scientific reports

OPEN

Adaptive data driven multi period power supply recovery method for distribution networks

Xi Ye[1,2✉], Meng Yang[3], Zhihong Yang[1], Qing Xiang[1] & Yazhuo Li[1]

In the process of distribution network fault recovery, in order to better address the issues caused by the uncertainty of new energy power output, this paper proposes a multi period power supply recovery method for distribution networks based on adaptive data driven approach. Firstly, this method uses the historical data of new energy power output to construct an ellipsoidal uncertainty set, and forms a data driven convex hull polyhedral set by connecting the vertices of the high-dimensional ellipsoid. Then, aiming at the problem of large conservatism during the reduction process of the convex hull polyhedral set, based on the range of the box set, cutting planes are made starting from the vertices, and a data driven hyperplane polyhedral set model is constructed. Furthermore, considering the constraints of cyber-physical integration, an adaptive data driven power supply recovery model for distribution networks is established, and the column and constraint generation (C&CG) algorithm is adopted to solve the robust scheduling model. Finally, simulations on the improved IEEE-33 bus system and actual example systems verify that the adaptive data driven power supply recovery model for distribution networks can reduce conservatism and improve the robustness of the optimization results. This power supply recovery model can reduce conservatism and enhance the robustness of the optimization results. In the actual distribution network fault recovery scenarios, it can make more efficient use of new energy and ensure the stability of power supply, which strongly demonstrates the effectiveness and application value of the proposed method.

In recent years, due to the frequent occurrence of natural disasters and human-induced disruptions, large-scale blackout incidents have posed significant challenges to the safe operation of distribution networks in various countries. To address the crisis of blackouts under different circumstances, countries have begun to establish resilient grids, and enhancing the load recovery capability of power systems is a means to improve grid resilience[1].

With the development of new power systems, the penetration rate of new energy represented by photovoltaic (PV) has gradually increased, and it has become possible to provide short-term support in the power supply recovery of distribution networks[2]. However, compared with traditional fossil fuel power generation, new energy sources are characterized by volatility and randomness, bringing unpredictable disturbance risks to grid operation. Compared with traditional thermal power generation, the output of new energy power generation is highly volatile and difficult to predict, which poses new challenges to the traditional fault recovery methods of power systems[3,4].

At the same time, due to the rapid development of power communication technology in recent years, information and communication technology greatly promotes the efficient operation of power systems, affecting various aspects of power system operations[5]. With the advancement of communication technology, the rapid recovery capability of distribution networks has been greatly improved. The close integration of information systems and power systems is gradually evolving modern power systems into highly intelligent information-physical power systems. This close integration not only enhances the real-time perception and precise control capabilities of distribution networks but also makes the grid more complex and diverse. In the process of power supply recovery in distribution networks, the normal operation of terminal communication equipment corresponding to power nodes plays a critical role[6,7]. Therefore, in the above context, considering the factors

---

[1]Institute of Intelligent Manufacturing, Jianghan University, Wuhan, Hubei, China. [2]Hubei Engineering Research Center for Intelligent Detection and Control of Specialized Equipment, Wuhan, China. [3]Economics and Technology Research Institute, State Grid Hubei Electric Power Company, Wuhan 430077, China. ✉email: 1208920611@qq.com

related to the operation of information networks and utilizing new energy sources and other renewable energy for power system recovery is of great significance.

Power supply recovery of distribution networks can be achieved by restoring the isolated areas resulting from power outages[8]. increased the load recovery of distribution networks by enhancing the utilization rate of distributed generators (DG) after detecting the isolated areas[9]. prioritized power supply recovery for critical loads using dynamic scheduling of DG islands and controllable load alliances, and simplifies the solution process with feeder equivalence and fault equivalence methods[10]. proposed constraint simplification methods suitable for the distributed power supply recovery network topology search process after determining the island operation areas. For areas isolated due to faults in the distribution system, soft open point (SOP) technology can effectively provide voltage support for the power outage areas[11,12] established SOP planning and operation models to further enhance power supply recovery levels after analyzing the characteristics and functions of SOP.

In order to cope with the uncertainty of distributed PV output during the power supply recovery process, there are two main types of uncertainty optimization methods for distribution network scheduling: stochastic optimization methods and robust optimization methods[13,14]. Robust optimization methods usually use an ensemble form to describe the range of distribution of uncertain parameters, which is getting more and more attention as it does not need to obtain the probability distribution of uncertain parameters compared to stochastic methods and avoids the high-dimensional problem introduced due to a large number of scenarios.

However, different sets have different impacts on the optimization results of distribution network power supply recovery. Therefore, selecting an appropriate set not only reduces the conservativeness of the optimization results but also ensures the robustness of the results. In recent years, some scholars have proposed constructing uncertain sets based on historical data by considering the correlation between uncertain parameters and their random changes over time. In addition to constructing sets using polyhedral sets and ellipsoidal sets, one common method is to construct sets based on extreme scenarios[15–18]. This method first performs envelope processing on the historical data of uncertain parameters, then selects extreme scenarios from the historical data, and constructs sets that include these scenarios. By introducing these extreme scenarios into the optimization model, the model is more robust and includes all possible historical data[19,20] presented an example of constructing such sets. This method addressed the conservativeness of polyhedral sets but still faces the problem of overly conservative results when dealing with extreme scenarios[21,22] proposed a data driven method to improve the conservativeness of polyhedral sets, but constructing sets based on extreme scenarios may still face the problem of computational complexity.

Most of the current research on power supply recovery is limited to topological research and lacks consideration of new uncertainties brought by new energy sources. Therefore, this paper proposes a multi period power supply recovery method of distribution network based on adaptive data drive to address the shortcomings of the above set. This paper first proposes a convex polyhedral set constructed based on historical scenario envelopes, then constructs a convex polyhedral set from the connection vertices of the historical scenario envelopes, and finally addresses the problem of the conservativeness of the convex polyhedral set through contraction to include all historical scenarios. Using the concept of super-planes, this paper constructs a data driven hyperplane polyhedral set. At the same time, considering the establishment of an information-physical model with information nodes, a distribution network power supply recovery model based on the information-physical model is constructed, and the effectiveness of the proposed method is verified through simulations.

Data driven modeling of uncertain sets
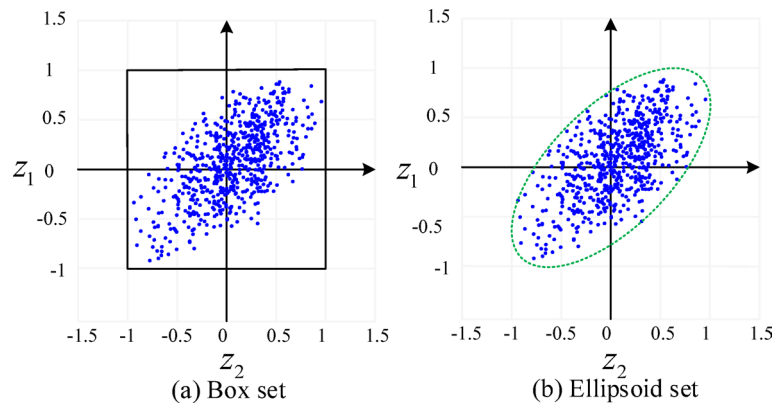
(1) Convex hull polyhedral set.

Traditional uncertainty modeling methods mainly include the probability distribution method and the scenario analysis method. The probability distribution method assumes that the output of new energy sources follows a certain known probability distribution. However, the actual output of new energy sources is affected by a variety of complex factors, and its distribution often does not fully conform to these assumptions. The ellipsoidal uncertainty set and the convex hull polyhedral set are constructed based on actual historical data, which can more accurately reflect the true uncertainty of the new energy output and avoid the modeling bias caused by the wrong assumption of the probability distribution. The scenario analysis method usually simulates the uncertainty by selecting a limited number of representative scenarios. It is very difficult for this method to comprehensively cover all possible situations, and it is easy to miss some extreme but possible scenarios. The convex hull polyhedral set can cover all possible states of new energy output based on historical data, has a wider coverage of scenarios, and can deal with uncertainties more comprehensively.

In order to adapt to the real-time changes in the power output of distributed PV, this paper, based on the power quality detection devices installed at the nodes of the distribution network, captures the real-time signals triggered by sudden weather changes and unexpected power grid faults. By establishing a dedicated data transmission and processing channel, the historical data of the power output of distributed PV under various weather conditions are obtained. Then, according to the scatter plot formed by the historical data of the power output of distributed PV, different envelopes can be used to represent different sets, namely the box set and the ellipsoidal set, as shown in Fig. 1. For different sets, this paper uses the uncertain set $\mathbf{U}$ to represent the fluctuation range of distributed PV output.
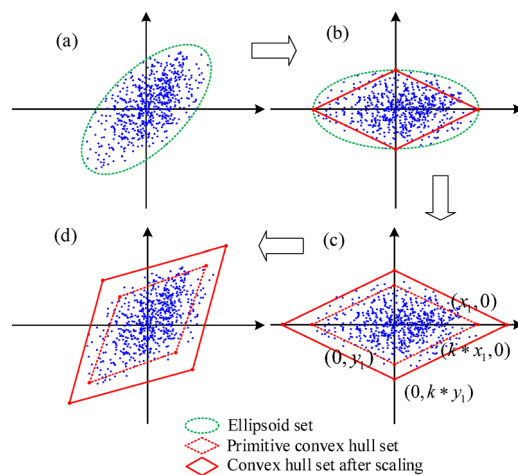
1) Box set.

The specific expression is:

$$\mathbf{U} = \left\{ \mathbf{z}^{\mathrm{PV}} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} | \beta \mathbf{z}_{\mathrm{down}}^{\mathrm{PV}} \leqslant \mathbf{z}^{\mathrm{PV}} \leqslant \beta \mathbf{z}_{\mathrm{up}}^{\mathrm{PV}} \right\} \tag{1}$$

Fig. 1. Uncertain set.

Fig. 2. Modeling process of convex hull polyhedron uncertainty set.

where N_{PV} represents the number of PV connections; $\mathbf{z}^{\mathrm{PV}}$ is the uncertain variable of PV output, $\mathbf{z}^{\mathrm{PV}}_{\mathrm{up}}$ and $\mathbf{z}^{\mathrm{PV}}_{\mathrm{down}}$ respectively represent the upper and lower bounds of the uncertain variable of PV output. β represents the tuning parameter used to adjust the retention degree of the box set within $(0, 1]$.

As shown in Fig. 1 (a), the box set covers all possible values of distributed PV output. However, at different times and locations, the output of distributed PV tends to follow a certain trend and shows spatial correlations. Therefore, PV output data often lie along the y = x and y=-x diagonals. If the uncertain set constructed by the box set fails to reflect the temporal and spatial correlation of distributed PV output, it may result in overly conservative predictions. This is because the box set not only covers the potential fluctuations of PV output but also includes a large number of empty regions where fluctuations are unlikely to occur. Therefore, a more accurate construction method of the uncertain set is required.

2) Ellipsoid set.

The specific expression is:

$$\mathbf{U} = \left\{ \mathbf{z}^{\mathrm{PV}} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} \middle| \left(\mathbf{z}^{\mathrm{PV}} - \mathbf{c} \right)^{\mathrm{T}} \Sigma^{-1} \left(\mathbf{z}^{\mathrm{PV}} - \mathbf{c} \right) \leqslant 1 \right\} \tag{2}$$

where $\mathbf{c}$ denotes the centroid of the high-dimensional ellipsoid, and $\Sigma \in \mathbf{R}^{N_{\mathrm{PV}} \times N_{\mathrm{PV}}}$ is a positive definite matrix that indicates the direction of the axis shift of the high-dimensional ellipsoid.

As shown in Fig. 1 (b), the ellipsoidal set, like the box set, covers all possible values of distributed PV output. At the same time, unlike the box set, the ellipsoidal set reduces the inclusion of empty regions where fluctuations are unlikely to occur, thereby reducing the conservativeness of the decision results. However, because the expression of the ellipsoid set is quadratic, it increases the difficulty of solving the problem in the optimization process.

