is proposed by Ding et al. (2019), and the distribution network is divided into several partitions considering the system network structure and the load characteristics, thus conducting a hierarchical and partitioned network structure. A planning model of renewable energy access is established by Hu et al. (2020) based on cluster partition considering the investments and power generation interests of power producers and the power match degree within clusters. With the increasing proportion of distributed PVs, the source-load uncertainty increases the difficulty in modeling the uncertainty of DNs (Liu et al., 2022) and increases the dimensionality of variables in the planning model (Jiang et al., 2022). How to establish a cluster planning model based on the source-load uncertainty (Zhu et al., 2018) and simplify the traditional centralized planning model need to be further researched.

Based on the above analysis, this paper proposes a cluster partition-based two-layer expansion planning model of

grid–resource–storage for DNs. The main contributions of this paper are summarized as follows:

(1) To deal with poor power balance and unbalanced cluster size in existing cluster partition, a comprehensive cluster partition index is proposed, which includes the modularity index, power balance index, and nodal size index. In addition, based on the comprehensive cluster partition index, an improved genetic algorithm is proposed to partition the DN into some clusters.
(2) To deal with complex models in centralized planning methods, a cluster partition-based two-layer expansion planning model is established for the DNs. In the upper layer, a line planning model is established to carry out the planning of cluster connection lines. In the lower layer, the PV and ES planning model within a cluster is established, which can realize the optimal planning of PVs and ESs in each cluster.
(3) To reduce the conservatism of the traditional robust optimization, a box uncertainty set is utilized to characterize the uncertainty of loads and PVs, and an uncertainty regulation parameter is used to control the range of uncertainty sets, which can reduce the conservatism of the optimization and simplify the calculation process.

The remainder of this paper is organized as follows: a comprehensive cluster partition index-based cluster partition method is proposed in Section 2; a cluster partition-based two-layer expansion planning method is proposed in Section 3; in Section 4, the case study is analyzed; and the conclusion is given in Section 5.

2 Comprehensive cluster partition index-based cluster partition method

2.1 Comprehensive cluster partition index

As the existing cluster partition index is not comprehensive, based on the DN structure and cluster function, a comprehensive cluster partition index is proposed to complete the cluster partition in this paper. The proposed comprehensive cluster partition index includes the modularity index, power balance index, and nodal size index.

2.1.1 Modularity index

The coupling degree between nodes can be measured by a modularity index based on voltage sensitivity, which is expressed as

$$\rho = \frac{1}{2\Omega}\sum_{i\in S}\sum_{j\in S}\left(v_{ij} - \frac{\kappa_i\kappa_j}{2\Omega}\right)\phi(i,j), \tag{1}$$

$$\Omega = \frac{\sum_{i\in S}\sum_{j\in S} v_{ij}}{2}, \tag{2}$$

$$\kappa_i = \sum_{j\in S} v_{ij}, \tag{3}$$

where ρ_{m} is the modularity index. v_{ij} is the edge weight between node i and node j. S is the node set of DNs. Ω is the sum of the edge weight of all networks. κ_i is the sum of the edge weights that are connected to node i. $\phi(i,j)$ is the judgment function of the cluster; when node i and node j are in the same cluster, then $\phi(i,j) = 1$; otherwise, $\phi(i,j) = 0$. v_{ij} is determined by the electrical distance, and the electrical distance can indicate the electrical coupling degree between two nodes in the network. For a DN with N nodes, the electrical distance based on the reactive voltage sensitivity matrix is expressed as

$$L_{ij}^{\mathrm{QV}} = \sqrt{\left(S_{i1}^{\mathrm{QV}} - S_{j1}^{\mathrm{QV}}\right)^2 + ... + \left(S_{i\mathrm{N}}^{\mathrm{QV}} - S_{j\mathrm{N}}^{\mathrm{QV}}\right)^2}, \tag{4}$$

where S_{i1}^{QV} is the element in row i and column j of the reactive voltage sensitivity matrix, which represents the sensitivity of reactive power generation at node i to the voltage at node j. L_{ij}^{QV} is the electrical distance between the two nodes based on the reactive voltage sensitivity matrix; the larger the value of L_{ij}^{QV}, the smaller the electrical distance between the two nodes. Similarly, the electrical distance L_{ij}^{PV} based on the active voltage sensitivity matrix can be obtained. The node voltages are affected by active and reactive power variations, and then, the electrical distance based on the sensitivity matrix is expressed as follows:

$$L_{ij} = \frac{L_{ij}^{\mathrm{QV}} + L_{ij}^{\mathrm{PV}}}{2}. \tag{5}$$

The relationship between the edge weight and electrical distance is that the larger the edge weight, the smaller the electrical distance. Then, the mathematical expression between the edge weight and electrical distance can be obtained as follows:

$$v_{ij} = 1 - \frac{L_{ij}}{\max\left(L_{ij}\right)}, \tag{6}$$

where $\max(L_{ij})$ is the maximum value of the elements in the electrical distance matrix.

2.1.2 Power balance index

In order to avoid large-scale power transfer between clusters, the PV output and load demand within a cluster should be as equal as possible. In order to evaluate the ability of clusters to hold the distributed PVs, this paper proposes the power balance index. The active power balance index φ^{P} is given as follows:

$$\varphi^{\mathrm{P}} = \frac{1}{N_k}\sum_{k=1}^{N_k}\left(1 - \frac{1}{\mathrm{T}}\sum_{t=1}^{\mathrm{T}}\frac{P_{k,t}}{\max P_{k,t}}\right), \tag{7}$$

where N_k is the number of clusters. $P_{k,t}$ is the net power of cluster k at time t. T is the period of optimization. Similarly, the reactive power balance index φ^{Q} is established as follows:

$$\varphi^{\mathrm{Q}} = \frac{1}{N_k}\sum_{k=1}^{N_k}\left(1 - \frac{1}{\mathrm{T}}\sum_{t=1}^{\mathrm{T}}\frac{Q^{\mathrm{need}}}{Q^{\mathrm{sup}}}\right), \tag{8}$$

where Q^{sup} is the maximum value of reactive power supply within the cluster and Q^{need} is the value of reactive power demand within the cluster.

2.1.3 Nodal size index

A reasonable cluster size will directly impact the complexity of the subsequent cluster planning, as well as avoid the differences in the complexity of optimization among different clusters caused by unbalanced cluster size. In addition, a reasonable cluster size can avoid the isolated nodes in the planning process. In order to balance the size of each cluster, a nodal size index is established as follows:

$$\varphi^{\mathrm{M}} = \frac{\mu(i, V[i])}{\mu(i, V - V[i])}, \tag{9}$$

$$\mu(i, V[i]) = \frac{1}{|V[i]|} \sum_{j \in V[i]} v_{ij}, \tag{10}$$

$$\mu(i, V - V[i]) = \frac{1}{|V - V[i]|} \sum_{j \in V - V[i]} v_{ij}, \tag{11}$$

where $V[i]$ is the cluster that node i belongs to. $\mu(x, y)$ is the affiliation degree of node x to cluster y. $|V[i]|$ is the sum of the edge number in the cluster that node i belongs to. $V\text{-}V[i]$ are the clusters that do not contain node i. $|V\text{-}V[i]|$ is the sum of the edge number in clusters that do not contain cluster $V[i]$.

The indexes shown in Eq 1, (7), (8), and (9) are combined into a comprehensive cluster partition index ϕ for the DNs, which is expressed as follows:

$$\phi = \omega^1 \rho + \omega^2 \varphi^{\mathrm{P}} + \omega^3 \varphi^{\mathrm{Q}} + \omega^4 \varphi^{\mathrm{M}}, \tag{12}$$

where ω^1, ω^2, ω^3, and ω^4 are the weights of each index. In the process of cluster partition, different weights can be set for each index depending on different needs.

2.2 Improved genetic algorithm-based cluster partition method

To carry out the cluster partition, a hybrid genetic-simulated annealing (HGSA) algorithm is utilized. The HGSA algorithm uses the annealing selection as the individual replacement strategy, while the global information obtained by the genetic algorithm can be completely used, and the premature convergence of the genetic algorithm can be avoided. Then, the global convergence of the algorithm is enhanced. The cluster partition is implemented as follows:

Step 1: Initial optimization parameters are set: population size n, initial temperature T^0, termination temperature T^{end}, temperature cooling factor r, maximum number of genetic generations M, and objective function for the cluster partitioning index ϕ.

