

# Research on Optimal Configuration of Distributed Generation based on Multi-scenario Analysis

Jianguo Jiang, Lili Yang

School of Electrical and Information Engineering, Northeast Petroleum University, Daqing, Heilongjiang 163318, China

**Abstract:** In order to make the grid-connected planning of distributed generation more reasonable, the uncertainties of intermittent distributed generation output and load forecasting are included in the solution process. Firstly, multi-scenario analysis is introduced to transform the source load uncertainty problem into deterministic problem, and the Latin super-force sampling method is used to generate the initial planning scene. The density peak clustering idea and elbow method are used to improve the K-means clustering algorithm and reduce the scene. Secondly, the optimal allocation model of grid-connected distributed generation is constructed with the minimum annual comprehensive cost as the objective function. Finally, in view of the slow convergence speed and easy to fall into local optimum of particle swarm optimization (PSO), an adaptive inertia weight factor is adopted to improve PSO, and the effectiveness of the proposed model and method is verified by IEEE 33-node standard simulation example.

**Keywords:** Distributed Generation; Uncertainty; Multi-Scenario Analysis; Improved Particle Swarm Optimization Algorithm.

## 1. Introduction

At present, China's distribution network is still mainly operated by large units and large power grid, which has poor power supply flexibility and reliability, is difficult to meet the power demand of distribution network terminal load, and is easily affected by fault disturbance, resulting in large-scale power outage or even paralysis of the system [1]. Distributed generation (DG), which is close to the user, has the advantages of small floor area, high reliability of power supply, flexible power generation mode, etc. [2]. Its power generation is between tens of kilowatts and tens of megawatts, which can solve the power supply problems of remote loads and loads in remote areas, so as to make up for the defects of large power grids. After the distributed generation is connected to the distribution network, it has a certain impact on the system operation mode and network loss, which is closely related to its access location and capacity. Therefore, the research on grid-connected optimal configuration of distributed generation has become a major research hotspot [3] [4].

Scholars at home and abroad have made in-depth research on the optimal allocation of distributed generation grid-connected. Reference [5] considers three indexes of power generation cost, environmental cost and active power loss cost at the same time, and takes the minimum comprehensive cost as the objective function, and constructs a distributed generation location and constant capacity model. Reference [6] considers the power cost, network loss cost, environmental cost and voltage offset cost, establishes a distributed grid-connected optimization model with the goal of minimizing the total cost, and solves the model by using an improved krill swarm search algorithm. In reference [7], aiming at the voltage quality related problems after the connection of distributed generation, the optimal configuration model of distributed generation is constructed with the objective function of minimizing the economic loss of voltage sag and minimizing the investment cost of installing DG, and the improved firefly algorithm is used to solve the model. The above models start from various angles,

but none of them consider the influence of output and load uncertainty of distributed generation, which leads to the lack of rationality of distributed generation planning results.

To sum up, this paper takes the source load uncertainty into account in the research of distributed generation grid-connected planning, introduces the concept of multi-scenario analysis, generates large-scale source load planning scenarios to describe various situations that may occur in practice, and substitutes them into the distributed generation optimal configuration model to solve them, and optimizes them by improving particle swarm optimization algorithm, and simulates them on IEEE 33-bus system.

## 2. Uncertainty Modeling

### 2.1. Wind Power Model

Wind speed generally accords with two-parameter Weibull distribution, and its probability function model is shown in Equation (1)[8]:

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

Where  $v$  is the actual wind speed;  $k$  and  $c$  are the shape parameter and scale parameter of Weibull distribution function, respectively.

The actual output power of the fan is closely related to the wind speed, and the specific relationship can be expressed as follows:

$$P_{\text{WTG}} = \begin{cases} 0 & 0 \leq v \leq v_{ci} \text{ or } v \geq v_{co} \\ P_{\text{WTG}}^r \frac{v - v_{ci}}{v_r - v_{ci}} & v_{ci} \leq v \leq v_r \\ P_{\text{WTG}}^r & v_r \leq v \leq v_{co} \end{cases} \quad (2)$$
$$Q_{\text{WTG}} = \frac{P_{\text{WTG}}}{\tan \varphi_{\text{WTG}}} \quad (3)$$

Where  $P_{\text{WTG}}^r$  is the rated active power of the fan;  $v_{ci}$ ,  $v_r$ ,  $v_{co}$  are the cut-in wind speed, rated wind speed and cut-out wind speed of the fan respectively;  $\varphi_{\text{WTG}}$  is the power factor angle for wind power generation.