Based on this[19], proposed a generalized convex hull set that can significantly reduce the conservativeness of the optimization results while maintaining good robustness, and avoids introducing quadratic expressions in the modeling process. Therefore, this paper constructs a data driven uncertain set based on the method in[19], with the modeling process shown in Fig. 2.

Step (1): First, construct a high-dimensional ellipsoid $\mathbf{U}_{e1}$ that covers all historical data fluctuations and has the smallest volume. The constructed high-dimensional ellipsoid is shown in Fig. 2 (a), and its specific expression is as follows (1):

$$\mathbf{U}_{e1} = \left\{ \mathbf{z}^{\mathrm{PV}} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} | (\mathbf{z}^{\mathrm{PV}} - \mathbf{c})^{\mathrm{T}} \Sigma^{-1} (\mathbf{z}^{\mathrm{PV}} - \mathbf{c}) \leqslant 1 \right\} \qquad (3)$$

Step (2): On the basis of the existing high-dimensional ellipsoid, perform eigenvalue decomposition on the positive definite matrix Σ:$\Sigma = \mathbf{P^T J P} = \mathbf{P}^{-1} \mathbf{J P}$. This rotates and translates the ellipsoid so that its center coincides with the origin of the coordinate system, as shown by the green dashed line in Fig. 2 (b). At this point, the uncertain set of the high-dimensional ellipsoid is denoted as $\mathbf{U}_{e2}$, i.e.,

$$\mathbf{U}_{e2} = \left\{ \mathbf{z}^{\mathrm{PV}'} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} | (\mathbf{z}^{\mathrm{PV}'})^{\mathrm{T}} \mathbf{J}^{-1} (\mathbf{z}^{\mathrm{PV}'}) \leqslant 1 \right\} \qquad (4)$$

$$\mathbf{z}^{\mathrm{PV}'} = \mathbf{P} \times (\mathbf{z}^{\mathrm{PV}} - \mathbf{c}) \qquad (5)$$

where $\mathbf{J}$ is a diagonal matrix, denoted as $\mathbf{J} = \mathrm{diag}(\lambda_1 \ldots \lambda_{N_{\mathrm{PV}}})$, and P is a transformation matrix representing the rotation angles of the matrix. According to the diagonal matrix $\mathbf{J}$, the coordinates of the vertex of the transformed high-dimensional ellipsoid are:

$$\begin{cases} z_{c,1}^{\mathrm{PV}'} = [1/\sqrt{\lambda_1}, 0 \ldots 0], z_{c,N_{\mathrm{PV}}+1}^{\mathrm{PV}'} = -[1/\sqrt{\lambda_1}, 0 \ldots 0] \\ z_{c,2}^{\mathrm{PV}'} = [0, 1/\sqrt{\lambda_2} \ldots 0], z_{c,N_{\mathrm{PV}}+2}^{\mathrm{PV}'} = -[0, 1/\sqrt{\lambda_2} \ldots 0] \\ \vdots \\ z_{c,N_{\mathrm{PV}}}^{\mathrm{PV}'} = [0, 0 \ldots 1/\sqrt{\lambda_{N_{\mathrm{PV}}}}], z_{c,2N_{\mathrm{PV}}}^{\mathrm{PV}'} = -[0, 0 \ldots 1/\sqrt{\lambda_{N_{\mathrm{PV}}}}] \end{cases} \qquad (6)$$

Next, connect each vertex of the high-dimensional ellipsoid to form a high-dimensional polyhedron, as shown by the red dashed line in Fig. 2 (b). At this point, the uncertain set of the high-dimensional polyhedron is denoted as $\mathbf{U}_{p2}$, which is:

$$\mathbf{U}_{p2} = \left\{ \mathbf{z}^{\mathrm{PV}'} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} \left| \begin{array}{c} \mathbf{z}^{\mathrm{PV}'} = \sum_{i=1}^{2N_{\mathrm{PV}}} m_i \mathbf{z}_{c,i}^{\mathrm{PV}'} \\ \sum_{i=1}^{2N_{\mathrm{PV}}} m_i = 1;\ 0 \leqslant m_i \leqslant 1 \end{array} \right. \right\} \qquad (7)$$

where m_i denotes the weight coefficient of the first vertex.

Step (3): Since the high-dimensional polyhedral set obtained in step 2 causes some data points to lie outside the boundary, it is necessary to perform a contraction process on the set. As shown by the red solid line in Fig. 2 (c), the vertices of the contracted high-dimensional polyhedral set are:

$$\begin{cases} kz_{c,1}^{\mathrm{PV}'} = [k/\sqrt{\lambda_1}, 0 \ldots 0], kz_{c,N_{\mathrm{PV}}+1}^{\mathrm{PV}'} = -[k/\sqrt{\lambda_1}, 0 \ldots 0] \\ kz_{c,2}^{\mathrm{PV}'} = [0, k/\sqrt{\lambda_2} \ldots 0], kz_{c,N_{\mathrm{PV}}+2}^{\mathrm{PV}'} = -[0, k/\sqrt{\lambda_2} \ldots 0] \\ \vdots \\ kz_{c,N_{\mathrm{PV}}}^{\mathrm{PV}'} = [0, 0 \ldots k/\sqrt{\lambda_{N_{\mathrm{PV}}}}], kz_{c,2N_{\mathrm{PV}}}^{\mathrm{PV}'} = -[0, 0 \ldots k/\sqrt{\lambda_{N_{\mathrm{PV}}}}] \end{cases} \qquad (8)$$

At this time, the high-dimensional linear polyhedron after scaling is shown in Fig. 2 (d), and its uncertainty set $\mathbf{U}_{p2}$ is
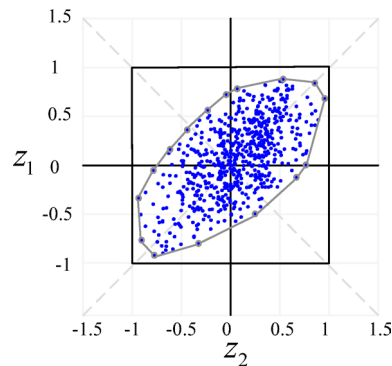
$$\mathbf{U}_{p2} = \left\{ \mathbf{z}^{\mathrm{PV}'} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} \left| \begin{array}{c} \mathbf{z}^{\mathrm{PV}'} = \sum_{i=1}^{2N_{\mathrm{PV}}} m_i k \mathbf{z}_{c,i}^{\mathrm{PV}'} \\ \sum_{i=1}^{2N_{\mathrm{PV}}} m_i = 1;\ 0 \leqslant m_i \leqslant 1 \end{array} \right. \right\} \qquad (9)$$

where k is the contraction multiple used to adjust the retention degree of the high-dimensional polyhedral set. The calculation method for k can be found in[19]. Therefore, there exists a minimum k_{min} that allows the contracted polyhedral set to encompass all possible data points. The value range of k is $[0, k_{\mathrm{min}}]$. The polyhedral sets formed by different k values are shown in Fig. 3.

Step (4): Rotate and translate the contracted high-dimensional polyhedral set to encompass the original data points. From (5), it is known that after the rotation and translation, the uncertain set of the high-dimensional polyhedral set is $\mathbf{U}_{p1}$, which is :
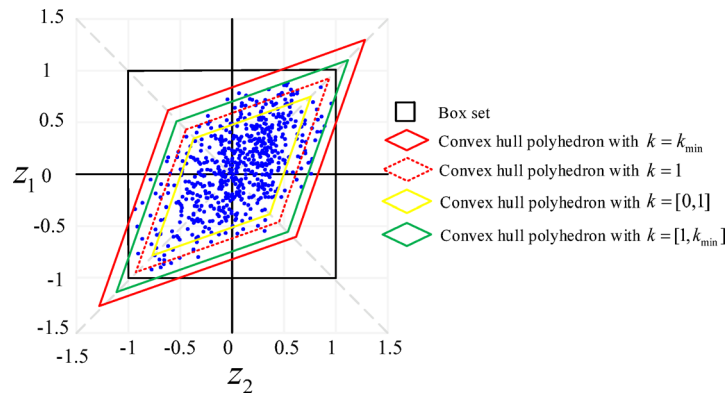
Fig. 3. Range of convex hull polyhedron set under different k.

Fig. 4. Uncertainty set based on extreme scenarios.

$$\mathbf{U}_{p1} = \left\{ \mathbf{z}^{\mathrm{PV}} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} \,\middle|\, \begin{array}{c} \mathbf{z}^{\mathrm{PV}} = \sum_{i=1}^{2N_{\mathrm{PV}}} m_i (\mathbf{c} + k\mathbf{P}^{-1}\mathbf{z}_{c,i}^{\mathrm{PV}\prime}) \\ \sum_{i=1}^{2N_{\mathrm{PV}}} m_i = 1; \ 0 \leqslant m_i \leqslant 1 \end{array} \right\} \tag{10}$$

(2) Hyperplane polyhedral set.

The polyhedral set described in the previous section is constructed by first establishing a high-dimensional polyhedral set based on the vertices of an ellipsoid, and then performing successive contractions to obtain an uncertain high-dimensional polyhedral set. This set encompasses all historical data points. However, if the contraction is uniformly applied, it might lead to some new data points being located outside the boundary due to excessive enlargement of the bounding region. Consequently, the corresponding bounding regions will include more empty space. Therefore[21,22], proposed constructing the uncertain set incrementally based on the positions of the boundary points, forming an irregular polyhedral set through iterative extreme value envelopes, as shown in Fig. 4.

The uncertain set based on extreme scenarios is expressed as:

$$\mathbf{U} = \left\{ \mathbf{z}^{\mathrm{PV}} \in \mathbf{R}^{N_{\mathrm{PV}} \times 1} \,\middle|\, \begin{array}{c} \mathbf{z}^{\mathrm{PV}} = \sum_{i=1}^{N_{ex}} \sigma_i \mathbf{z}_i^{\mathrm{PV}} \\ \sum_{i=1}^{N_{ex}} \sigma_i = 1; \ 0 \leqslant \sigma_i \leqslant 1 \end{array} \right\} \tag{11}$$

where N_{ex} represents the number of extreme scenarios. $\mathbf{z}_i^{\mathrm{PV}}$ represents the ith extreme scenario. σ_i represents the weight coefficient of the extreme scenario.

Compared to the uncertain set of the convex hull, the uncertain set based on extreme scenarios can significantly reduce the inclusion of empty regions with low probability distributions. Therefore, this method maintains the best fidelity. However, as shown in (11), the construction of the uncertain set based on extreme scenarios depends on the number of extreme scenarios, i.e., the number of vertex points. If the number of extreme scenarios is large, it will increase the computational complexity of the optimization process. Therefore,

this paper proposes a hyperplane polyhedral set, which cuts and flattens the boundary from the vertices based on the original box set range, significantly reducing empty regions, as shown in Fig. 5.

In general, for a K-dimensional space, the hyperplane is generally expressed as:

$$\alpha_m^T \mathbf{z} = \beta_m, \forall m \tag{12}$$

where α_m is a non-zero K-dimensional nonzero vector. β_m is a scalar. m is the vertex index of the K-dimensional box set. Let $\mathbf{z}_m^B$ be the vertex of the K-dimensional box set. The vertex produced by this hyperplane cut is $\mathbf{z}_n^H$. Then the relationship between the vertex indices m, n is:

$$n = (m-1) \times K + s, \ s = 1, 2 \ldots K, \forall m = 1, 2 \ldots 2^K \tag{13}$$

At this point, the vertices produced by the hyperplane cut can be obtained by solving the following model:

$$\forall m, \quad \max_{\alpha_m, \beta_m, \xi_{ms}, \mathbf{z}_n^H} \frac{1}{K!} \prod_{s=1}^{K} \xi_{ms} \tag{14}$$

$$\alpha_m^T \mathbf{z}_m^B \geqslant \beta_m, \alpha_m^T \tau \leqslant \beta_m, \alpha_m^T \mathbf{z}_n^H = \beta_m, \forall s, \forall \tau \in T \tag{15}$$

$$z_{mo}^B - z_{no}^H = 0, \forall s, o = 1, \ldots s-1, s+1, \ldots K \tag{16}$$

$$\xi_{ms} = \theta_{ms}(z_{ms}^B - z_{ns}^H), \forall s \tag{17}$$

$$\xi_{m_1 s} + \xi_{m_2 s} \leqslant z_{m_2 s}^B - z_{m_1 s}^B, \forall s, \forall h \in H(s), m_1, m_2 \in h \tag{18}$$

In (14), ξ_{ms} represents the Euclidean distance between $\mathbf{z}_n^H$ and the corresponding $\mathbf{z}_m^B$. Since (14) is a nonlinear equation, to reduce computational pressure without affecting the decision variables to be solved, it is transformed into the form of (14-a):

$$\forall m, \quad \max_{\alpha_m, \beta_m, \xi_{ms}, \mathbf{z}_n^H} \sum_{s=1}^{K} \ln \xi_{ms} \tag{14-a}$$

(15) divides the K-dimensional space into inside and outside regions, representing the vector direction of the $\mathbf{z}_m^B$ of the box set, with the data vector τ inside, and T representing the data set. (16) indicates that, except for the first $o = s$-dimension, other dimensions have coordinates with a standard deviation of 0. (17) shows the calculation formula for the Euclidean distance ξ_{ms}, where θ_{ms} takes values of 1 or −1, depending on whether $\mathbf{z}_{ms}^B$ in the s-dimension is greater or less than $\mathbf{z}_{ns}^H$. (18) indicates that any hyperplane does not intersect with the K-dimensional box set. h and $H(s)$ represent the order and collection of intervals of the s-dimensional uncertain set of the box, while m_1 and m_2 represent the boundaries of the regions.