Step 2: The number of temperature updates is set equal to 0. Considering that the cluster partition is carried out based on the original network connectivity, this paper uses the unweighted adjacency matrix A_{ij} to represent the connectivity of the network, where $A_{ij} = 1$ indicates that nodes i and j are connected and $A_{ij} = 0$ indicates that nodes i and j are not connected. The clustering partition is randomly modified when $A_{ij} = 1$. The result of the modification is 0 or 1, and 0 means that the two nodes are disconnected from each other, and then, the two nodes belong to different clusters. In the iteration of the algorithm, in order to ensure that the result after crossover still satisfies the meaning of the adjacency matrix, only the upper triangle of the adjacency matrix is chosen to carry out the crossover. After the crossover step, the upper triangle of the matrix is symmetrically transferred to the lower triangle to form the newborn individuals, thus generating the initial population $P_l(k)$.

Step 3: The individuals that satisfy the cluster constraints are screened. Considering that the reverse power flow only occurs within the cluster, the net power within each cluster needs to be greater than 0 at each moment among the cluster partition. For the individuals who cannot satisfy the constraint, the population eliminates these individuals.

Step 4: Eq. (12) is chosen as an indicator to calculate the fitness of individuals. Considering that the value of fitness intuitively reflects the superiority or inferiority of cluster partition, the individuals with greater fitness are replicated to the offspring to form the new population $P_{ls}(k+1)$.

Step 5: The crossover and mutation are carried out for $P_{ls}(k+1)$ using the traditional genetic algorithm to obtain the population $P_{lv}(k+1)$. The mutated populations are limited to accepting bad solutions by the simulated annealing algorithm to form new populations $P_{lz}(k+1)$.

Step 6: $P_l(k) = P_{lz}(k+1)$ is set. If the genetic algebra accumulates to the maximum number M, then step 7 is repeated; otherwise, step 3 is repeated.

Step 7: The temperature is updated, i.e., $T_l = rT^0$, $k = 0$, $P_{l+1}(k+1) = P_l(k)$. If the convergence condition $T_l < T_{end}$, then the computation to output the optimal solution is terminated. Otherwise, the temperature reduction operation is performed, i.e., $T_{l+1} = rT_l$, and step 3 is repeated.

3 Cluster partition-based two-layer expansion planning method

In this paper, a cluster partition-based two-layer expansion planning model is proposed, which involves the uncertainty of PVs and loads. In the upper layer, a line planning model is established with the objective of minimizing the line investment and network loss costs. In the lower layer, a source-storage planning model is proposed for PVs and ESs with the objective function of minimizing source-storage investment and operation costs within a cluster. Meanwhile, the box-type uncertainty set is utilized to characterize the uncertainty of PVs and loads in the lower layer, and an uncertainty parameter is used to control the range of uncertainty sets, which can reduce the conservatism of the optimization.

3.1 Upper-layer line planning model

In the upper layer, the objective function of the line planning model is established as follows:

$$\min f = \sum_{i,j \in \Omega_s} \left[C^{\mathrm{L}} \sum_{j \in \varphi_i} D_{ij} x_{ij} \frac{\chi(1+\chi)^{\beta}}{(1+\chi)^{\beta} - 1} \right] + C^{\mathrm{e}} \sum_{t=1}^{\mathrm{T}} \sum_{k=1}^{N_k} \sum_{i=1}^{\mathrm{N_{n,k}}} \sum_{j \in \varphi_i} \frac{1}{2} x_{ij} \varepsilon_{ij} i_{ij,t}, \tag{13}$$

where Ω_s is the node set of the branch network. φ_i is the child node set of node i. C^{L} is the investment cost parameter of lines. D_{ij} is the

line length of line *i–j*. x_{ij} is the 0–1 variable, where $x_{ij} = 1$ means that the line is installed and $x_{ij} = 0$ means that the line is not installed. χ is the bank rate. β is the payback period. C^{e} is the electricity price. $N_{n,k}$ is the node numbers of cluster *k*. ε_{ij} is the resistance of the line *i–j*. $i_{ij,t}$ is the squared value of the current in the line *i–j* at time *t*. The constraints include a power flow constraint and a penetration rate constraint of PVs in the clusters.

(1) Power flow constraint

The power flow containing line variables is constrained by second-order conic relaxation (SOCR) (Shaker et al., 2023) as follows:

$$\sum_{i\in\Psi_j}\left(P_{ij,t}-\varepsilon_{ij}i_{ij,t}\right)-x_{ij}P^{\mathrm{L}}_{j,t}+\eta P^{\mathrm{CH}}_{j,t}-\frac{P^{\mathrm{DS}}_{j,t}}{\eta}+P^{\mathrm{PV}}_{j,t}+P^{\mathrm{PVf}}_{j,t}=\sum_{c\in\varphi_j}P_{jc,t}, \tag{14}$$

$$\sum_{i\in\Psi_j}\left(Q_{ij,t}-\delta_{ij}i_{ij,t}\right)-x_{ij}Q^{\mathrm{L}}_{j,t}+Q^{\mathrm{PV}}_{j,t}+Q^{\mathrm{PVf}}_{j,t}=\sum_{c\in\varphi_j}Q_{jc,t}, \tag{15}$$

$$\sum_{j\in\varphi_i}x_{ij}=1, \tag{16}$$

$$v_{j,t}=v'_{i,t}-2\left(\varepsilon_{ij}P_{ij,t}+\delta_{ij}Q_{ij,t}\right)+\left(\varepsilon^2_{ij}+\delta^2_{ij}\right)i_{ij,t}, \tag{17}$$

$$0\le v'_{i,t}\le \mathrm{W}x_{ij}, \tag{18}$$

$$\mathrm{W}\left(1-x_{ij}\right)+v_{i,t}\le v'_{i,t}\le v_{i,t}, \tag{19}$$

$$\left\| 2P_{ij,t}\quad 2Q_{ij,t}\quad i_{ij,t}-v_{i,t}\right\|2\le i_{ij,t}+v_{i,t}, \tag{20}$$

$$x_{ij}v^{\min}\le v_{i,t}\le x_{ij}v^{\max}, \tag{21}$$

$$0\le i_{ij,t}\le x_{ij}\left(I^{\max}\right)^2, \tag{22}$$

where ψ_j is the upstream node set of node *j*. δ_{ij} is the reactance of line *i–j*. $P_{ij,t}$ and $Q_{ij,t}$ are the active and reactive power through line *i–j* at time *t*, respectively. η is the ES charging and discharging efficiency. $P^{\mathrm{CH}}_{j,t}$ and $P^{\mathrm{DS}}_{j,t}$ are the charging and discharging power of ES at node *j* at time *t*, respectively. $P^{\mathrm{L}}_{j,t}$ and $Q^{\mathrm{L}}_{j,t}$ are the active and reactive power of loads at node *j* at time *t*, respectively. $P^{\mathrm{PV}}_{j,t}$ is the actual active power generated by the PV at node *j* at time *t*. $P^{\mathrm{PVf}}_{j,t}$ is the additional active power required to be generated by the PV at node *j* at time *t*. φ_j is the downstream node set of node *j*. $Q^{\mathrm{PV}}_{j,t}$ is the actual reactive power generated by the PV inverter at node *j* at time *t*. $Q^{\mathrm{PVf}}_{j,t}$ is the additional reactive power required to be generated by the PV inverter at node *j* at time *t*. $v_{i,t}$ is the square of the voltage at node *i* at time *t*. $v'_{i,t}$ is the square of the voltage at node *i* that is constrained by the line variable x_{ij}. W is a large constant. $v^{\min}$ and $v^{\max}$ are the upper and lower voltage limits, respectively. $I^{\max}$ is the maximum current limit.

(2) Penetration rate constraint of PVs

Many isolated nodes exist in the DN that need to be connected to the system, and the penetration rate of PVs should be constrained when carrying out line planning to access these isolated nodes, which is defined as follows:

$$\sum_{j=1}^{\mathrm{N}_{n,k}}S^{\mathrm{PV}}_j\le\sum_{j=1}^{\mathrm{N}_{n,k}}S^{\mathrm{L}}_j, \tag{23}$$

where S^{PV}_j is the planed capacity of the PV of node *j* in the cluster and S^{L}_j is the load apparent power of node *j* in the cluster.