## 2.2. Photovoltaic Model

The probability model of illumination intensity usually adopts Beta distribution, and its functional expression is as follows [9]:

$$f(I) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{I}{I_{\max}}\right)^{\alpha-1} \left(1 - \frac{I}{I_{\max}}\right)^{\beta-1} \quad (4)$$

The actual active power output of photovoltaic changes with the intensity of illumination. Photovoltaic generators are generally equipped with reactive power compensation devices to keep their power factors constant, so the calculation formulas of actual active power and reactive power output of photovoltaic are shown in Equations (5) and (6) respectively:

$$P_{\text{PVG}} = \begin{cases} P_{\text{PVG}}^r \frac{I}{I_r} & I \leq I_r \\ P_{\text{PVG}}^r & I > I_r \end{cases} \quad (5)$$

$$Q_{\text{PVG}} = \frac{P_{\text{PVG}}}{\tan \varphi_{\text{PVG}}} \quad (6)$$

Where  $P_{\text{PVG}}^r$  is the rated active power of photovoltaic;  $I_r$  is the rated illumination intensity;  $\varphi_{\text{PVG}}$  is the power factor angle of photovoltaic generator.

## 2.3. Load Model

The probability model of load can be expressed by normal distribution, and its active power probability density function can be expressed as formula (7):

$$f(P_L) = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left[-\frac{(P_L - \mu_p)^2}{2\sigma_p^2}\right] \quad (7)$$

$$Q_L = P_L \tan \varphi_L \quad (8)$$

Where  $\mu_p$  and  $\sigma_p$  are the expectation and variance of active load respectively.

## 3. Multi-scenario Analysis

Multi-scenario analysis is one of the methods to solve the problem that it is difficult to describe uncertainty by mathematical model. The core of multi-scenario analysis is to enumerate uncertain variables according to certain rules, and turn uncertain problems into multiple deterministic problems, which greatly reduces the complexity of problem solving. Among them, the larger the scale of the generated scene, the more accurate it is to describe the uncertainty problem, while too few generated scenes may lead to the omission of some representative scenes in the calculation, which greatly reduces the accuracy of the model solution [10]. However, large-scale scenarios will increase the computational time of optimization problems. In order to solve the contradiction between the accuracy of describing uncertain problems and the optimization speed, we can build a large enough scene to simulate uncertain problems, and then use the scene reduction technology to generate typical planning scenes to calculate the DG optimal configuration model to solve the solution speed problem. It can be seen that multi-scene analysis mainly includes scene generation and scene reduction.

### 3.1. Scene Generation

Based on the probabilistic model function of uncertain variables, this paper adopts Latin super-force sampling to generate large-scale scenes. Compared with Monte Carlo simulation, Latin hypercube sampling can ensure that all

sampling areas can be covered by sampling points, which makes the sampling results more representative [11].

The cumulative probability density curve  $Y_k = F_k(x_k)$  of random variables  $x_k$  is divided into  $N$  segments in the interval  $[0, 1]$ , and a probability value  $Y_{kn}$  ( $1 \leq n \leq N$ ) is extracted from each segment, and then the sample value of the variable  $Y_k = F_k^{-1}(x_k)$  in the first segment is obtained by using the variable  $x_k$ . In order to make the error between the sample expected value and the actual expected value of random variables as small as possible, the importance sampling method is used to improve Latin hypercube sampling, and the sample value is obtained according to the importance in the probability density function. The specific rules are as follows:

$$x_{kn} = \begin{cases} F_k^{-1}\left(\frac{n}{N}\right) & \frac{n}{N} \leq 0.5 \\ F_k^{-1}\left(\frac{n-1}{N}\right) & \frac{n}{N} > 0.5 \end{cases} \quad (9)$$

From Equation (9), it can be seen that the probability below the expected value is selected as the upper boundary point of each interval, and the probability above the expected value is selected as the lower boundary point of each interval. Each sampling point is selected in a way close to the expected value, which combines the idea of stratified sampling and importance, so that the mean value of the obtained samples is unchanged and the variance is small, and the error caused by the tail feature of random function is reduced.