The above model is a nonlinear model, so the interior point method is used for solving. After determining the coordinates of the vertices of the hyperplane and combining with (11), the uncertain set of the hyperplane is expressed as:

$$\mathbf{U} = \left\{ \mathbf{z}^{H,PV} \in \mathbf{R}^{N_{PV} \times 1} \ \middle| \ \begin{array}{l} \mathbf{z}^{H,PV} = \sum_{n=1}^{N_H} \varepsilon_i \mathbf{z}_n^{H,PV} \\ \sum_{n=1}^{N_H} \varepsilon_i = 1; \ 0 \leqslant \varepsilon_i \leqslant 1 \end{array} \right\} \tag{15}$$

where ε_i denotes the weighting coefficient of the ith vertex.
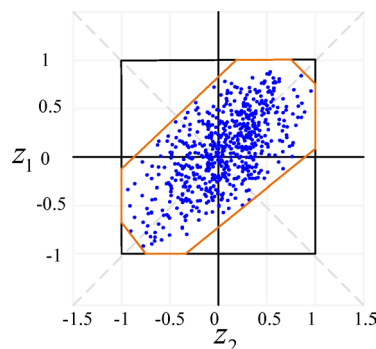
Fig. 5. Hyperplane polyhedron set.

Multi period power supply recovery modelling for distribution networks

1) Objective function.

This paper constructs a distribution network power supply recovery model that maximizes load recovery and minimizes network loss. The model is formulated as:

$$\max C = \sum_{t=1}^{T} \left(\frac{\sum_{i \in D} r_i \lambda_{i,t} P_{i,t}^{\text{INF}}}{\sum_{i \in D} r_i P_{i,t}^{\text{INF}}} \right) + \sum_{t=1}^{T} \sum_{ij \in L} c_{\text{loss}} (P_{ij,t} + P_{ji,t}) \quad (16)$$

where $\lambda_{i,t}$ represents the power supply recovery status of node i in time period t, where 0 indicates no recovery and 1 indicates full recovery. r_i represents the importance weight of node i. This paper divides the load into three levels[23], with corresponding values of 10, 5, and 1 for each level. $P_{i,t}^{\text{INF}}$ denotes the power of load recovery of node i in time period t. $P_{ij,t}$ and $P_{ji,t}$ denote the active power of branch ij in time period t flow. $P_{ji,t}$ denotes the power flow from node i to node j at time t. $P_{ji,t}$ denotes the power flow through node i to node j. C represents the total power supply recovery cost of the distribution network. c_{loss} denotes the power loss coefficient of the distribution network. D represents the set of distribution network nodes. L denotes the set of all paths in the distribution network.

2) Constraint condition.

1) Second-order cone relaxation power flow constraints.

The second-order cone relaxation power flow model under the polar coordinate system can be expressed as:

$$P_{i,t}^{\text{sum}} - \lambda_{i,t} P_{i,t}^{\text{INF}} = \sum_{j \in N(i)} P_{ij,t} \quad (17)$$

$$Q_{i,t}^{\text{sum}} - \lambda_{i,t} Q_{i,t}^{\text{INF}} = \sum_{j \in N(i)} Q_{ij,t} \quad (18)$$

$$P_{i,t}^{\text{sum}} = P_{i,t} + P_{i,t}^{\text{PV}} - P_{i,t}^{\text{ch}} + P_{i,t}^{\text{dis}} + P_{i,t}^{\text{SOP}} \quad (19)$$

$$Q_{i,t}^{\text{sum}} = Q_{i,t} + Q_{i,t}^{\text{SOP}} \quad (20)$$

$$P_{ij,t} = \sqrt{2} g_l u_{i,t}^l - g_l R_{l,t} - b_l T_{l,t} \quad (21)$$

$$P_{ji,t} = \sqrt{2} g_l u_{j,t}^l - g_l R_{l,t} + b_l T_{l,t} \quad (22)$$

$$Q_{ij,t} = -\sqrt{2} b_l u_{i,t}^l + b_l R_{l,t} - g_l T_{l,t} \quad (23)$$

$$Q_{ji,t} = -\sqrt{2} b_l u_{j,t}^l + b_l R_{l,t} + g_l T_{l,t} \quad (24)$$

$$I_l^2 = \sqrt{2} (g_l^2 + b_l^2)(u_{i,t}^l + u_{j,t}^l - \sqrt{2} R_{l,t}) \quad (25)$$

$$\left\| \begin{array}{c} \sqrt{2} R_{l,t} \\ \sqrt{2} T_{l,t} \\ u_{i,t}^l - u_{j,t}^l \end{array} \right\| \leqslant u_{i,t}^l + u_{j,t}^l \quad (26)$$

where $P_{i,t}^{\text{sum}}$ and $Q_{i,t}^{\text{sum}}$ denote the active and reactive power injected by the node i at the t time, respectively. $Q_{i,t}^{\text{D}}$ denotes the reactive power load of node i at time t. $Q_{ij,t}$ denotes the reactive power flowing from node i to node j at time t of branch ij in the distribution network. $P_{i,t}$ and $Q_{i,t}$ denote the gateway power injected by node i at time t, respectively. $P_{i,t}^{\text{SOP}}$ and $Q_{i,t}^{\text{SOP}}$ represent the active and reactive power of the SOP at node i at time t. $N(i)$ denotes the set of all nodes connected to node i. g_l and b_l denote the conductance of branch l and the conductance to ground. I_l indicates the current amplitude of branch l. Also introduce $u_{i,t} = V_{i,t}^2 / \sqrt{2}$,

$R_{l,t} = V_{i,t} V_{j,t} \cos\theta_{ij,t}$, $T_{l,t} = V_{i,t} V_{j,t} \sin\theta_{ij,t}$, where $V_{i,t}$ and $V_{j,t}$ denote the voltage amplitude of node i and j at time t, respectively. $\theta_{ij,t}$ represents the voltage phase angle difference between both ends of the line ij at time t. Additionally, since $u_{i,t}$ does not include the branch switch variable $\alpha_{l,t}$, a new variable $u_{i,t}^l$ that includes the branch switch variable is introduced to jointly determine the branch status and the voltage at the branch end nodes.

When the branch l is open at time t both $u_{i,t}^l$ and $u_{j,t}^l$ are 0. When the branch l is closed at time t, $u_{i,t}^l$ and $u_{j,t}^l$ are equal to the corresponding node voltages $u_{i,t}$ and $u_{j,t}$. To express the relationship between $u_{i,t}^l$, $u_{j,t}^l$, and $u_{i,t}$, $u_{j,t}$, additional constraints are added, i.e.,

$$\begin{cases} 0 \leqslant u_{i,t}^l \leqslant \dfrac{(V_i^{\max})^2}{\sqrt{2}} \alpha_{l,t} \\[2mm] 0 \leqslant u_{j,t}^l \leqslant \dfrac{(V_j^{\max})^2}{\sqrt{2}} \alpha_{l,t} \\[2mm] 0 \leqslant u_{i,t} - u_{i,t}^l \leqslant \dfrac{(V_i^{\max})^2}{\sqrt{2}} (1 - \alpha_{l,t}) \\[2mm] 0 \leqslant u_{j,t} - u_{j,t}^l \leqslant \dfrac{(V_j^{\max})^2}{\sqrt{2}} (1 - \alpha_{l,t}) \end{cases} \tag{27}$$

where $V_i^{\max}$ and $V_j^{\max}$ represent the maximum value of voltage amplitude at nodes i and j, respectively. Additionally, the constraints for node voltages and branch currents are respectively given by:

$$\frac{(V_i^{\min})^2}{\sqrt{2}} \leqslant u_{i,t} \leqslant \frac{(V_i^{\max})^2}{\sqrt{2}} \tag{28}$$

$$u_{i,t}^l + u_{j,t}^l - \sqrt{2} R_{l,t} \leqslant \frac{(I_l^{\max})^2}{\sqrt{2}(g_l^2 + b_l^2)} \tag{29}$$

where $I_l^{\max}$ represents the maximum value of current amplitude in branch l.

2) Distributed PV constraints

$$\tilde{P}_{i,t}^{\mathrm{PV}} = P_{i,t}^{\mathrm{PV,f}} + \Delta P^{\mathrm{PV,max}} z_{i,t}^{\mathrm{PV}} \tag{30}$$

$$0 \leqslant P_{i,t}^{\mathrm{PV}} \leqslant \tilde{P}_{i,t}^{\mathrm{PV}}, \forall i \in \Omega_N^{\mathrm{PV}} \tag{31}$$

$$(P_{i,t}^{\mathrm{PV}})^2 + (Q_{i,t}^{\mathrm{PV}})^2 \leqslant (S_{i,t}^{\mathrm{PV}})^2, \forall i \in \Omega_N^{\mathrm{PV}} \tag{32}$$

(30)-(32) represent the operational constraints of distributed PV, where $P_{i,t}^{\mathrm{PV,f}}$ represents the maximum available power at node x connected to distributed PV before the fluctuation at time t. $\Delta P^{\mathrm{PV,max}}$ denotes the maximum fluctuation of the distributed PV. $Q_{i,t}^{\mathrm{PV}}$ and $S_{i,t}^{\mathrm{PV}}$ denote the reactive power and capacity of the connected distributed PV at node i at time t.

3) Battery energy storage constraints

$$S_{i,t}^{\mathrm{SOC}} = S_{i,t-1}^{\mathrm{SOC}} + \eta_{i,\mathrm{ch}} \frac{P_{i,t}^{\mathrm{ch}} \Delta t}{E_{i,\max}} - \frac{P_{i,t}^{\mathrm{dis}} \Delta t}{\eta_{i,\mathrm{dis}} E_{i,\max}}, \forall t \in H \tag{33}$$

$$\begin{cases} 0 \leqslant P_{i,t}^{\mathrm{ch}} \leqslant P_{i,\max}^{\mathrm{ch}} D_{i,t}^{\mathrm{ch}} \\[1mm] 0 \leqslant P_{i,t}^{\mathrm{dis}} \leqslant P_{i,\max}^{\mathrm{dis}} D_{i,t}^{\mathrm{dis}} \\[1mm] D_{i,t}^{\mathrm{ch}} + D_{i,t}^{\mathrm{dis}} \leqslant 1 \end{cases} \tag{34}$$

$$S_{i,\min}^{\mathrm{SOC}} \leqslant S_{i,t}^{\mathrm{SOC}} \leqslant S_{i,\max}^{\mathrm{SOC}} \tag{35}$$

(33)-(35) represent the operational constraints of the battery energy storage. Where $S_{i,t}^{\mathrm{SOC}}$ represents the charging state of the connected battery energy storage at node i at time t. $D_{i,t}^{\mathrm{ch}}$ and $D_{i,t}^{\mathrm{dis}}$ are 0–1 variables, which represent the charging and discharging state of the connected battery energy storage at node i at time t, with 1 indicating charging and 0 indicating discharging. $\eta_{i,\mathrm{ch}}$ and $\eta_{i,\mathrm{dis}}$ represent the charging and discharging efficiencies of the connected battery energy storage at node i, respectively. $E_{i,\max}$ represents the maximum amount of battery storage energy at node i. $S_{i,\min}^{\mathrm{SOC}}$ and $S_{i,\max}^{\mathrm{SOC}}$ represent the minimum and maximum states of charge of the battery at node i, respectively. Δt represents the time interval for the battery state adjustment. H indicates the set of charging and discharging time of the battery energy storage.