3.2 Lower-layer PV and ES planning model

In the lower layer, the PV and ES planning model is established to identify the "worst scenario " of the uncertain variables, and based on the worst scenario, the proposed model minimizes the investment and operation costs of source storage within the cluster. Therefore, the objective function for the PV and ES planning model within the cluster is established as

$$\min F=F_1-F_2+F_3, \tag{24}$$

$$F_1=\sum_{j=1,j\in\Pi_k}^{N_{n,k}}\left[C^{\mathrm{PV}}S^{\mathrm{PV,f}}_j\frac{\chi(1+\chi)^{\beta}}{(1+\chi)^{\beta}-1}+C^{\mathrm{OMPV}}\left(S^{\mathrm{PV}}_j+S^{\mathrm{PV,f}}_j\right)\right], \tag{25}$$

$$F_4=\left(C^{\mathrm{sell}}+C^{\mathrm{sub}}\right)\sum_{t=1}^{\mathrm{T}}\sum_{j=1,j\in\Pi_k}^{N_{n,k}}\left[\omega^{\mathrm{PV}}_t\left(S^{\mathrm{PV}}_j+S^{\mathrm{PV,f}}_j\right)-P^{\mathrm{L}}_{j,t}\right], \tag{26}$$

$$F_5=\sum_{j=1,j\in\Pi_k}^{\mathrm{N}_{n,k}}\left[\frac{\rho(1+\rho)^r}{(1+\rho)^r-1}c^{\mathrm{bat}}P^{\mathrm{batt}}_j\right], \tag{27}$$

where F_1 is the annual investment and operation costs of PVs in cluster *k*. F_2 is the annual revenue of PVs in cluster *k*. F_3 is the annual investment costs of ESs in cluster *k*. C^{PV} is the investment cost parameter of PVs. C_{OMPV} is the annual fixed maintenance cost parameter of PVs. Π_k is the node set of cluster *k*. S^{PV}_j is the original installed PV capacity of node *j* in cluster *k*. $S^{\mathrm{PV,f}}_j$ is the planned PV capacity of node *j* in cluster *k*. C^{sell} and C^{sub} are the feed-in tariff and subsidized tariff of the PVs, respectively. ω^{PV}_t is the output limit of PVs per megawatt for a given sunshine condition at time *t*. ρ is the discount rate. *r* is the discounted number of years. c^{bat} is the investment cost parameter of ESs. $P_{\mathrm{batt},j}$ is the allocated capacity of ES at node *j*. The constraints include uncertainty constraints for PVs and loads, PV and ES capacity constraints, and power flow constraints.

(1) Uncertainty constraints for PVs and loads

The box uncertainty set is utilized to characterize the uncertainty range of active load power, reactive load power, and PV outputs. Meanwhile, an uncertainty parameter is utilized that can be set to adjust the conservativeness of the optimal solution. The larger the uncertainty parameter, the more conservative the solution. The specific formula is as follows:

$$\begin{cases} U=\left\{\boldsymbol{u}=\left[P^{\mathrm{L}}_{j,t},Q^{\mathrm{L}}_{j,t},P^{\mathrm{PV}}_{j,t}\right]^{\mathrm{T}}\middle| P^{\mathrm{L}}_{j,t}\in\left[P^{\mathrm{L,F}}_{j,t}-\Delta P^{\mathrm{L}}_{j,t},P^{\mathrm{L,F}}_{j,t}+\Delta P^{\mathrm{L}}_{j,t}\right],\sum_{j=1}^{N_a}\frac{\left|P^{\mathrm{L}}_{j,t}-P^{\mathrm{L,F}}_{j,t}\right|}{\Delta P^{\mathrm{L}}_{j,t}}\le\Gamma^{\mathrm{PL}};\right. \\ Q^{\mathrm{L}}_{j,t}\in\left[Q^{\mathrm{L,F}}_{j,t}-\Delta Q^{\mathrm{L}}_{j,t},Q^{\mathrm{L,F}}_{j,t}+\Delta Q^{\mathrm{L}}_{j,t}\right],\sum_{j=1}^{N_b}\frac{\left|Q^{\mathrm{L}}_{j,t}-Q^{\mathrm{L,F}}_{j,t}\right|}{\Delta Q^{\mathrm{L}}_{j,t}}\le\Gamma^{\mathrm{QL}}; \\ \left.P^{\mathrm{PV}}_{j,t}\in\left[P^{\mathrm{PV,F}}_{j,t}-\Delta P^{\mathrm{PV}}_{j,t},P^{\mathrm{PV,F}}_{j,t}+\Delta P^{\mathrm{PV}}_{j,t}\right],\sum_{e=1}^{N_c}\frac{\left|P^{\mathrm{PV}}_{j,t}-P^{\mathrm{PV,F}}_{j,t}\right|}{\Delta P^{\mathrm{PV}}_{j,t}}\le\Gamma^{\mathrm{PV}};\right\}, \end{cases} \tag{28}$$

where *U* represents the boxed uncertainty set. $P^{\mathrm{L,F}}_{j,t}$, $Q^{\mathrm{L,F}}_{j,t}$, and $P^{\mathrm{PV,F}}_{j,t}$ are the predicted values of active load power, reactive load power, and PV outputs of node *j* at time *t* in cluster *k*, respectively. $\Delta P^{\mathrm{L}}_{j,t}$, $\Delta Q^{\mathrm{L}}_{j,t}$, and $\Delta P^{\mathrm{PV}}_{j,t}$ are the fluctuation deviations of active load power,

reactive load power, and PV output of node j at time t in cluster k, respectively. N_a, N_b, and N_e are the total number of active load nodes, the total number of reactive load nodes, and the total number of PV nodes in cluster k, respectively. Γ^{PL}, Γ^{QL}, and Γ^{PV} are the uncertain adjustment parameters for active load power, reactive load power, and PV output S, respectively, which are integers from 0 to N_a, 0 to N_b, and 0 to N_e, respectively. The decision-maker can choose a variety of uncertainty regulation parameters to adjust the scheme flexibly; the larger the value of each uncertainty regulation parameter, the more conservative the resulting planning scheme.

(2) PV and ES constraints

$$T_k^{\mathrm{ES}} = \bar{t}_1, \tag{29}$$

$$\sum_{j\in\Pi_k}\eta P_{j,t}^{\mathrm{CH}} - \sum_{j\in\Pi_k}\frac{P_{j,t}^{\mathrm{DS}}}{\eta} - \sum_{j\in\Pi_k}P_{j,t}^{\mathrm{PV}} - \sum_{j\in\Pi_k}P_{j,t}^{\mathrm{PVf}} + \sum_{j\in\Pi_k}P_{j,t}^{\mathrm{L}} = 0, t\in T_k^{\mathrm{ES}}, \tag{30}$$

$$soc^{\min}P_j^{\mathrm{batt}} \le soc_{j,t} \le soc^{\max}P_j^{\mathrm{batt}}, j\in\Pi_k, t\in T_k^{\mathrm{ES}}, \tag{31}$$

$$soc_{j,t} = soc_{j,t-1} + \frac{\eta P_{j,t}^{\mathrm{CH}} - \frac{P_{j,t}^{\mathrm{DS}}}{\eta}}{P_j^{\mathrm{batt}}}, j\in\Pi_k, t\in T_k^{\mathrm{ES}}, \tag{32}$$

$$0 \le \eta P_{j,t}^{\mathrm{CH}} \le P^{\mathrm{CH,max}}, j\in\Pi_k, t\in T_k^{\mathrm{ES}}, \tag{33}$$

$$0 \le \frac{P_{j,t}^{\mathrm{DS}}}{\eta} \le P^{\mathrm{DS,max}}, j\in\Pi_k, t\in T_k^{\mathrm{ES}}, \tag{34}$$

$$0 \le P_{j,t}^{\mathrm{PV}} \le P_j^{\mathrm{L,max}}, \tag{35}$$

$$S_j^{\mathrm{PVf}} = \max\left(P_{j,t}^{\mathrm{PVf}}\right), j\in\Pi, t\in T_k^{\mathrm{ES}}, \tag{36a}$$

where T_k^{ES} is the simulated operating time of the ESs in cluster k. $\overline{t_1}$ is the annual PV generation time. $soc_{j,t}$ is the state of charge of ES at node j in cluster k at time t. $soc^{\min}$ and $soc^{\max}$ are the minimum and maximum charge states of ES, respectively. $P^{\mathrm{CH,max}}$ and $P^{\mathrm{DS,max}}$ are the maximum power limits for charging and discharging power of ES, respectively. $P_j^{\mathrm{L,max}}$ is the maximum load demand of node j in the cluster.

(3) Power flow constraints

Eqs (14), (15), and (17) are referred to for the power flow constraints.