### 3.2. Scene Reduction

K-means clustering algorithm is often used to reduce scenes in multi-scene analysis. The main advantages of this algorithm are simple principle, easy convergence and efficient processing of large-scale data. However, K-means clustering algorithm needs to artificially establish the optimal cluster number  $K$ , and the random selection of  $K$  value and initial clustering center point may lead to unstable final clustering results and low clustering accuracy [12]. Therefore, this paper introduces the idea of density peak clustering and elbow method to improve K-means clustering algorithm.

#### (1) Density peak clustering algorithm

Clustering by fast search and find of density peaks (CFSFDP) algorithm is an efficient clustering algorithm proposed by Alex Rodriguez in Science in 2014. The core idea of CFSFDP algorithm is to calculate the distance between two samples in the sample set, and at the same time, take a certain distance value about 1% ~ 2% before all distance values as the truncation distance of the sample, and define two variables, namely the local density of sample points and the relative distance to high local density points, by the truncation distance, and take these two variables as the selection criterion of clustering center [13].

The local density of sample points is calculated from the truncation distance as shown in Equation (10):

$$\begin{cases} \rho_i = \sum_j \chi(d_{ij} - d_c) \\ \chi(x) = \begin{cases} 1, x < 0 \\ 0, x \geq 0 \end{cases} \end{cases} \quad (10)$$

When the size of the data set is small, the local density can be calculated by Gaussian kernel function, as shown in Equation (11):

$$\rho_i = \sum_j \exp\left(-\frac{d_{ij}^2}{d_c^2}\right) \quad (11)$$

Calculate the mutual distance  $\delta_i$  between arbitrary points, as shown in equation (12):

$$\delta_i = \begin{cases} \min_{j: \rho_j > \rho_i} (d_{ij}), & \text{if } \exists j \text{ s.t. } \rho_j > \rho_i \\ \max_j (d_{ij}), & \text{otherwise} \end{cases} \quad (12)$$

Multiply  $\rho$  and  $\delta$  to get a product  $r$  and arrange it in descending order of its value. The larger the value of  $r$ , the more likely it is to be the center of the cluster class. Therefore,  $r$  value is taken as the initial clustering center selection criterion.

(2) Elbow method

The elbow method has a good effect on K value of K-means clustering algorithm. Its core index is Sum of the Squared Errors (SSE):

$$SSE = \sum_{j=1}^K \sum_{x_i \in M_j} |x_i - c_j|^2 \quad (13)$$

Where  $M_j$  represents the j-th cluster;  $x_i$  is the sample point in  $M_j$ ;  $c_j$  is the cluster center of  $M_j$ .

The core idea of elbow method is that the sum of error squares of samples changes with the fineness of sample division, and the finer the sample division, the smaller its value. When K samples the total number of samples, each sample as the cluster center, at this time the sample clustering degree is the highest, but contrary to the original intention of clustering. When K is less than the true number of clusters, the aggregation degree of each cluster will greatly increase with the increase of K value, so the decline of SSE is steep, while when K reaches the true number of clusters, the decline of aggregation degree will slow down with the increase of K value. In other words, the relationship between SSE and K is elbow-shaped, and the K value corresponding to this elbow joint is the true cluster amount of data.

## 4. Optimal Configuration Model of Distributed Generation

### 4.1. Objective Function

Considering the operating economic benefits of distribution companies and environmental factors in the planning cost, the objective function of DG optimal allocation problem is to minimize the annual comprehensive cost, as shown in Equation (14):

$$\min C = \sum_{s=1}^k P(s)(C_I^s + C_{OM}^s + C_P^s + C_L^s - C_S^s - C_E^s) \quad (14)$$

Where  $s$  represents the typical scene of source load;  $P(s)$  is the probability of scenario  $s$ ;  $C_I^s$ ,  $C_{OM}^s$ ,  $C_P^s$ ,  $C_L^s$ ,  $C_S^s$  and  $C_E^s$  are DG investment cost, DG operation and maintenance cost, power loss cost of distribution network, power purchase cost of distribution network from superior power grid, government subsidy cost and environmental cost respectively under scenario  $s$ .