4) Load model constraints.

This paper classifies loads into two types: controllable loads and uncontrollable loads. For controllable loads, the distribution network power supply recovery can restore part of the load. For uncontrollable loads, the distribution network power supply recovery must fully restore or not restore them at all. The specific model is as follows:

$$P_{i,t}^{\mathrm{INF}} = k_{i,t} P_{i,t}^{\mathrm{D}} \tag{36}$$

where $k_{i,t}$ denotes the load recovery factor of node i at time t, when the load is controllable, $k_{i,t} \in [0, 1]$. $k_{i,t}$ is equal to 1 when the load is equivalent to an uncontrollable load. $P_{i,t}^{\mathrm{D}}$ denotes the amount of load of node i at time t.

5) Radial topology structure.

This paper uses the ST constraints to describe the radial topology constraint of the distribution network, as shown in (37).

$$\begin{cases} \beta_{ij,t} + \beta_{ji,t} = \alpha_{l,t}, \forall l \in L \\ \sum_{j \in N(i)} \beta_{ij} = 0, \forall i \in N_s \\ \sum_{j \in N(i)} \beta_{ij} = 1, \forall i \in N/N_s \\ 0 \leqslant \alpha_{l,t} \leqslant 1, \beta_{ij,t}, \beta_{ji,t} \in \{0,1\}, \forall l \in L \end{cases} \tag{37}$$

(37) determines whether the first and last nodes i and j of a branch are child nodes or parent nodes by introducing two 0–1 variables $\beta_{ij,t}$ and $\beta_{ji,t}$, ensuring that the generated radial structure connects to the root nodes to form a tree. $\beta_{ij,t} = 1$ indicates that node i is the parent node of node j, and $\beta_{ij,t} = 0$ indicates that it is not. N represents the set of all nodes in the distribution network, and N_s represents the set of root nodes.

6) SOP constraints

$$P_{i,t}^{\mathrm{SOP}} + P_{j,t}^{\mathrm{SOP}} + P_{i,t}^{\mathrm{SOP,L}} + P_{j,t}^{\mathrm{SOP,L}} = 0 \tag{38}$$

$$P_{i,t}^{\mathrm{SOP,L}} = \eta^{\mathrm{SOP}} \sqrt{(P_{i,t}^{\mathrm{SOP}})^2 + (Q_{i,t}^{\mathrm{SOP}})^2} \tag{39}$$

$$P_{j,t}^{\mathrm{SOP,L}} = \eta^{\mathrm{SOP}} \sqrt{(P_{j,t}^{\mathrm{SOP}})^2 + (Q_{j,t}^{\mathrm{SOP}})^2} \tag{40}$$

$$\sqrt{(P_{i,t}^{\mathrm{SOP}})^2 + (Q_{i,t}^{\mathrm{SOP}})^2} \leqslant S_i^{\mathrm{SOP}} \tag{41}$$

$$\sqrt{(P_{j,t}^{\mathrm{SOP}})^2 + (Q_{j,t}^{\mathrm{SOP}})^2} \leqslant S_j^{\mathrm{SOP}} \tag{42}$$

where $P_{i,t}^{\mathrm{SOP,L}}$ and $P_{j,t}^{\mathrm{SOP,L}}$ represent the losses of the SOP. η^{SOP} represents the loss coefficient of the SOP. S_i^{SOP} represents the capacity of the SOP.

3) Information-physical integration framework.

In the physical-information integrated framework constructed in this paper, the physical network is the distribution network, and the information network refers to the network formed by the terminal devices that detect and control the distribution network. There is a one-to-one correspondence between the nodes of the information network and the nodes of the distribution network, meaning each load node corresponds to an information node[24].

1) Energy supply constraints for information nodes.

During the power supply recovery of the distribution network, the information node will only be recovered if the recovered load at the load node is greater than the energy demand of the information node. The specific model is as follows:

$$\varphi_{i',t} = \begin{cases} 1 & \lambda_{i,t} \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{INF}} \geqslant x \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{D}} \\ 0 & \lambda_{i,t} \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{INF}} < x \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{D}} \end{cases} \tag{43}$$

$$\sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{INF}} = \sum_{i \in N_{\mathrm{D}}} k_{i,t} P_{i,t}^{\mathrm{D}} \tag{44}$$

where $\varphi_{i',t}$ is a 0–1 variable indicating the effective situation of energy supply of the information node i', equal to 0 indicates that the energy supply is invalid, and equal to 1 indicates that the energy supply is effective. x indicates the coupling parameter of the information node and the load node, $x \in [0,1]$. The larger x is, the larger the coupling strength is. N_{D} denotes the set of load nodes connected to the information node.

(43) is the information node validity constraint, which indicates that when the load recovery of a load node is greater than the supply demand of the information node, the node is flagged as a valid state and vice versa. (44) is the definition of the load recovery of the load node.

Since (43) is non-linear, it needs to be linearized as shown in the following equation:

$$\lambda_{i,t} \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{INF}} - x \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{D}} \leqslant \varphi_{i',t}(1-x) \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{D}} + \sigma \tag{45}$$

$$\lambda_{i,t} \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{INF}} - x \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{D}} \geqslant (1 - \varphi_{i',t})(-x) \sum_{i \in N_{\mathrm{D}}} P_{i,t}^{\mathrm{D}} \tag{46}$$

where σ is a very small positive number.

2) Information network traffic constraints.

Whether an information node works or not depends not only on the energy supply status of that node, but also on whether the node has communication traffic with the control node, as modelled below:

$$\psi_{i',t} = \begin{cases} 1 & \sum_{j' \in N_{\mathrm{C}}} T_{j'i',t} - \sum_{j' \in N_{\mathrm{C}}} T_{i'j',t} > 0 \\ 0 & \sum_{j' \in N_{\mathrm{C}}} T_{j'i',t} - \sum_{j' \in N_{\mathrm{C}}} T_{i'j',t} = 0 \end{cases} \quad i' \in N_{\mathrm{C,goal}} \tag{47}$$

$$\begin{cases} 0 \leqslant T_{i'j',t} \leqslant \varphi_{i',t} n_{\mathrm{C,goal}} \\ 0 \leqslant T_{i'j',t} \leqslant \varphi_{j',t} n_{\mathrm{C,goal}} \end{cases} \tag{48}$$

$$\begin{cases} \sum_{i' \in N_{\mathrm{C,0}}} T_{i'j',t} \leqslant n_{\mathrm{C,goal}} \\ \sum_{i' \in N_{\mathrm{C,0}}} T_{j'i',t} = 0 \end{cases} \tag{49}$$

$$\sum_{i'j' \in L'} T_{i'j',t} = \sum_{i'j' \in L'} T_{j'i',t} + \psi_{i',t} B_{i'} \quad i' \notin N_{\mathrm{C,0}} \tag{50}$$

where $\psi_{i',t}$ indicates the working state of information node i' in time period t. $T_{i'j',t}$ and $T_{j'i',t}$ indicate that the communication traffic flows from node i' to node j' and from node j' to node i' in time period t, respectively. $n_{\mathrm{C,goal}}$ denotes the number of terminals of the information node. N_{C} denotes the set of information nodes. $N_{\mathrm{C,0}}$ denotes the set of information source nodes. $B_{i'}$ denotes the bandwidth of the node i'.

3) Information node control constraints.

If the information node fails, its corresponding load node will be uncontrollable and its corresponding model is as follows.

$$P_{i,t}^{\mathrm{INF}} \leqslant \psi_{i',t} P_{i,t}^{\mathrm{D}} \tag{51}$$

$$P_{i,t}^{\mathrm{sum}} \leqslant \psi_{i',t} P_{i,\max}^{\mathrm{sum}} \tag{52}$$

(51) indicates that if the information node fails, its corresponding load node will not be recoverable. (52) indicates that the external power supply connected to the load node can be put into operation only when the information node is valid. $P_{i,\max}^{\mathrm{sum}}$ denotes the maximum value of the output power of the external power supply.

Robust power supply recovery optimization method for distribution networks

1) Robust power supply recovery optimization model for distribution network.

Let the power flow constraints variable be vector $\mathbf{P} = \{P_{ij,t}, P_{ji,t}, Q_{ij,t}, Q_{ji,t}, u_{i,t}, R_{l,t}, T_{l,t}, I_{l,t},\}$. Let the PV constraints variable be vector $\mathbf{P}^{\mathrm{PV}} = \{P_{i,t}^{\mathrm{PV}}, Q_{i,t}^{\mathrm{PV}}\}$. Let the battery energy storage operating variable be vector $\mathbf{P}^{\mathrm{ESS}} = \{S_{i,t}^{\mathrm{SOC}}, P_{i,t}^{\mathrm{ch}}, P_{i,t}^{\mathrm{dis}}\}$. Let the load model constraints be vector $\mathbf{P}^{\mathrm{D}} = \{P_{i,t}^{\mathrm{INF}}\}$. Let the radial topology constraints be vector $\mathbf{P}^{\mathrm{switch}} = \{\beta_{ij,t}, \beta_{ji,t}, \alpha_{l,t}\}$. Let the information network constraints be vector $\mathbf{P}^{\mathrm{INF}} = \{\varphi_{i,t}, \lambda_{i,t}, T_{ij,t}, \psi_{i,t}\}$. Let the SOP constraints be vector $\mathbf{P}^{\mathrm{SOP}} = \{P_{i,t}^{\mathrm{SOP}}, P_{i,t}^{\mathrm{SOP,L}}, Q_{i,t}^{\mathrm{SOP}}\}$.

On the basis of establishing the data driven polyhedral set for distributed PV output, this paper establishes a two-stage robust power supply recovery optimization model for distribution networks, whose matrix form is:

$$\min_{x} \left(\max_{u \in \mathbf{U}} \min_{\mathbf{y} \in \Omega(\mathbf{x},\mathbf{u})} \mathbf{c}^{\mathrm{T}} \mathbf{y} \right)$$
$$s.t. \ \mathbf{Ax} \leqslant \mathbf{d} \ (a) \tag{53}$$
$$\mathbf{Gy} \leqslant \mathbf{h} - \mathbf{Ex} - \mathbf{Mu} \ (b)$$
$$\|\mathbf{Ry}\|_2 \leqslant \mathbf{r}^{\mathrm{T}} \mathbf{y} \ (c)$$

where $\mathbf{x}$, $\mathbf{y}$ are decision variables of the model and $\mathbf{u}$ is an uncertainty variable. $\mathbf{x} = \{D_{i,t}^{\mathrm{ch}}, D_{i,t}^{\mathrm{dis}}, \varphi_{i,t}, \psi_{i,t}, \mathbf{P}^{\mathrm{switch}}\}$ is the first stage decision variable. $\mathbf{y} = \{\mathbf{P}, \mathbf{P}^{\mathrm{PV}}, \mathbf{P}^{\mathrm{ESS}}, \mathbf{P}^{\mathrm{D}}, \mathbf{P}^{\mathrm{INF}}, \mathbf{P}^{\mathrm{SOP}}\}$ is the second stage decision variable. $\mathbf{u} = \{\tilde{P}_{i,t}^{\mathrm{PV}}\}$ is the second stage uncertainty variable. Constant matrix $\mathbf{A}$ denotes the coefficient matrix

associated with decision variable $\mathbf{x}$, and column vector $\mathbf{d}$ is a constant denoting the coefficient vector associated with decision variable $\mathbf{x}$. Constant matrices $\mathbf{G},\mathbf{E},\mathbf{R}$ denote the coefficient matrix associated with decision variable $\mathbf{y}$. Column vectors $\mathbf{h}$ and $\mathbf{r}$ are constant vectors denoting the coefficient vector associated with decision variable $\mathbf{y}$. The constant matrix $\mathbf{M}$ represents the matrix of coefficients associated with the uncertain variable $\mathbf{u}$. $\Omega\left(\mathbf{y},\mathbf{u}\right)$ is the feasible domain of continuous variable $\mathbf{y}$ given $(\mathbf{x},\mathbf{u})$. $\mathbf{c}^{\mathrm{T}}\mathbf{y}$ denotes the objective function of the second stage, corresponding to (16). (53-a) corresponds to the constraints associated with variable $\mathbf{x}$ in the first stage. (53-b) corresponds to the constraints associated with variable $\mathbf{y}$ in the second stage. (53-c) corresponds to the second-order cone constraints associated with variable $\mathbf{y}$ in the second stage.