3.3 Iterative solving process

In this paper, an iterative solution method is utilized to solve the proposed planning models. For the upper layer, the planning model can be expressed by a specific form as follows:

$$\begin{cases} \min\limits_{\boldsymbol{x}} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\alpha} \\ s.t.\ \boldsymbol{P}\boldsymbol{x} = \boldsymbol{W} \\ \boldsymbol{\alpha} \ge \boldsymbol{k}^{\mathrm{T}}\boldsymbol{y}_l \\ \boldsymbol{B}\boldsymbol{x} + \boldsymbol{C}\boldsymbol{y}_l \le \boldsymbol{D} \\ \boldsymbol{R}_x\boldsymbol{u}_l^* = \boldsymbol{O}\boldsymbol{y}_l \\ \boldsymbol{S}_u^*\boldsymbol{y}_l \le \boldsymbol{V} \\ \left\|\boldsymbol{G}\boldsymbol{y}_l\right\| \le \boldsymbol{H}^{\mathrm{T}}\boldsymbol{y}_l \\ \forall l \le k, \end{cases} \tag{36b}$$

where $\boldsymbol{x}$ is the optimization vectors. $\boldsymbol{c}$ is the coefficient matrices corresponding to the objective functions. $\boldsymbol{P}$, $\boldsymbol{k}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{R_x}$, $\boldsymbol{S^*u}$, $\boldsymbol{G}$, and $\boldsymbol{H}$ are the coefficient matrices corresponding to the variables under the constraints. $\boldsymbol{\alpha}$, $\boldsymbol{D}$, $\boldsymbol{O}$, $\boldsymbol{V}$, and $\boldsymbol{W}$ are the constant column vectors. l is the current number of iterations. k is the maximum iteration. $\boldsymbol{y}_l$ is the variable at the lth iteration. $\boldsymbol{u}^*\ l$ is the value of the uncertain variable u in the "worst" scenario after the lth iteration.

The specific form of the lower-layer model is

$$\begin{cases} \max\limits_{\boldsymbol{u}\in U}\ \min\limits_{\boldsymbol{y}\in\boldsymbol{\Omega}(\boldsymbol{x},\boldsymbol{u})} \boldsymbol{d}^{\mathrm{T}}\boldsymbol{y} \\ s.t.\ \boldsymbol{B}\boldsymbol{x}^* + \boldsymbol{C}\boldsymbol{y} \le \boldsymbol{D} \\ \boldsymbol{R}_u^*\boldsymbol{u} = \boldsymbol{O}\boldsymbol{y} \\ \boldsymbol{S}_u\boldsymbol{y} \le \boldsymbol{V} \\ \left\|\boldsymbol{G}\boldsymbol{y}\right\| \le \boldsymbol{H}^{\mathrm{T}}\boldsymbol{y}, \end{cases} \tag{37}$$

where $\boldsymbol{y}$ is the optimization vectors. $\boldsymbol{\Omega}(\boldsymbol{x},\boldsymbol{u})$ is the feasible domain of $\boldsymbol{y}$ under the given $\boldsymbol{x}$ and $\boldsymbol{u}$. $\boldsymbol{x}^*$ is the optimization vectors obtained from the upper layer. $\boldsymbol{d}$ is the coefficient matrices corresponding to the objective functions. $\boldsymbol{S_u}$ and $\boldsymbol{R^*\ u}$ are the coefficient matrices corresponding to the variables under the constraints. Given a set of $\boldsymbol{u}$, the inner min. problem becomes a second-order cone programming problem, which can be transformed into a "max" form by the duality theory:

$$\begin{cases} \max\limits_{\boldsymbol{u}\in U,\gamma,\nu,\pi,\mu,\mu_1} (\boldsymbol{D} - \boldsymbol{B}\boldsymbol{x}^*)^{\mathrm{T}}\gamma + (\boldsymbol{R_x}\boldsymbol{u})^{\mathrm{T}}\nu + \boldsymbol{u}^{\mathrm{T}}\pi \\ s.t.\ \boldsymbol{C}^{\mathrm{T}}\gamma + \boldsymbol{O}^{\mathrm{T}}\pi + \boldsymbol{S_u}^{\mathrm{T}}\nu + \sum\limits_i\left(\boldsymbol{G}_i\mu_i + \boldsymbol{H}_i\mu_{1_i}\right) \le \boldsymbol{d} \\ \left\|\mu_i\right\|_2 \le \mu_{1_i}, \forall i = 1, 2, ..., j \\ \gamma, \pi, \nu \ge 0, u \in U, \end{cases} \tag{38}$$

where$[\gamma,\nu,\pi,\mu,\mu_1]$ are the dual variables in the lower-layer model.

After the above transformation, the proposed model can be solved by the iterative solution method as follows:

Step 1: Given a set of $\boldsymbol{u}$ values as the initial worst-case scenario, a lower bound is set on the operating cost $LB = -\infty$, an upper bound $UB = +\infty$, and the number of iterations $l = 1$.

Step 2: The upper model is solved based on the worst-case scenario $\boldsymbol{u}_l$ to obtain the optimal solution $\boldsymbol{x}^*\ l$ and $\boldsymbol{\alpha}^*\ l$. The value of $\boldsymbol{\alpha}^*\ l$ is used as the new lower bound $LB = \max(LB, \boldsymbol{\alpha}^*\ l)$.

Step 3: The lower layer is optimized based on the optimization results of the upper layer, and the optimized results $f_l\ (\boldsymbol{x}^*\ l)$ and the worst-case scenario $\boldsymbol{x}^*\ l$ are obtained. The upper bound is updated as $UB = \min(UB, f_l\ (\boldsymbol{x}^*\ l))$.

Step 4: If $UB\text{-}LB < \varepsilon$, where ε *is* a threshold of convergence, then, the optimal solutions can be obtained, and the iteration is stopped. Otherwise, the variable $\boldsymbol{y}^{l+1}$ and the following constraints are added:

$$\begin{cases} \boldsymbol{\alpha} \ge \boldsymbol{k}^{\mathrm{T}}\boldsymbol{y}^{l+1} \\ \boldsymbol{B}\boldsymbol{x} + \boldsymbol{C}\boldsymbol{y}^{l+1} \le \boldsymbol{D} \\ \boldsymbol{R_x}\boldsymbol{u}_{l+1}^* = \boldsymbol{O}\boldsymbol{y}_{l+1} \\ \boldsymbol{S}_u^{*l+1}\boldsymbol{y}_{l+1} \le \boldsymbol{V} \\ \left\|\boldsymbol{G}\boldsymbol{y}_{l+1}\right\| \le \boldsymbol{H}^{\mathrm{T}}\boldsymbol{y}_{l+1}. \end{cases} \tag{39}$$

Let $l = l + 1$, and step 2 is repeated until the algorithm converges.

4 Case study

4.1 Case study system

In order to verify the effectiveness of the proposed method, an actual 35 kV/10 kV DN in China is utilized for analysis. The total

TABLE 1 Installed capacity of photovoltaics (PVs).

Node location	Installed capacity (kW)	Node location	Installed capacity (kW)	Node location	Installed capacity (kW)
7	500	16	500	21	500
9	500	17	500	27	1,500

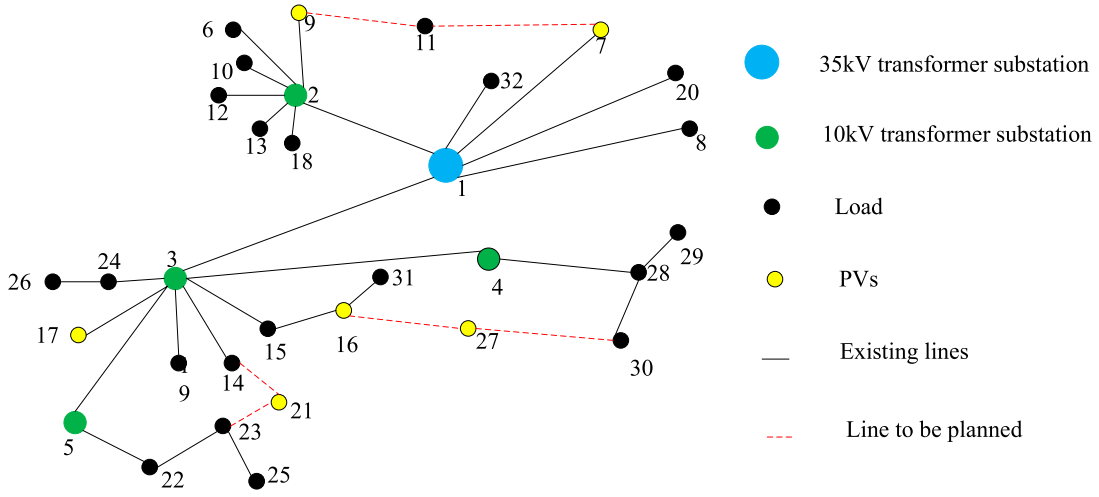

FIGURE 1
Physical topology of the system.

TABLE 2 Lines to be planned.

Lines	Distance/km	Line	Distance/km	Line	Distance/km
7–11	15.64	30–27	17.62	16–27	23.64
9–11	22.37	14–21	6.50	23–21	5.30

loads in this DN are 7.3 MVA, and the total capacity of PVs in this DN is 4 MW. The installed PV capacity is shown in Table 1. The topology of this system is shown in Figure 1, the lines to be planned are shown in Table 2, and the line parameters of the network are shown in Table 3.