(1) DG Investment Expenses

$$C_I = \frac{r(1+r)^m}{(1+r)^m - 1} \sum_{k=1,2} \sum_{j \in N_{DGk}} (C_{DGk}^I P_{DGkj}) \quad (15)$$

Where  $r$  is the cash discount rate;  $m$  is the planning period

of DG;  $k$  is the type of DG incorporated into the power grid (1 stands for wind power generation; 2 stands for photovoltaic power generation);  $N_{DGk}$  is the node set incorporated into the power grid;  $C_{DGk}^I$  is the unit capacity investment cost of  $k$  types of DG;  $P_{DGkj}^I$  is the installation capacity of the type  $k$  DG of the node  $j$  to be installed.

(2) DG Operation and Maintenance Cost

$$C_{OM} = \sum_{k=1,2} \sum_{j \in N_{DGk}} (C_{DGk}^M E_{DGkj}) \quad (16)$$

Where  $C_{DGk}^M$  is the operation and maintenance cost under the unit capacity of  $k$  types of DG;  $E_{DGkj}$  is the actual annual power generation of type  $k$  DG at node  $j$ .

(3) Distribution network loss cost

$$C_P = C_e P_{Loss} \quad (17)$$

Where  $C_e$  is the on-grid electricity price;  $P_{Loss}$  is the total annual active power loss of the line.

(4) Electricity purchase cost from higher authorities in distribution network

$$C_L = C_e (P_{max} T_{max} - \sum_{k=1,2} \sum_{j \in N_{DGk}} E_{DGkj}) \quad (18)$$

Where  $P_{max}$  is the maximum load in the whole year;  $T_{max}$  is the annual maximum load utilization hour.

(5) Government subsidy expenses

$$C_S = C_f \sum_{k=1,2} \sum_{j \in N_{DGk}} E_{DGkj} \quad (19)$$

Where  $C_f$  is the government subsidy cost of DG unit power generation.

(6) Environmental costs

$$C_E = \sum_{k=1,2} \sum_{j \in N_{DGk}} E_{DGkj} m_g (b_g + e_g) \quad (20)$$

$m_g$  respectively represent the mass of the  $g$  pollution gas emitted by traditional thermal power generation per kilowatt hour, as well as the penalty standard and environmental value of this gas.

### 4.2. Constraint Condition

When planning DG, we need to consider the following four constraints: power flow, voltage, current and DG installation capacity constraints.

(1) Power flow equation constraint

$$\begin{cases} P_i = U_i \sum_{j=1}^n U_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij}) \\ Q_i = U_i \sum_{j=1}^n U_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) \end{cases} \quad (21)$$

Where  $P_i$  and  $Q_i$  are the injected active power and reactive power of Node  $i$  respectively;  $U_i$  is the voltage amplitude of node  $i$ ; The  $\theta_{ij}$  is the phase angle difference between power saving  $i$  and node  $j$ ;  $G_{ij}$  and  $B_{ij}$  are the conductance and susceptance of the branch.

(2) Node voltage constraint

$$U_{i,\min} < U_i < U_{i,\max} \quad (22)$$

Where  $U_{i,\min}$  and  $U_{i,\max}$  are the lower limit and upper limit of the voltage amplitude of node  $i$  in the line respectively.

(3) Branch current constraint

$$I_j \leq I_{j,\max} \quad (23)$$

Where  $I_{j,\max}$  is the maximum allowable current of branch  $j$ .

(4) DG Installation Capacity Constraint

$$\begin{cases} 0 \leq P_{DGi} \leq P_{DGimax} \\ \sum P_{DGi} \leq \varepsilon P_{L, total} \end{cases} \quad (24)$$

Where  $P_{DGi}$  is the capacity of DG access node  $i$ ;  $P_{DGimax}$  is the maximum active power allowed by node  $i$  to access DG;  $P_{L, total}$  is the total active load of the system;  $\varepsilon$  is the penetration rate of system DG connected to power grid.