For the two-stage robust optimization model in (53), it includes continuous variables and integer variables. Since the second stage contains uncertain variables $\mathbf{u}$, it cannot be solved directly. Therefore, this paper adopts the C& CG method[25,26] to transform it into a master-slave problem for solution. The master problem seeks the worst-case scenario that maximizes the load recovery amount. The problem is then decomposed into solving the integer variables of the master problem first (including the charging and discharging state of the storage device, switch status, and the number of information nodes, etc.), optimizing the objective function for the second-stage variables, and iteratively updating the feasible region to obtain the total load recovery amount under the worst-case scenario.

2) C&CG solution methods.
1) Main and subproblem model.

The main and subproblem model corresponding to (53) are:

$$\text{MP1:}\begin{cases} \min_{x,y,u}(\eta) \\ s.t.\ \mathbf{Ax}\leqslant\mathbf{d} \\ \mathbf{Gy}^{l}\leqslant\mathbf{h}-\mathbf{Ex}-\mathbf{Mu}^{l}\ \forall l\leqslant k \\ \left\|\mathbf{Ry}^{l}\right\|_{2}\leqslant\mathbf{r}^{\mathrm{T}}\mathbf{y}^{l}\ \forall l\leqslant k \\ \eta\geqslant\mathbf{c}^{\mathrm{T}}\mathbf{y}^{l}\ \forall l\leqslant k \end{cases} \tag{54}$$

$$\text{SP1:}\begin{cases} \max_{u\in\mathbf{U}}\ \min_{y\in\Omega(x,u)}\ \mathbf{c}^{\mathrm{T}}\mathbf{y} \\ s.t.\ \mathbf{Gy}\leqslant\mathbf{h}-\mathbf{Ex}^{*}-\mathbf{Mu}:\pi \\ \left\|\mathbf{Ry}\right\|_{2}\leqslant\mathbf{r}^{\mathrm{T}}\mathbf{y}:\tau^{a},\tau^{b} \end{cases} \tag{55}$$

First, solve the main problem MP1 corresponding to (54). MP1 is a mixed-integer second-stage robust optimization problem. The first stage variable solution $\mathbf{x}^{*}$ corresponding to MP1 and the auxiliary variable η introduced in the $k+1$ iteration are C& CG cuts. Then, the first-stage decision variable $\mathbf{x}^{*}$ is brought into the second-stage subproblem SP1 to seek the worst-case scenario $\mathbf{u}^{l}$, where l is the historical iteration number and k is the current iteration number. Finally, the worst-case scenario $\mathbf{u}^{l}$ of the second-stage subproblem SP1 is brought back into the first-stage main problem MP1 for solving. The last three constraints in (54) are the set of extreme cut and non-intersecting cut formed by the worst-case scenario variables π,τ^{a},τ^{b}.

2) Subproblem solving methods.

(55) is a $\max-\min$ optimization problem, we use the duality theorem to convert the inner $\min$ problem of (55) into its dual form to merge it into a maximization problem, which is shown in the form of (56):

$$\begin{cases} \max_{u,\pi}-(\mathbf{h}-\mathbf{Mu}-\mathbf{Ex})^{\mathrm{T}}\pi \\ s.t.\ \mathbf{c}+\mathbf{N}^{\mathrm{T}}\pi+\mathbf{R}^{\mathrm{T}}\tau^{a}+\mathbf{r}\tau^{b}=0 \\ \left\|\tau^{a}\right\|_{2}\leqslant\tau^{b} \\ \pi,\tau^{a},\tau^{b}\geqslant0 \end{cases} \tag{56}$$

where there exists a bilinear term $(\mathbf{Mu})^{\mathrm{T}}\pi$, which is solved using the outer approximation method for bilinear terms[26]. The main problem MP2 and subproblem SP2 are obtained, as shown in (57) and (58):

$$\text{SP2:}\begin{cases} \max_{u,\pi}-(\mathbf{h}-\mathbf{Mu}-\mathbf{Ex})^{\mathrm{T}}\pi \\ s.t.\ \mathbf{c}+\mathbf{N}^{\mathrm{T}}\pi+\mathbf{R}^{\mathrm{T}}\tau^{a}+\mathbf{r}\tau^{b}=0 \\ \left\|\tau^{a}\right\|_{2}\leqslant\tau^{b} \\ \pi,\tau^{a},\tau^{b}\geqslant0 \end{cases} \tag{57}$$

$$\text{MP2:} \begin{cases} \max\limits_{u,\pi} -(\mathbf{h} - \mathbf{Ex})^{\mathrm{T}}\pi + \beta \\ s.t. \ \mathbf{c} + \mathbf{N}^{\mathrm{T}}\pi + \mathbf{R}^{\mathrm{T}}\tau^a + \mathbf{r}\tau^b = 0 \\ \|\tau^a\|_2 \leqslant \tau^b \\ \pi, \tau^a, \tau^b \geqslant 0 \\ \beta \leqslant G^m(\mathbf{u}, \pi), \forall m \leqslant n \end{cases} \quad (58)$$

where MP2 and SP2 are used to solve the upper and lower bounds of (56), respectively. m is the number of historical iterations, and n is the number of current iterations. An auxiliary variable β is introduced to replace the bilinear term in the original equation, and a bilinear term exists in (58) $G^m(\mathbf{u}, \pi) = (\mathbf{Mu})^{\mathrm{T}}\pi$. Therefore, it is necessary to use the outer approximation method for linearization, and the linearization formula is shown in (59):

$$G^n(\mathbf{u}, \pi) = (\mathbf{u}^n)^{\mathrm{T}}\pi_{sp}^n + (\mathbf{u} - \mathbf{u}^n)^{\mathrm{T}}\pi_{sp}^n + (\pi - \pi_{sp}^n)^{\mathrm{T}}\mathbf{u}^n \quad (59)$$

3) Solution steps and processes.

The specific solving steps of the C& CG algorithm are as follows:
1) Set the initial values of the upper and lower bounds of the master-slave problem to be $U1 = +\infty$ and $L1 = -\infty$, respectively, and the initial number of iterations to be $k = 1$, given that the convergence value is $\varepsilon1^{\max}$.
2) Solve the worst-case main problem, where the constraints of the main problem do not contain C& CG cut, and obtain the integer solution $\mathbf{x}^*$
3) Based on the integer solution $\mathbf{x}^*$ obtained from solving the main problem, solve the subproblem to obtain the worst case scenario $(\mathbf{u}^{l+1})^*$, and the objective function value of the subproblem, and update the upper limit $U1 = \min(U1, \mathbf{c}^{\mathrm{T}}\mathbf{y})$.
4) Determine whether $U1 - L1$ is smaller than the convergence value $\varepsilon1^{\max}$, if it is valid, then end. If not, make $k = k + 1$, and add a new set of scenario variables $\mathbf{u}^k$ and C& CG cuts for the main problem. Solve the main problem to get η^* to update the lower bound $L1 = \max(L1, \eta^*)$ and go to step 3.
The specific solution steps are shown in Fig. 6.

Example analysis

To verify the feasibility of the proposed adaptive data driven multi period robust power supply recovery method for distribution networks, this section uses the IEEE-33 bus system and a distribution network in the northeastern city of Hubei Province for case analysis. For convenience, it is assumed that the availability factors of the two PV systems before the fault are equal. According to the calculation method provided in[19], the value of $k_{\min}$ is 1.41.

1) 33-bus system example analysis.

The connection diagram of the modified IEEE-33 bus system is shown in Fig. 7. In Fig. 7, with the substation disconnected, the distribution network is in a state of total power outage. The system load is powered by diesel generators, PV systems and BESS. Table 1 provides the parameter settings for the integrated system's PV, BESS, and SOP. The system's base voltage is 12.66 kV, and the base capacity is 10MVA. The active power range for the connection is 0–5000 kW, the reactive power range is 0-5000kvar, the upper limit for branch current magnitude is 0.5 p.u., and the node voltage magnitude range is 0.95–1.05 p.u.

1) Power supply recovery results when using hyperplane set.

Figs.8,9,10 illustrate the load recovery situation and the energy supply of the system during the power outage phase when using the hyperplane uncertain set. According to the diagrams, the energy supply of the 33-bus case study system during the power outage phase relies mainly on distributed PV systems, BESS, and diesel generators. When there is no PV output (1–7 h), the energy supply mainly relies on diesel generators and BESS. During the PV output period (13–14 h), the large-scale access to PV energy leads to a significant reduction in the load recovery ratio, and the system's load is more adequately supplied. For areas with adequate PV resources, the large-scale PV output can significantly reduce the load recovery ratio. When the PV output is insufficient, the system compensates for the insufficient energy supply with BESS, which meets the load recovery needs, while diesel generators operate to meet the system's basic operating requirements. During the grid fault phase, the intelligent switch S2 at node 4 is triggered and opened, causing node 3 to lose power. When the PV output increases, the system begins to supply power from node 4 to node 9, reversing the power flow from node 4 to node 9. This change is due to the large-scale PV energy injection.