The simulated genetic annealing algorithm used in this paper sets the population size n = 40, the initial temperature T_0 = 100°C, the termination temperature T_{end} = 1°C, the temperature cooling factor r = 0.8, the maximum number of genetic generations N = 500, the crossover probability p_c = 0.4, and the variation probability p_m = 0.2. This paper uses MATLAB version 2019 to program the algorithm.

4.2 Analysis of cluster partition

In order to illustrate the superiority of the proposed cluster partition index in this paper, the modularity function used by Wang et al. (2021) is selected to compare with the proposed cluster partition index. The weights of the proposed cluster partition index in this paper are taken as $\omega^1 = \omega^2 = \omega^3 = \omega^4 = 0.25$. The cluster partition results under the proposed cluster partition index are shown in Figure 2, and the cluster partitioning results under the modularity function are shown in Figure 3.

Figure 2 and Figure 3 show that the size difference among sub-networks attained by the proposed cluster partition index is smaller than that attained by the modularity function, and isolated nodes forming separate clusters (cluster 3 and cluster 6) occur in modularity function. In addition, there are PVs in each cluster attained by the proposed method, while no PVs exist in cluster 1 attained by the modularity function. The differences between the two methods prove that the cluster partition index can result in a more reasonable cluster partition for the subsequent cluster control and operation.

In order to illustrate the superiority of the proposed cluster partitioning algorithm, the traditional genetic algorithm is selected to be compared. The comparison of cluster partition results obtained by the two algorithms is shown in Table 4. Table 4 shows that the optimization results of each cluster partitioning index of the proposed algorithm are greater than those of the traditional

10.3389/fenrg.2024.1390073

TABLE 3 Line parameters of the network.

Start node	Final node	Line impedance (Ω/km)	Start node	Final node	Line impedance (Ω/km)
1	2	0.192 + 20.533 j	4	29	0.515 + 0 j
1	3	0.192 + 0 j	4	30	0.13 + 0 j
1	7	0.192 + 13.200 j	5	22	0.13 + 6.482 j
1	8	0.61 + 1.820 j	7	11	0.299 + 0.823 j
1	20	4.681 + 8.314 j	9	11	0.12 + 0.471 j
1	32	7.686 + 14.750 j	14	21	0.186 + 0.261 j
2	6	1.084 + 3.236 j	15	16	0.21 + 0.730 j
2	9	2.132 + 3.786 j	16	27	0.9 + 0.804 j
2	10	1.703 + 4.690 j	16	31	0.731 + 1.298 j
2	12	0.13 + 10.323 j	23	21	3.773 + 5.297 j
2	13	0.238 + 16.259 j	24	29	1.24 + 2.662 j
2	18	0.238 + 0 j	28	29	2.388 + 4.242 j
3	4	5.628 + 16.797 j	28	30	0.515 + 29.563 j
3	5	5.616 + 9.975 j	30	27	2.495 + 4.432 j
3	14	0.238 + 10.209 j	33	29	4.231 + 3.621 j
3	15	0.476 + 0 j	34	30	3.681 + 9.235 j
3	17	4.456 + 7.915 j	35	3	7.214 + 3.751 j
3	19	2.7 + 3.790 j	4	28	4.396 + 6.172 j
3	24	0.64 + 1.290 j			

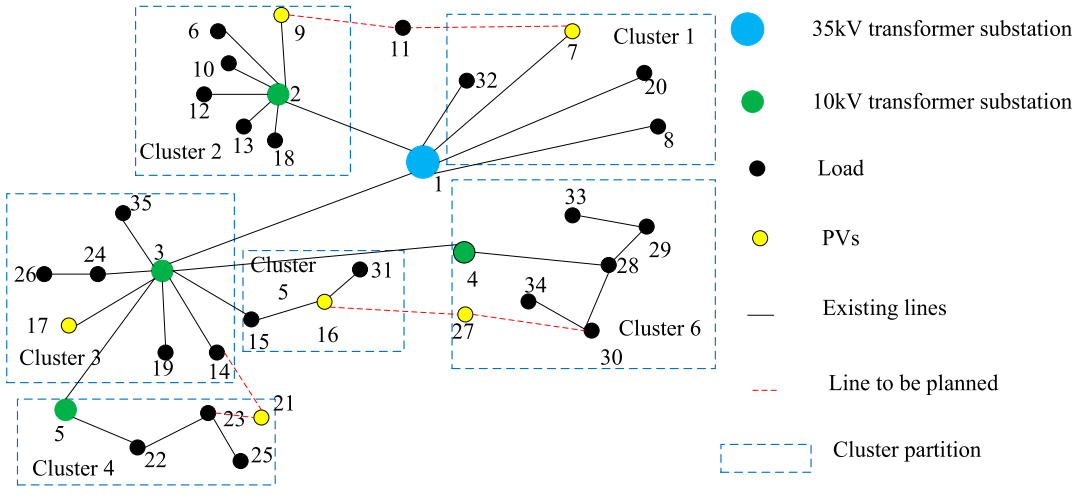

FIGURE 2
Partition results under the proposed cluster partition index.

genetic algorithm, which indicates that the proposed cluster partitioning algorithm has better optimization performance.

Different combinations of partition index weights are given in Table 5. As shown in Table 5, when increasing the weight of the modularity index, the nodes within the cluster are better connected, but the cluster power complementarity decreases. When increasing the weight of the active and reactive power balance index, the cluster power complementarity characteristics improve, but the nodes within the cluster are significantly less connected. When increasing the weight of the nodal size index, the number of clusters changes. When increasing the weight of the nodal size index, the number of

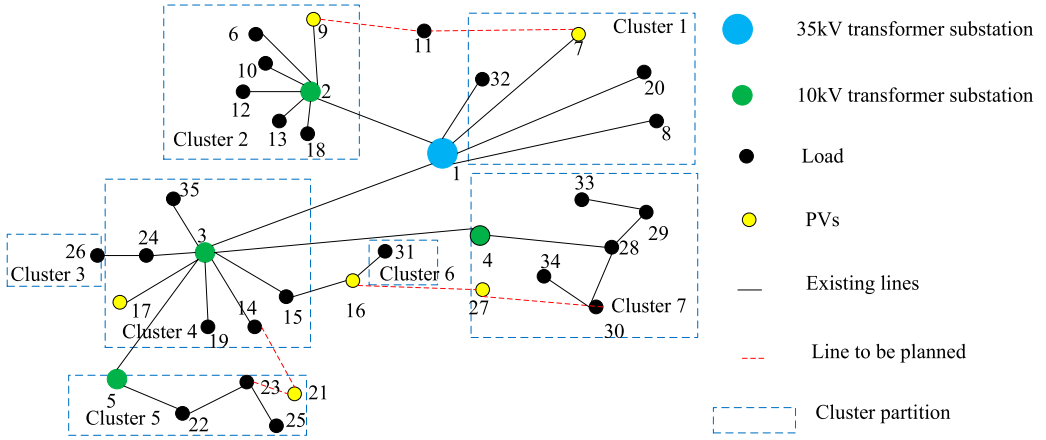

FIGURE 3
Partition results under the modularity function.

TABLE 4 Comparison of cluster partition results obtained by different algorithms.

Result	ρ_m	φ_P	φ_Q	φ_M
Proposed algorithm	0.6953	0.8639	0.6866	0.7233
Traditional genetic algorithms	0.6132	0.8071	0.6120	0.6982

clusters changes, and the cluster partition is more focused on the change in cluster size.

4.3 Analysis of the proposed planning strategy

In order to verify the flexibility of the proposed two-layer expansion planning model, three sets of uncertain regulation parameters are selected and compared in terms of network connection, planned capacity of PVs and ESs, and total planning costs. Three sets of uncertain regulation parameters are set as follows:

Case 1: $\Gamma^{PL} = \Gamma^{QL} = \Gamma^{PV} = 0$ is set, and in this case, the active load power, reactive load power, and PV output are all equal to the predicted values.

Case 2: $\Gamma^{PL} = \Gamma^{QL} = 16$ and $\Gamma^{PV} = 4$ are set, and in this case, the active and reactive load power of 16 nodes is taken to the minimum value of the prediction interval, while the PV outputs of 4 nodes are taken to the maximum value of the prediction interval.

Case 3: $\Gamma^{PL} = \Gamma^{QL} = 27$ and $\Gamma^{PV} = 6$ are set. In this case, the active and reactive load power of 27 nodes is taken to the minimum value of the prediction interval, while the PV outputs of 4 nodes are taken to the maximum value of the prediction interval, which is the worst scenario.

The comparisons of the three cases in planned PV capacity and ES capacity are shown in Figure 4 and Figure 4, and the planning results and total planning costs are shown in Table 6 and Table 7, respectively.