2) Effect of scaling factor k on power supply recovery

Table 2 shows the load recovery situation of the 33-bus system under different sets. According to the table data, when the convex hull set is used, as the number of contractions increases, the load recovery ratio of the system gradually decreases, the number of recovered loads also decreases, and network losses are also reduced. However, the rate of reduction slows down as the number of contractions increases. This is because the convex
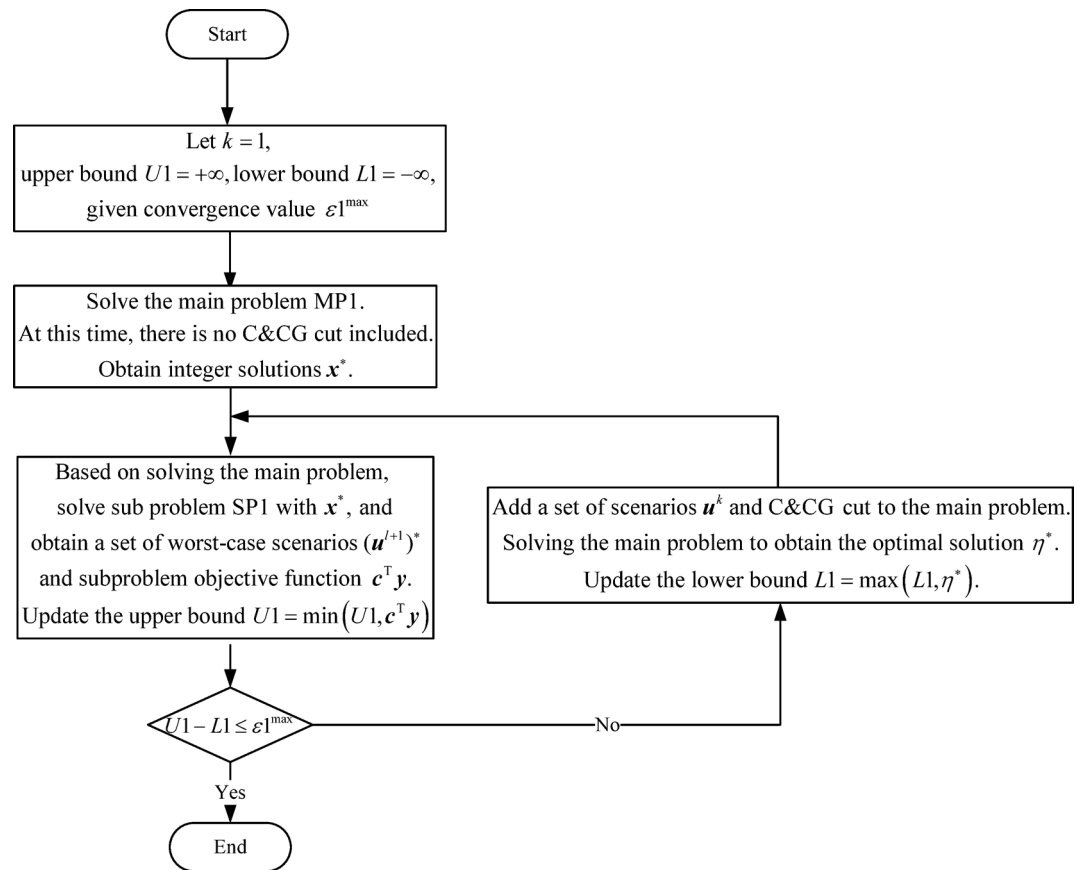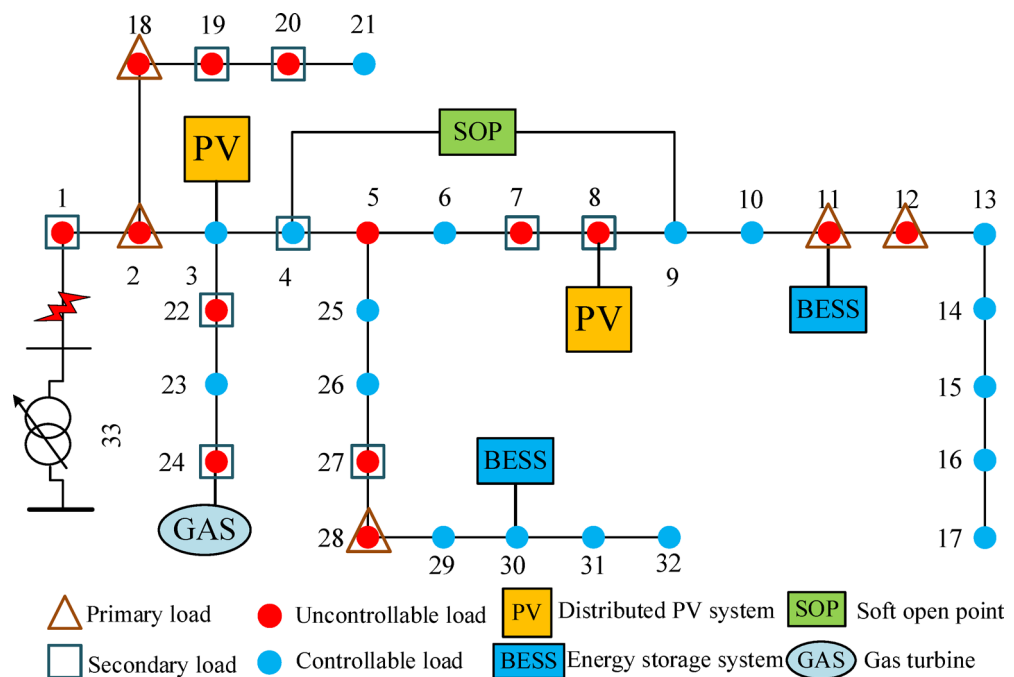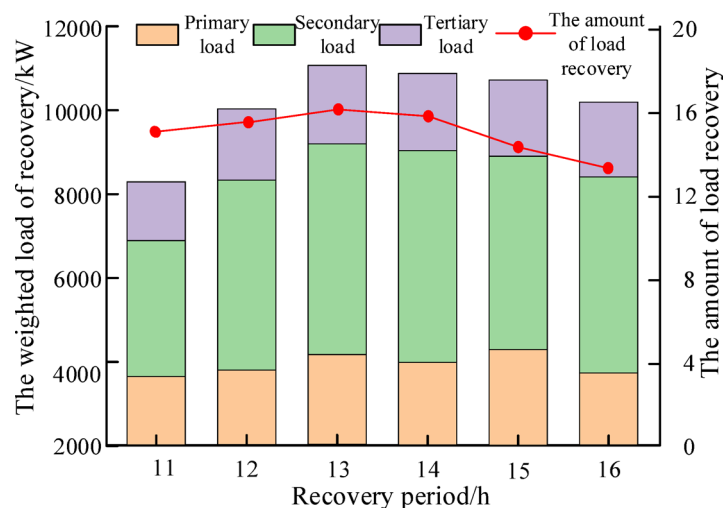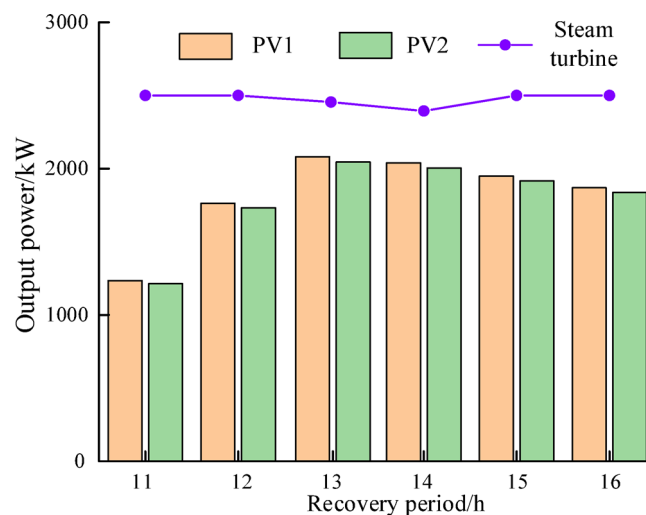
Fig. 6. Solution flow chart of C&CG.

Fig. 7. Modified IEEE-33 bus test system.

Device	Access node	Parameter
PV	3, 8	Capacity: 2500 kW
BESS	11/30	Charge state: 5%~95%
		Capacity: 1000 kW
		Charge/discharge efficiency: 95%
SOP	4, 9	Capacity: 500 kW
		Loss coefficient: 0.02
GAS	24	Capacity: 2500 kW

Table 1. System configuration parameters.

Fig. 8. Power supply restoration situation.

Fig. 9. Energy output situation.

hull set expands the historical maximum data boundary, making the worst-case scenario more likely to occur. To achieve a balance between reducing network loss and the uncertainty introduced by PV injection, the system needs to filter out part of the power injected by distributed PV, reducing the load recovery ratio and the number of recovered loads. For network loss, the power loss and load loss both decrease due to the reduction in new energy injection, resulting in a certain degree of stability in the system. Compared with the convex hull set and the box set, the recovered weighted load of the hyperplane set can be increased by 0.55% and 0.56% respectively.
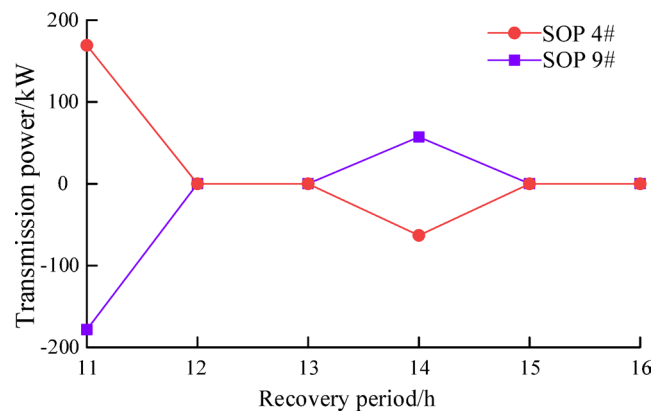
Fig. 10. SOP transmission status

Different sets	Restore weighted load/kW	Number of loads restored	Network losses/kW
$k = 0.6$	61454.11	83.26	192.09
$k = 0.8$	61243.55	82.55	191.28
$k = 1$	61032.92	81.83	190.50
$k = 1.2$	60715.02	81.24	188.24
$k = 1.4$	60700.19	80.21	188.18
Hyperplane set	61033.95	81.90	190.07
Box set	60695.21	80.09	188.19

Table 2. The total impact of scaling factor k on fault recovery.
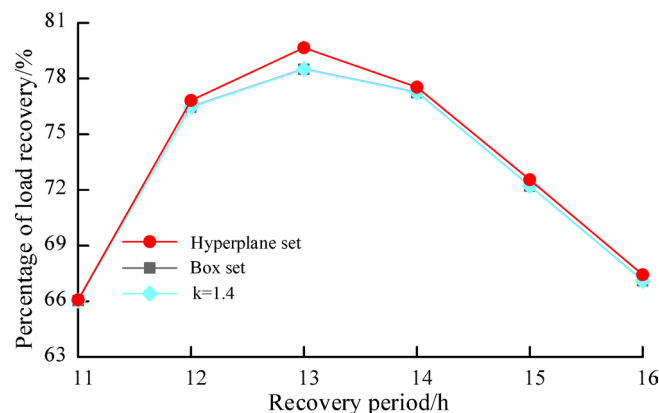
Fig. 11. Load recovery percentage under three uncertain sets.

This verifies that the hyperplane set can improve the robustness of the optimization results when considering all the power outputs of photovoltaics.

When using the hyperplane set, it applies the mathematical model of the hyperplane to the historical maximum data boundary, significantly reducing the coverage range. This method can effectively balance the system's robustness and reduce system loss. Therefore, using the hyperplane set, the load recovery ratio and the number of recovered loads are the smallest. Compared with the convex hull set, it can better respond to the worst-case scenario of PV output. Due to the balance of robustness and efficiency, the load recovery ratio and the number of recovered loads of the system are higher.

3) Impact of three uncertainty sets on power supply recovery.

The load recovery situation at each time period is shown in Figs. 11 and 12. From the figures, it can be seen that the 33-bus system load recovery reaches its peak at 13:00, which is consistent with the PV output situation. When using the three types of uncertain sets for system fault recovery, it is observed that the hyperplane set
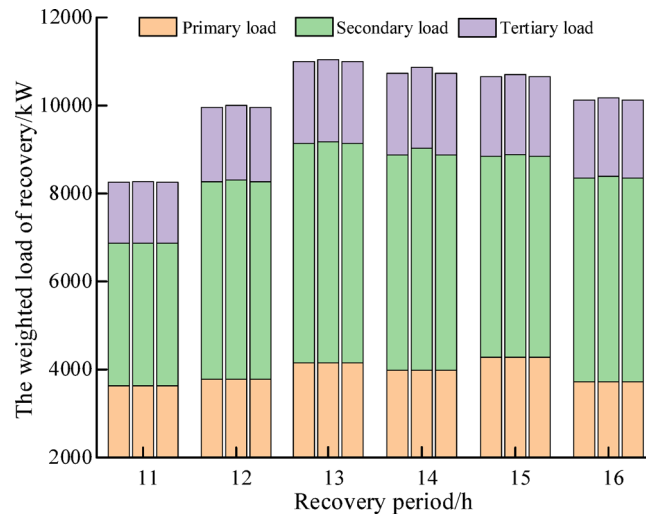
Fig. 12. :Restoring weighted load cases using three uncertain sets. (From left to right are the convex hull set, the hyperplane set, and the box set).
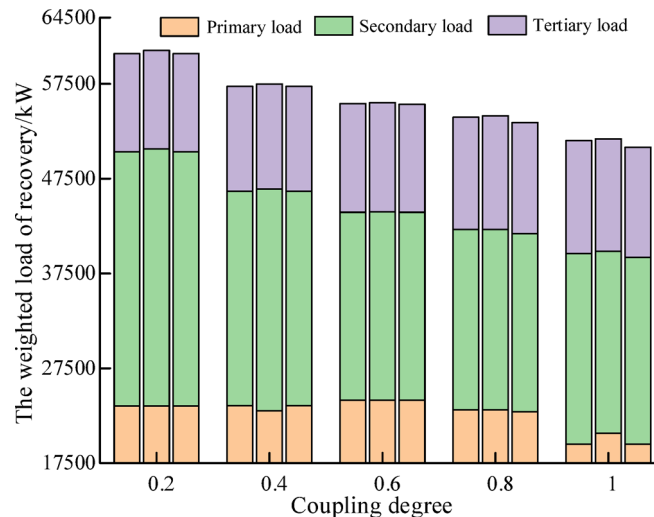
Fig. 13. Restoration of weighted load under different. (From left to right are the convex hull set, the hyperplane set, and the box set).

results in a higher load recovery at each time period compared to the convex hull set and the box set. This is consistent with the previous analysis.

4) Effect of coupling factor γ on power supply recovery results

The Figs. 13, 14 and 15 give the impact of taking each of the three uncertainty sets on power supply recovery as the coupling degree changes from 0.2 to 1. From the Fig. 13, it can be seen that the weighted load of 33-bus system recovery decreases as the coupling degree increases. This is due to the fact that when the coupling degree increases, the information node has a higher recovery demand on the power node, resulting in the power node not being able to meet the energy supply demand of the information node's end devices during power supply recovery, which leads to a decrease in the amount of load recovered. In the Fig. 14, when the coupling degree increases, in order to increase the energy supply to the information nodes as much as possible, the power nodes increase the transmission flow of power to each node, which leads to an increase in the network loss of the system. In the Fig. 15, when the coupling coefficient is small, the information nodes are less dependent on the power nodes, and the information network is more resistant to interference, so more information nodes can be recovered. However, when the coupling coefficient increases, since the premise of information node restoration is that the power node load must be completely restored, and it can be seen from the attached Fig. 13 that at this time the amount of load restored by the power system is less, which leads to a reduction in the number of information nodes restored as well, and makes the information network relatively fragile.
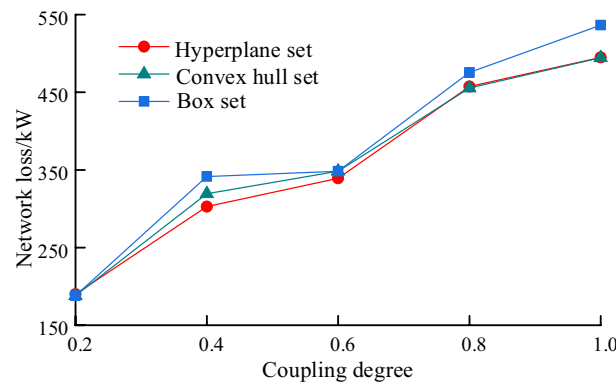
Fig. 14. Network loss situation under different.
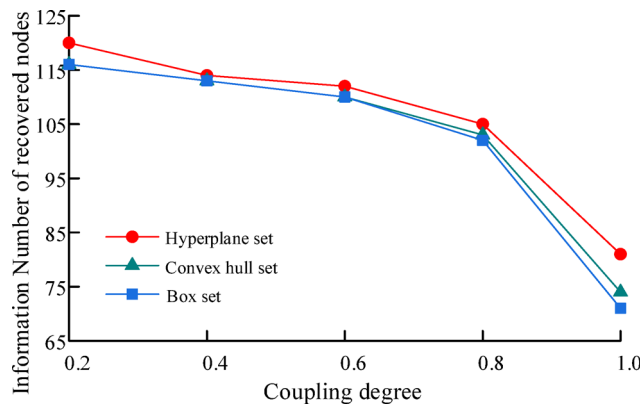
Fig. 15. Information node recovery under different.

2)   Analysis of the 26-bus example in an urban area in Hubei.

The connection diagram of the distribution network system in a city in Hubei Province is shown in Fig. 16. Similar to the 33-bus system, in Fig. 16, the substation is disconnected, and the distribution network is in a state of total power outage. The system load is supplied by diesel generators, distributed PV systems, and BESS. Table 3 provides the parameter settings for the integrated system's PV, BESS, and SOP. The system's base voltage is 12.66 kV, and the base capacity is 10 MVA. The active power range for the connection is 0–5000 kW, the reactive power range is 0–5000 kvar, the upper limit for branch current magnitude is 0.5 p.u., and the node voltage magnitude range is 0.95–1.05 p.u.

1)   Fault recovery results when using hyperplane aggregation.

The load recovery situation of the 26-bus power supply system in a city in Hubei is shown in Figs. 17, 18 and 19. It is worth mentioning that due to the almost negligible total load of the third-level loads in this city's system, the third-level loads are not recovered during fault recovery. The recovery results are similar to the 33-bus system, with the load recovery reaching its peak during the PV output period. During other periods, diesel generators operate at full power to compensate for the system's load deficiency.

2)   Effect of scaling factor k on optimization results

The effect of deflation multiplier k on the fault recovery in this urban area is shown in Table 4, from which it can be seen that when convex packet aggregation is used, the deflation multiplier decreases as it gradually increases, both in terms of the total amount of weighted loads recovered, the number of loads and the network loss, which is consistent with the results of the 33-node system fault recovery, while the use of hyperplane aggregation is still one of the uncertainty aggregations with the highest amount of weighted loads recovered. Compared with the convex hull set and the box set, the recovered weighted load of the hyperplane set can be increased by 0.33% and 0.35% respectively. This verifies that the hyperplane set can improve the robustness of the optimization results when considering all the power outputs of photovoltaics.

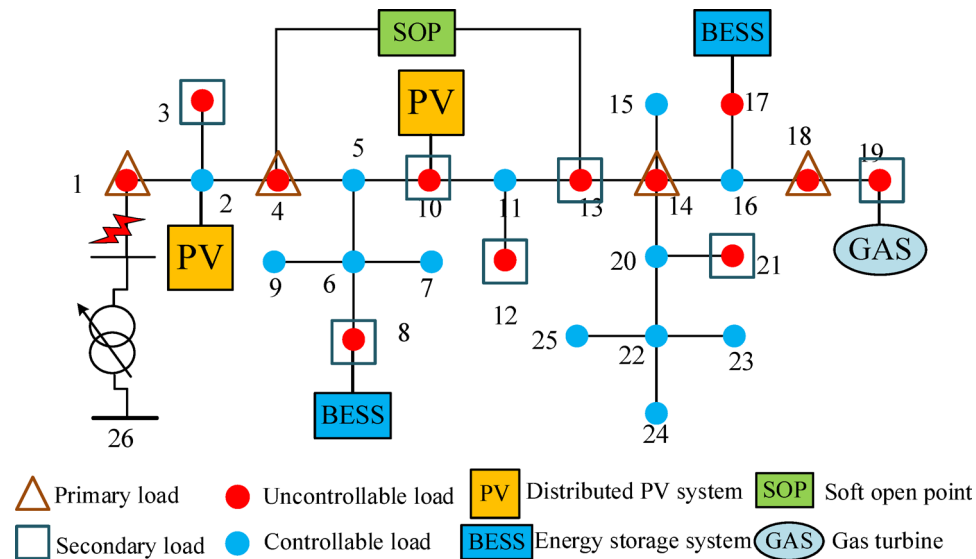3)   Impact of three uncertainty sets on power supply recovery.

Fig. 16. Testing system for a certain urban area in Hubei.

Device	Access node	Parametric
PV	2、10	Capacity: 1500 kW
BESS	11、30	Charge state: 5%~95%
		Capacity: 500 kW
		Charge/discharge efficiency: 95%
SOP	4、9	Capacity: 500 kW
		Loss factor: 0.02
GAS	24	Capacity: 1000 kW

Table 3. System configuration parameters.
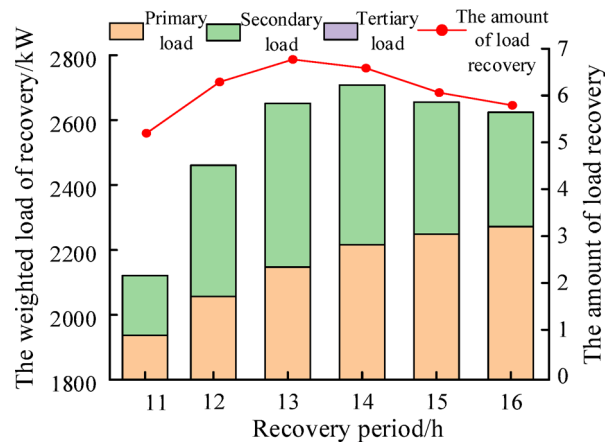
Fig. 17. Restoration of power supply in a certain urban area system.

The results of the three uncertainty sets for the recovery of a 24-node fault in an urban area are shown in the Figs. 20 and 21. The only difference with the 33-node system is that the 24-node system, due to its particular load distribution, therefore, the total amount of recovered tertiary loads is almost zero, while the rest of the variations are similar to that of the 33-node system.

4) Effect of coupling factor γ on optimization results
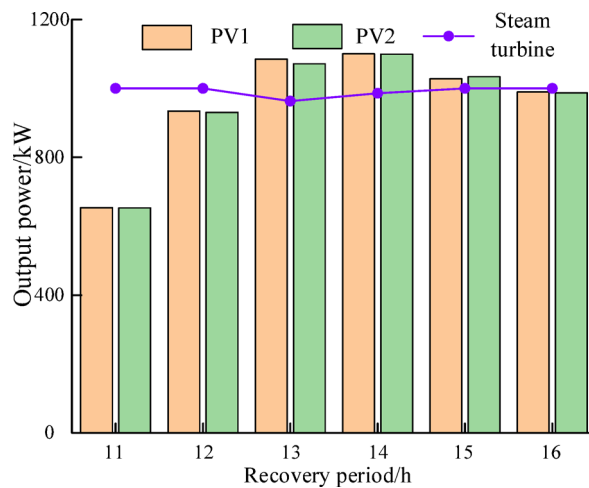
Fig. 18. Energy output of a certain urban area system.
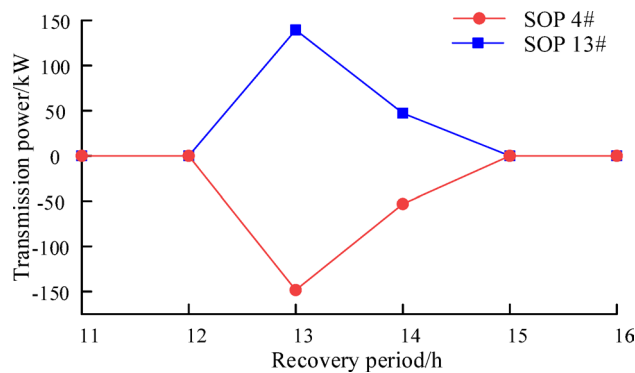
Fig. 19. SOP transmission status.

Different sets	Restore weighted load/kW	Number of loads restored	Network losses/kW
$k = 0.6$	15311.9	36.57	97.86
$k = 0.8$	15264.7	36.38	96.76
$k = 1$	15217.4	36.14	96.56
$k = 1.2$	15170.1	36.02	94.89
$k = 1.4$	15167.2	35.96	94.57
Hyperplane set	15217.8	36.15	96.85
Box set	15165.2	35.92	94.52

Table 4. The impact of scaling factor k on various costs.