Figure 4 shows that with the increase in uncertainty regulation parameters, the planned PV capacity decreases. The reason is that the scenario becomes increasingly severe, and to maintain the safe operation of the cluster, the corresponding PVs are reduced.

As shown in Table 6, with the increase in uncertainty regulation parameters, the connection of nodes 21 and 27 is different under the three cases. The reason is that as the operating scenario becomes more severe, and the load power increases, the planning results tend to favor a reduction in the power supply range.

As shown in Figure 5, as the conservatism of the planning increases, the ES installed capacity also increases. The reason is that the active power and reactive power of the load in cases 2 and 3 are smaller than those of case 1, and the PV outputs are larger; hence, more ESs are needed to mitigate the uncertainty of PVs and loads.

In addition, Table 7 shows that as the uncertainty regulation parameter increases, although the annual investment and operation

TABLE 5 Influence of different weights on cluster partition.

Combination	ω_1	ω_2	ω_3	ω_4	ρ_m	φ_P	φ_Q	φ_M	Number of clusters
1	0.1	0.4	0.4	0.1	0.632 1	0.896 4	0.711 7	0.701 6	7
2	0.2	0.3	0.3	0.2	0.677 9	0.889 1	0.679 3	0.708 3	6
3	0.3	0.2	0.2	0.3	0.742 4	0.833 0	0.657 1	0.739 3	6
4	0.4	0.1	0.1	0.4	0.789 1	0.798 3	0.625 3	0.756 1	5

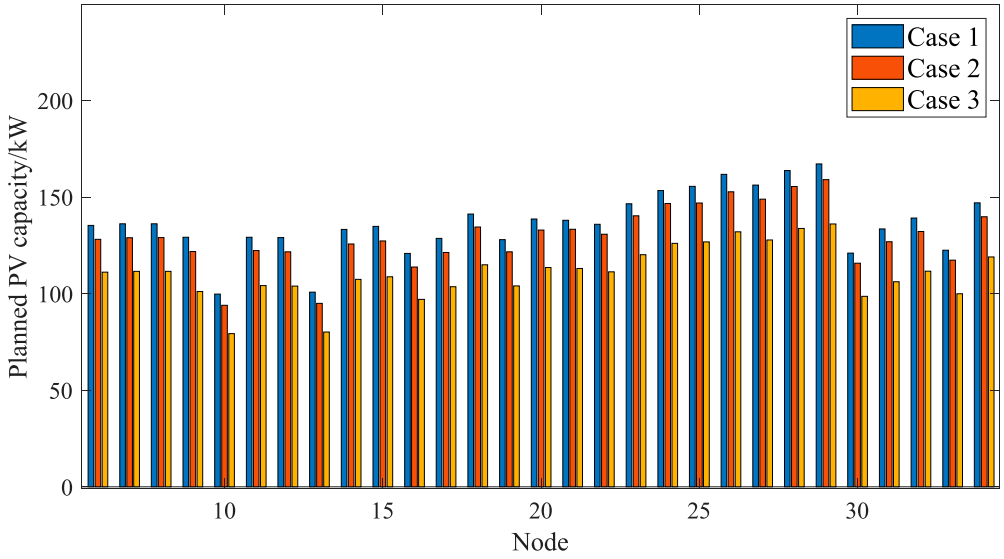

FIGURE 4
Planned photovoltaic (PV) capacity under different cases.

TABLE 6 Planning results of different schemes.

Case	Line connection results		Cluster planning results
	Connectivity	Disconnection	
1	7–11	9–11	Cluster 1 (7,8,11,20,32); cluster 2 (2,6,9,10,12,13,18)
	23–21	14–21	Cluster 3 (3,14,17,19,24,26,35); cluster 4 (5,21,22,23,25)
	30–27	16–27	Cluster 5 (15,16,31); cluster 6 (4,27,28,29,30,33,34)
2	7–11	9–11	Cluster 1 (7,8,11,20,32); cluster 2 (2,6,9,10, 12,13,18)
	14–21	23–21	Cluster 3 (3,14,17,19,21,24,26,35); cluster 4 (5, 22,23,25)
	30–27	16–27	Cluster 5 (15,16,31); cluster 6 (4,27,28,29,30,33,34)
3	7–11	9–11	Cluster 1 (7,8,11,20,32); cluster 2 (2,6,9,10,12,13,18)
	14–21	23–21	Cluster 3 (3,14,17,19,21,24,26,35); cluster 4 (5,22,23,25)
	16–27	30–27	Cluster 5 (15,16,27,31); cluster 6 (4,28,29,30,33,34)

TABLE 7 Planning costs in different cases.

Comparison item	Case 1	Case 2	Case 3
Annual investment costs of photovoltaics (PVs)	587.54 $	553.48 $	528.22 $
Annual operation costs of PVs	414.93 $	232.90 $	187.04 $
Annual income of PVs	240.43 $	247.92 $	285.45 $
Annual investment costs of lines	622.98 $	856.43 $	914.28 $
Network losses	139.82 $	154.68 $	185.82 $
Annual investment costs of energy storages (ESs)	1,062.34 $	1,579.61 $	1,726.73 $

costs of PVs correspondingly decrease, the annual income of PVs decreases, and the annual investment costs of networks, network losses, and the annual investment costs of ESs correspondingly increase. The reason is that the more uncertainty of DNs considered in the planning process, the more conservative the resulting solution becomes, and the corresponding total cost increases.

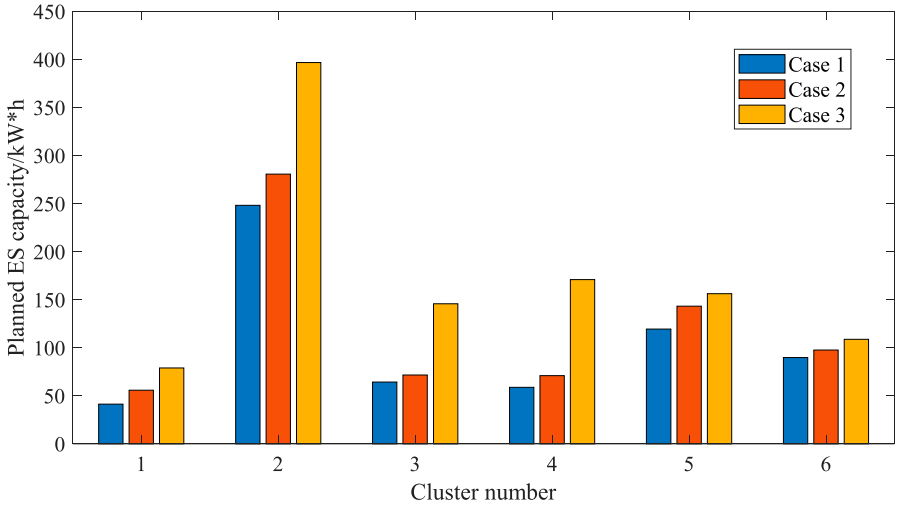

FIGURE 5
Planned energy storage (ES) capacity under different cases.

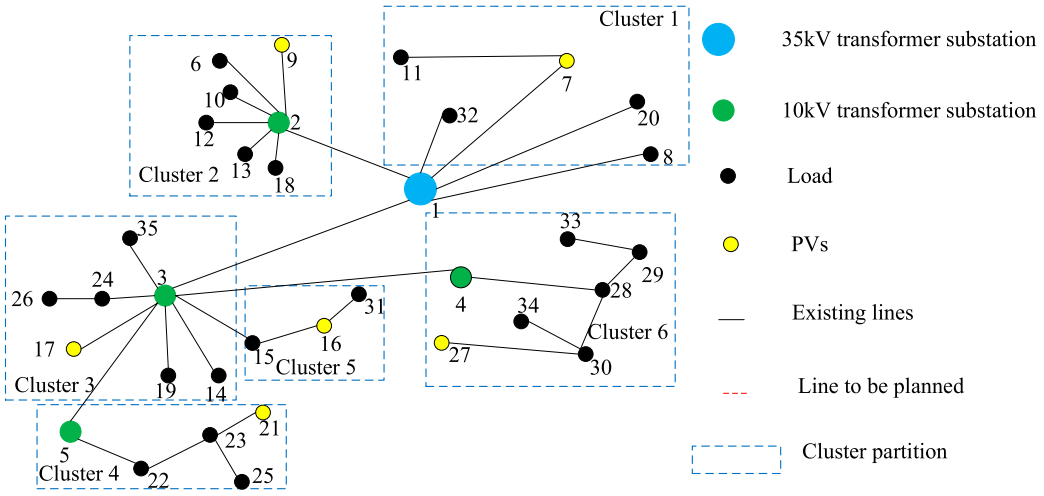

FIGURE 6
Planning results based on the proposed method.

In order to further illustrate the superiority of the proposed planning method, the centralized planning method is selected to compare with the proposed planning method. The uncertain regulation parameters in this comparison are set as $\Gamma^{\mathrm{PL}} = \Gamma^{\mathrm{QL}} = 12$ and $\Gamma^{\mathrm{PV}} = 3$, and in this case, the active and reactive load power of 12 nodes is taken to the minimum value of the prediction interval, while the PV outputs of 3 nodes are taken to the maximum value of the prediction interval. The network connection of the two methods is shown in Figure 6 and Figure 7. The comparison of the calculation time and total planning cost of the two methods is shown in Table 6.

Figure 6 and Figure 7 show that the connection of node 11 and node 27 under the proposed method is different from the centralized planning scheme. The length of the power supply lines for node 11 and node 27 under the proposed method has been reduced by 6.73 km and 6.02 km, respectively; the power supply range is greatly shortened, which can effectively reduce the line investment costs and network loss costs, as well as improve the economic efficiency of the planning scheme.

The calculation method for the difference rate given in Table 8 is to subtract the results of the proposed method from those of the centralized planning method and divide them by the results of the centralized planning method. As shown in Table 8, in the proposed method, the annual PV investment costs and PV operation costs, as well as the annual income of PVs, increased by 5.3% and 14.1%, respectively, compared to the centralized planning method, and the annual investment costs of lines, network losses, and annual investment costs of ESs decreased by 13.5%, 25.7%, and 15.5%, respectively. These different rates indicate that the proposed method can improve the overall economic efficiency of the planning, as well as reduce the costs of planning. In addition, the calculation time of

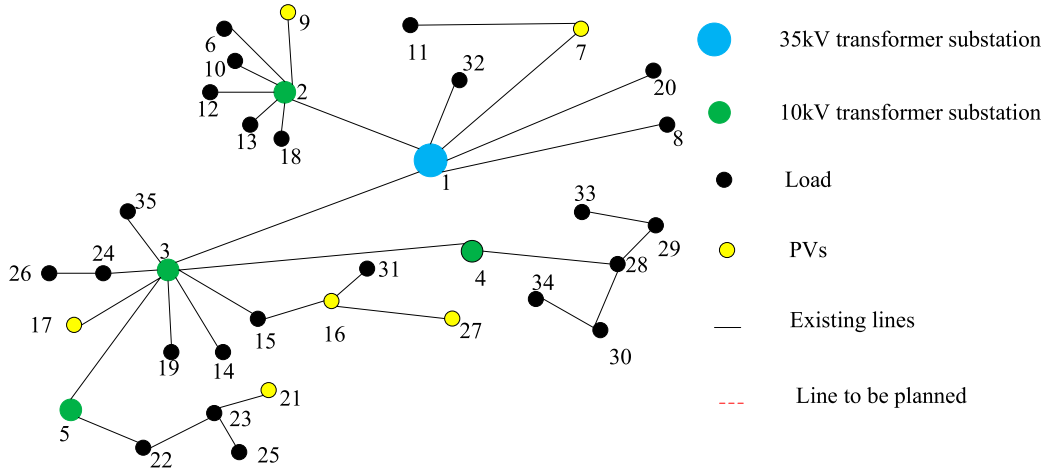

FIGURE 7
Planning results based on the centralized planning method.

TABLE 8 Comparison between the proposed method and centralized planning method.

Comparison item	Proposed method	Centralized planning method	Difference rate (%)
Annual investment costs of photovoltaics (PVs)	569.86 $	541.11 $	−5.3
Annual operation costs of PVs	293.10 $	239.87 $	−14.1
Annual income of PVs	240.43 $	277.97 $	13.5
Annual investment costs of lines	692.89 $	932.30 $	25.7
Network losses	150.14 $	177.58 $	15.5
Annual investment costs of energy storages (ESs)	1,360.22 $	1,689.09 $	10.3
Calculation time	20.41 s	43.62 s	53.2

the proposed method is 20.41 s, but that of the centralized planning method is 43.62 s. Owing to the fact that the centralized planning method requires a whole optimization process, the huge variables complex the optimization. However, the proposed method can partition the complex network into smaller sub-networks, and the complex problem is divided into several relatively simple sub-problems that can be solved rapidly and easily by a parallel computing way. Therefore, the computing time can be greatly shortened. The optimization speed of the proposed method is improved by 53.2% compared to centralized planning methods, which indicates that proposed method is more suitable for DN planning with large-scale PVs.

5 Conclusion

In this paper, a cluster partition-based two-layer expansion planning for DNs is proposed, and an actual 35 kV/10 kV DN in China is utilized for analysis. The results show that

1) To deal with poor power balance and unbalanced cluster size in existing cluster partition, a comprehensive cluster partition index is proposed in this paper, which includes the modularity index, power balance index, and nodal size index, and in order to avoid the cluster partition falling into local optimum, an improved genetic algorithm is utilized to carry out the network partition. The differences in cluster size among the clusters attained by the proposed cluster partition are smaller, and every cluster has PVs.
2) To deal with complex models in centralized planning methods, a cluster partition-based two-layer expansion planning model is established for the DNs, which decomposes the complex centralized planning model into cluster planning. The proposed method can improve the overall economic efficiency of the planning, as well as reduce the costs of planning. Meanwhile, the optimization speed is also improved.
3) To reduce the conservatism of traditional robust optimization, a box uncertainty set is utilized to characterize the uncertainty of loads and PVs, and an uncertainty regulation parameter is used to control the range of uncertainty sets, which can reduce the conservatism of the optimization, as well as simplify the calculation process.

In the actual operation, the PV outputs are uncontrollable and may exceed the worst scenario set in the planning stage. Then, the PV outputs within the cluster may not be fully absorbed. In future research, a cluster partition-based scheduling method for DNs should be further researched to enhance the source-load complementarity of the DNs.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Author contributions

SY: writing–original draft and writing–review and editing. SS: writing–original draft. YC: writing–review and editing. PY: writing–review and editing. CW: writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was supported by the project supported by the Science and Technology Project of State Grid Shandong Electric Power Company "Research and application of access and control technology targeting the distributed photovoltaics developed by the entire county (city, district)" (520626220014).

Conflict of interest

The authors declare that this study received funding from State Grid Shandong Electric Power Company. The funder had the following involvement in the study: The funder was involved in the study design.

Publisher's note

References

Bi, R., Liu, X., Ding, M., et al. (2019). A method of renewable energy power generation cluster classification with the goal of improving the consumption capacity. *Chin. J. Electr. Eng.* 39 (22), 6583–6592. doi:10.13334/j.0258-8013.pcsee.182512

Cai, Z., Shuang, H., and Shan, J. (2022). Optimal scheduling strategy for electric bus cluster power exchange considering operation cost. *Power Syst. Autom.* 46 (17), 205–217. doi:10.7500/AEPS20210916007

Ding, M., Fang, H., Bi, R., et al. (2019). Distributed photovoltaic and energy storage siting and capacity planning for distribution networks based on cluster delineation. *Chin. J. Electr. Eng.* 39 (8), 2187–2201. doi:10.13334/j.0258-8013.pcsee.180757

Faramarzi, D., Rastegar, H., Riahy, G. H., and Doagou-Mojarrad, H. (2023). A three-stage hybrid stochastic/IGDT framework for resilience-oriented distribution network planning. *Int. J. Electr. Power Energy Syst.* 146 (5), 108738–108738.15. doi:10.1016/j.ijepes.2022.108738

Ge, J., Liy, Y., Pang, D., et al. (2024). Cluster division voltage control strategy of photovoltaic distribution network with high permeability. *High. Volt. Eng.* 50 (1), 74–82. doi:10.13336/j.1003-6520.hve.20230816

Haji-Aghajani, E., Hasanzadeh, S., and Heydarian-Forushani, E. (2023). A novel framework for planning of EV parking lots in distribution networks with high PV penetration. *Electr. Power Syst. Res.* 217 (4), 109156–109156.12. doi:10.1016/j.epsr.2023.109156

Hu, D., Ding, M., Bi, R., et al. (2020). Analysis of the impact of photovoltaic and wind power complementarity on the planning of high penetration renewable energy cluster access. *Chin. J. Electr. Eng.* 40 (3), 821–836. doi:10.13334/j.0258-8013.pcsee.182308

Hu, D., Sun, L., and Ding, M. (2023). Integrated planning of an active distribution network and DG integration in clusters considering a novel formulation for reliability assessment. *CSEE J. Power Energy Syst.* 9 (2), 561–576. doi:10.17775/CSEEJPES.2020.00150