The power supply recovery results of coupling coefficient γ for 24 nodes in an urban area are shown in Figs. 22, 23 and 24, and it can be seen from Fig. 22 that when the coupling coefficient is changed from 0.2 to 0.8, the change of the recovered weighted load is not so obvious compared with that of the 33-node system, which is because of the difference of the load distribution of the actual example and the load distribution of the standard example, and the load recovery is not so obvious due to the fact that the load distribution of many electric power nodes is zero for the actual example. node loads are 0, which leads to the fact that when the coupling coefficient changes, the restored node loads may not change, so the change in load restoration is not so obvious, while when $\gamma = 1$ is in place, the information nodes have a high demand for restoration of the power nodes due to the complete coupling of information and power networks, so the impact on the load restoration of the power system is also the largest. The variation of network loss and node recovery under different coupling coefficients is similar to that of 33-node.
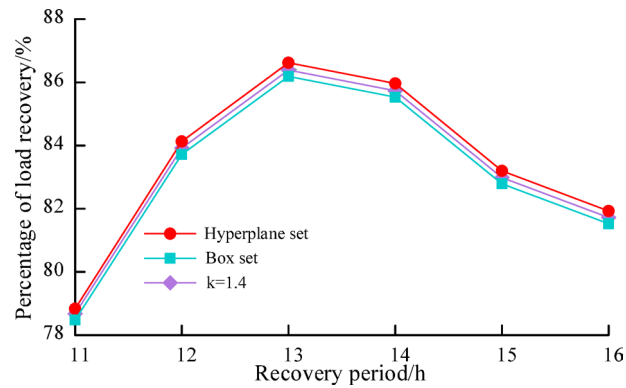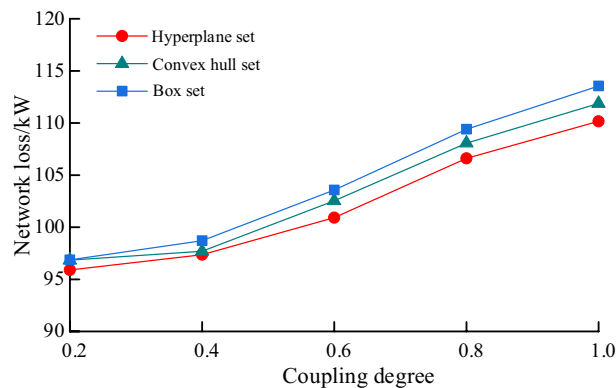
**Fig. 20**. Load recovery percentage under three uncertain sets.
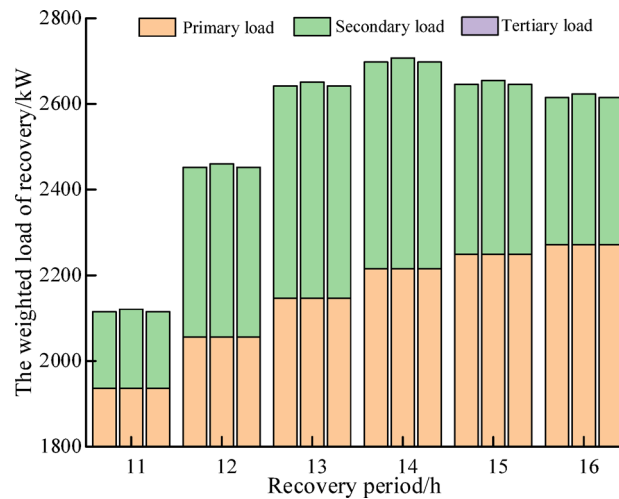
Fig. 21. Restoring weighted load cases using three uncertain sets. (From left to right are the convex hull set, the hyperplane set, and the box set).
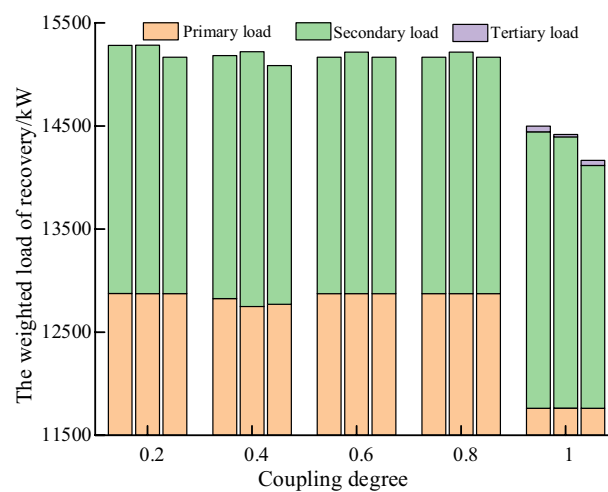
Fig. 22. Restoration of weighted load under different. (From left to right are the convex hull set, the hyperplane set, and the box set).

Fig. 23. Network loss situation under different.
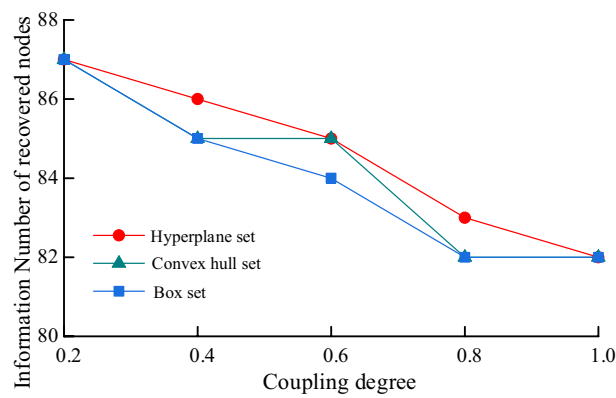
Fig. 24. Information node recovery under different.

Conclusion

This paper constructs a distribution network fault recovery model based on adaptive data driven methods and employs the C& CG method for solving. Finally, through the case studies of the IEEE-33 bus system and a 24-bus system in a city in Hubei Province, the power supply recovery models under three types of uncertain sets are calculated, and the results are as follows:

(1) Under the same coupling coefficient, in the IEEE-33 bus system, the recovered weighted load of the hyperplane set is increased by 0.55% and 0.56% respectively compared with that of the convex hull set and the box set. In the actual 26-bus example system, the recovered weighted load of the hyperplane set is increased by 0.33% and 0.35% respectively compared with that of the convex hull set and the box set. This verifies that the hyperplane set can improve the robustness of the optimization results when considering all the power outputs of photovoltaics. Furthermore, with the increase in the coupling factor, the load recovery ratio and the number of recovered loads of the system using the convex hull set gradually decrease, increasing the system's robustness.

(2) As the system load increases, the dependence of the information network and the power network on each other continues to increase. Therefore, the power supply recovery ratio of the system shows a downward trend. However, the hyperplane set performs better, especially in accurately depicting the parameter range of the uncertain set, reducing the empty regions with low probability and reducing the system loss. Hence, the hyperplane set based on adaptive data driven methods provides a more robust and reliable fault recovery method than the convex hull set, showing stronger robustness.

The data driven multi period robust fault recovery method for distribution networks mentioned in this paper is aimed at the hourly fault recovery scenarios of an actual distribution network in a certain place. In time-sensitive real-time fault recovery scenarios, due to the iterative characteristics of the C&CG algorithm, there may be the problem of relatively large computational complexity. Future research can take advantage of the strengths of modern multi-core processors and distributed computing architectures, conduct in-depth research on the parallelizable parts of the C&CG algorithm. By designing a reasonable parallel algorithm framework, multiple computing nodes can simultaneously handle different subtasks, thereby significantly reducing the overall computing time. At the same time, in order to demonstrate the applicability of this method to large-scale networks in the real world, in our subsequent research, we will actively cooperate with relevant power enterprises to obtain representative large-scale actual distribution network data, providing a reliable basis for testing our method.

Data availability

Data will be made available on request. If you would like to request the data for this study, please contact the corresponding author, Xi Ye.

Reference

1. Xiaowei, W. A. N. G. et al. Distribution network restoration supply method considers 5G base station energy storage participation[J]. *Energy* **289**, 129825 (2024).
2. Yumian, L. I. N. et al. Distributionally robust service restoration for integrated electricity-heating systems considering secondary strikes of subsequent random events[J]. *Appl. Energy.* **380**, 125038–125038 (2025).
3. Xi, L. U. et al. An operation model for integrated electricity and heat systems emphasizing modeling of both networks and uncertainties[J]. *Appl. Energy.* **370**, 123595 (2024).
4. Wenjia, Z. H. A. N. G. et al. Clustering and decision tree based analysis of typical operation modes of power systems[J]. *Energy Rep.* **9** (S8), 60–69 (2023).
5. Tianlei, Z. A. N. G. et al. Integrated planning and operation dispatching of source-grid-load-storage in a new power system: a coupled socio-cyber-physical perspective[J]. *Energies* **17** (12), 3013–3013 (2024).
6. Gang, H. U. A. N. G. et al. Cyberconstrained optimal power flow model for smart grid resilience enhancement[J]. *IEEE Trans. Smart Grid.* **10** (5), 5547–5555 (2019).
7. Baozhong, T. I. et al. Resilience assessment and improvement for cyberphysical power systems under typhoon disasters[J]. *IEEE Trans. Smart Grid.* **13** (1), 783–794 (2022).
8. Bangjie, Q. I. U. et al. An optimization-based intentional islanding scheme for service restoration in distribution systems considering anti-parallel operation of distributed generations[J]. *Appl. Sci.* **13** (9), 5501 (2023).
9. Huaxia, L. I. U. XU Yan. Research on multi-source collaborative distribution network fault recovery technology[J]. Journal of Physics: Conference Series, 2564(1): 012015. (2023).
10. Chen Chun, H. et al. Optimal configuration of distributed energy storage considering intending Island recovery in faulty distribution networks[J]. *Int. J. Electr. Power Energy Syst.* **158**, 109982 (2024).
11. Yuduo, Z. H. A. O. et al. A fault recovery strategy of flexible interconnected distribution network with SOP flexible closed-loop operation[J]. *Int. J. Electr. Power Energy Syst.* **142** (PB), 108360 (2022).
12. Jie, J. I. A. N. et al. Multi-stage supply restoration of active distribution networks with SOP integration[J]. Sustainable Energy, Grids and Networks, 29: 100562. (2022).
13. Baofeng, T. A. N. G. et al. Robust fault recovery strategy for multi-source flexibly interconnected distribution networks in extreme disaster scenarios[J]. *Front. Energy Res.* **12**, 1345839 (2024).
14. Feng, W. A. N. G. et al. Service restoration for distribution network with DGs based on stochastic response surface method[J]. *Int. J. Electr. Power Energy Syst.* **107**, 557–568 (2019).
15. Cheng, W. A. N. G. et al. Data-driven wind generation admissibility assessment of integrated electric-heat systems: a dynamic convex hull based approach[J]. *IEEE Trans. Smart Grid.* **PP** (99), 1–1 (2020).
16. Zhanghao, H. U. A. N. G., Yachao, Z. H. A. N. G. & Shiwei, X. I. E. Data-adaptive robust coordinated optimization of dynamic active and reactive power flow in active distribution networks[J]. *Renew. Energy.* **188**, 164–183 (2022).
17. Cheng, W. A. N. G. et al. Data-driven adjustable robust unit commitment of integrated electric-heat systems[J]. *IEEE Trans. Power Syst.* **36** (2), 1385–1398 (2021).
18. Zhichun, Y. A. N. G. et al. Two stage affinely adjustable robust optimal scheduling for AC/DC hybrid distribution network based on source-grid-load-storage coordination[J]. *Energy Rep.* **8**, 15686–15701 (2022).
19. Yuming, Y. E. et al. A robust optimization method for new distribution systems based on adaptive data-driven polyhedral sets[J]. *Front. Energy Res.* **12**, 1351907 (2024).
20. Haiyue, Y. A. N. G. et al. Two-stage robust optimal scheduling of flexible distribution networks based on pairwise convex hull[J]. *Sustainability* **15** (7), 6093 (2023).
21. Yipu, Z. H. A. N. G. et al. Data-adaptive robust optimization method for the economic dispatch of active distribution networks[J]. *IEEE Trans. Smart Grid.* **10** (4), 3791–3800 (2018).
22. BIFEI, T. et al. Two-stage robust optimization dispatch for multiple microgrids with electric vehicle loads based on a novel data-driven uncertainty set[J]. *Int. J. Electr. Power Energy Syst.* **134**, 107359 (2022).
23. Xiaoguang, S. U. N. et al. Research on emergency repair strategy based on load grade and economy[J]. Journal of Physics: Conference Series, 2378(1): 012086. (2022).
24. Geng, Z. H. A. N. G. et al. Fault recovery methods for a converged system comprised of power grids, transportation networks and information networks[J]. *Electronics* **12** (21), 4508 (2023).
25. AHMADPOUR, L., SAHRAEIAN, R. & ESHGHI, K. Resilient downstream oil supply chain under interdiction and recovery resources allocation using a column and constraint generation algorithm[J]. *Computers Oper. Res.* **177**, 106978–106978 (2025).
26. TONISSEN D D, ARTS J J, SHEN M, Z. A column-and-constraint generation algorithm for two-stage stochastic programming problems[J]. *TOP* **29** (3), 1–18 (2021).

Author contributions

Xi Ye and Meng Yang wrote the main manuscript text. Zhihong Yang and Qing Xiang performed the visualization. Yazhuo Li performed the validation. All authors reviewed the manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to X.Y.