Ji, Y., Chen, X., He, P., Liu, X., Wu, X., and Zhao, C. (2023). A novel voltage/var sensitivity calculation method to partition the distribution network containing renewable energy. *Recent Adv. Electr. Electron. Eng.* 16 (4), 380–394. doi:10.2174/2352096516666221130150549

Jiang, Y., Ren, Z., Li, Q., et al. (2022). New energy consumption strategy in distribution networks considering coordinated scheduling of multi-flexibility resources. *J. Electr. Eng. Technol.* 37 (7), 1820–1835. doi:10.19595/j.cnki.1000-6753.tces.211464

Kong, X., Liu, C., Chen, S., et al. (2022). A multi-time-node response potential assessment method for adjustable resource clusters considering dynamic processes. *Power Syst. Autom.* 46 (18), 55–64. doi:10.7500/AEPS20220501001

Koutsoukis, N., Georgilakis, P., and Hatziargyriou, N. (2018). Multistage coordinated planning of active distribution networks. *IEEE Trans. Power Syst.* 33 (1), 32–44. doi:10.1109/tpwrs.2017.2699696

Lei, C., Wang, Q., Zhou, G., Bu, S., Lin, T., et al. (2023). Probabilistic wind power expansion planning of bundled wind-thermal generation system with retrofitted coal-fired plants using load transfer optimization. *Int. J. Electr. Power Energy Syst.* 151 (9), 109145–109145.18. doi:10.1016/j.ijepes.2023.109145

Li, C., Dong, Z., Li, J., et al. (2021). Optimal control strategy for distributed energy storage clusters to enhance new energy consumption capacity of distribution networks. *Power Syst. Autom.* 45 (23), 76–83. doi:10.7500/AEPS20210224004

Li, P., Shen, J., Wu, Z., Yin, M., Dong, Y., and Han, J. (2023). Optimal real-time Voltage/Var control for distribution network: droop-control based multi-agent deep reinforcement learning. *Int. J. Electr. Power Energy Syst.* 153 (11), 109370–109370.11. doi:10.1016/j.ijepes.2023.109370

Li, S., and Yang, X. (2022). Siting and capacity determination of photovoltaic energy storage based on bi-directional dynamic reconfiguration and cluster delineation. *Power Syst. Prot. Control* 50 (3), 51–58. doi:10.19783/j.cnki.pspc.210385

Li, Y., Lu, N., Liu, X., et al. (2022c). Output evaluation method of distributed photovoltaic cluster considering renewable energy accommodation and power loss of network. *Electr. Power Constr.* 43 (10), 136–146. doi:10.12204/j.issn.1000-7229.2022.10.013

Li, Y., Yao, T., Qiao, X., et al. (2022b). Optimal allocation of distributed PV and energy storage based on joint timing scenarios and source-grid-load synergy. *J. Electr. Eng. Technol.* 37 (13), 3289–3303. doi:10.19595/j.cnki.1000-6753.tces.210712

Li, Z., Wang, W., Han, S., et al. (2022a). Study on voltage adaptation of distributed photovoltaic access to distribution network considering reactive power support. *Power Syst. Prot. Control* 50 (11), 32–41. doi:10.19783/j.cnki.pspc.211781

Liu, R., Sheng, W., Ma, X., et al. (2021). Research on the planning of distributed photovoltaic access to distribution networks based on multiple swarm genetic algorithms. *J. Sol. Energy* 42 (6), 146–155. doi:10.19912/j.0254-0096.tynxb.2019-0370

Liu, Z., Zhang, T., and Wang, Y. (2022). Model-based predictive control for multi-scenario variable time-scale optimal scheduling of active distribution network. *Power Autom. Equip.* 42 (4), 121–128. doi:10.16081/j.epae.202201001

Mukherjee, B., and Sossan, F. (2023). Optimal Planning of single-port and multi-port charging stations for electric vehicles in medium voltage distribution networks. *IEEE Trans. Smart Grid* 14 (2), 1135–1147. doi:10.1109/tsg.2022.3204150

Pan, M., Liu, N., and Lei, J. (2021). A dynamic classification method for distributed energy clusters containing cogeneration units. *Power Syst. Autom.* 45 (1), 168–176. doi:10.7500/AEPS20200217013

Sasaki, Y., Fukuba, S., Yokota, H., Yorino, N., Fukuba, S., et al. (2023). Comprehensive solution method considering load distribution for determining low-voltage network configurations. *IEEE Trans. Electr. Electron. Eng.* 18 (6), 1023–1032. doi:10.1002/tee.23808

Shaker, M., Farzin, H., and Mashhour, E. (2023). Joint planning of electric vehicle battery swapping stations and distribution grid with centralized charging. *J. Energy Storage* 58 (2), 106455–106455.11. doi:10.1016/j.est.2022.106455

Shen, X., Shahidehpour, M., Zhu, S., Han, Y., and Zheng, J. (2018). Multi-stage planning of active distribution networks considering the co-optimization of operation strategies. *IEEE Trans. Smart Grid* 9 (2), 1425–1433. doi:10.1109/tsg.2016.2591586

Su, S., Lei, J., Yan, Y., Pan, S., Yang, Y., et al. (2023). Voltage regulation strategy of distribution network with decentralized wind power based on cluster partition. *Recent Adv. Electr. Electron. Eng.* 16 (1), 30–44. doi:10.2174/2352096515666220902125455

Wang, L., Zhang, F., Kou, L., et al. (2021). Scaled distributed PV power cluster classification based on Fast Unfolding clustering algorithm. *J. Sol. Energy* 42 (10), 29–34. doi:10.19912/j.0254-0096.tynxb.2018-0896

Wang, Y., Yang, Y., and Xu, Q. (2023b). Integrated planning of natural gas and electricity distribution systems for enhancing resilience. *Int. J. Electr. Power Energy Syst.* 151 (9), 109103–109103.11. doi:10.1016/j.ijepes.2023.109103

Wang, Z., Tan, W., Li, H., Ge, J., and Wang, W. (2023a). A voltage coordination control strategy based on the reactive power-active network loss partitioned aggregation domain. *Int. J. Electr. Power Energy Syst.* 144 (1), 108585–108585.10. doi:10.1016/j.ijepes.2022.108585

Wu, H., Sun, L., Xiang, S., et al. (2022a). Active distribution network expansion planning with high-dimensional time-series correlation of renewable energy and load. *Power Syst. Autom.* 46 (16), 40–51. doi:10.7500/AEPS20220104004

Wu, L., Xin, J., and Wang, C. (2022b). Installed capacity forecasting of distributed photovoltaic taking into account user herding psychology. *Power Syst. Autom.* 46 (14), 83–92. doi:10.7500/AEPS20210902006

Xiao, C., Zhao, B., Zhou, J., et al. (2017). Voltage control of high proportional distributed PV clusters in distribution networks based on network partitioning. *Power Syst. Autom.* 41 (21), 147–155. doi:10.7500/AEPS20170101002

Xu, L., Yang, J., and Xiong, Z. (2021a). A method of power distribution network classification based on improved grey cluster. *Electr. Eng.* (13), 81–84. doi:10.19768/j.cnki.dgjs.2021.13.021

Xu, X., Zheng, X., Wang, S., et al. (2021b). Coordinated planning method for transmission and distribution networks based on improved genetic annealing algorithm. *Power Syst. Prot. Control* 49 (15), 1 24–131. doi:10.19783/j.cnki.pspc.201236

Yang, J., Hao, J., and Bo, Z. (2017). Hierarchical control strategy for reactive voltage of wind farm clusters based on adjacency empirical particle swarm algorithm. *Power Grid Technol.* 41 (6), 1823–1829. doi:10.13335/j.1000-3673.pst.2016.2237

Zdraveski, V., Vuletic, J., Angelov, J., and Todorovski, M. (2023). Radial distribution network planning under uncertainty by implementing robust optimization. *Int. J. Electr. Power Energy Syst.* 149 (7), 109043–109043.12. doi:10.1016/j.ijepes.2023.109043

Zhang, T., Xie, M., Wang, P., et al. (2021). System dynamics simulation of the shared value of distributed photovoltaic and its impact on the distribution network. *Power Syst. Autom.* 45 (18), 35–44. doi:10.7500/AEPS20200901004

Zhang, Y., Li, C., Wan, H., Shi, Q., Liu, W., and Xu, Y. (2023). Collaborative stochastic expansion planning of cyber-physical system considering extreme scenarios. *IET Generation, Transm. Distribution* 17 (10), 2419–2434. doi:10.1049/gtd2.12819

Zhu, J., Gu, W., Zhang, H., et al. (2018). An optimization method for siting and capacity determination of distributed power supplies considering dynamic reconfiguration of networks. *Power Syst. Autom.* 42 (5), 111–119. doi:10.7500/APES20170605016