## 5. Improved Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO) was proposed by Kennedy and Eberhart in 1995, which originated from the study of foraging behavior of birds. It updates the speed and position of particles through individual extremum and group extremum, and finally converges to the optimal solution. PSO is an efficient optimization search algorithm through memory and feedback mechanism [14]. The iterative formulas of particle velocity and position in the iterative process are as follows: (25) and (26), respectively:

$$v_{id}(k+1) = \omega v_{id}(k) + c_1 r_1 [p_{id}(k) - x_{id}(k)] + c_2 r_2 [g_d(k) - x_{id}(k)] \quad (25)$$

$$x_{id}(k+1) = x_{id}(k) + v_{id}(k) \quad (26)$$

Where  $\omega$  is the inertia weight coefficient;  $r_1$ ,  $r_2$  are random numbers between intervals (0, 1);  $c_1$  and  $c_2$  are learning factors, the former refers to the motivation factor of self-learning, while the latter refers to the motivation factor generated by referring to others' learning.

Aiming at the shortcomings of slow convergence speed and easy to fall into local optimum in the optimization process of particle swarm optimization, the improvement of particle swarm optimization is proposed, and the specific improvement measures are as follows:

(1) Adaptive inertia weight factor

Inertia weight factor  $\omega$  is one of the important parameters of particle swarm optimization, which reflects the ability of particles to maintain their original state in the process of optimization. In the early stage of iteration, in order to search for the optimal solution, it is required that the search space of particle swarm should be as large as possible, and the value of inertia weight factor should be larger to make the global search ability of particle swarm strong; At the later stage of iteration, the algorithm is required to have the ability to converge to the optimal solution quickly, so the inertia weight factor should be reduced to enhance the local search ability of particles, and then the whole algorithm can converge to the optimal solution quickly. Therefore, this paper carries out linear dynamic adjustment, so that it can automatically adjust its size in the iterative process to achieve better coordination between global search and local search, thus making the whole algorithm have good dynamic performance. The

specific adjustment method is shown in Equation (27):

$$\omega = \omega_{\max} - \frac{(\omega_{\max} - \omega_{\min})}{k_{\max}} k \quad (27)$$

Where  $\omega_{\min}$  and  $\omega_{\max}$  are the lower limit and upper limit of inertia weight factor respectively;  $k$  is the current iteration number of the algorithm;  $k_{\max}$  is the maximum number of iterations of the algorithm.

(2) Realization of genetic variation operation

Firstly, a certain number of particles are selected according to a certain crossover rate  $p_c$ , and they are regarded as parent particles. Secondly, parts of the parent particles are exchanged to form new individuals, and the parent particles are replaced by offspring particles. According to a certain mutation probability  $p_m$ , some particle dimensions are initialized to improve the population diversity of particle swarm optimization.

(3) Elite retention strategy

In order to avoid the loss of the optimal particles due to genetic mutation operation, the optimal particles are retained and directly inherited to the next generation. The purpose of doing this is to preserve the individual optimal position with high fitness function value, and prevent the "negative optimization" effect of its particle swarm optimization algorithm due to position update and mutation operation.

## 6. Example Analysis

Taking IEEE 33 standard node system as an example, the network topology diagram of IEEE 33 standard node system is shown in Figure1, where 1-33 represents node number, and node 1 is the system balance node [15]. The system reference voltage is 12.66 kV, the three-phase power reference value is 10MVA, the total active load of the network is 3715kW, and the total reactive load is 2300kvar. The range of node voltage is 0.9 ~ 1.1, and the convergence accuracy of power flow algorithm is 10e-7.

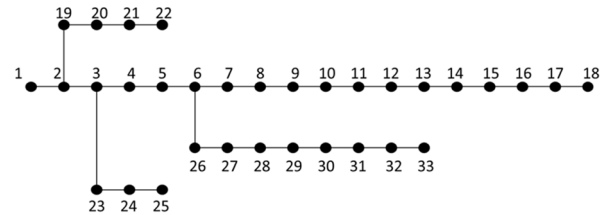


Fig 1. IEEE 33-node distribution network

The types of DG to be selected are wind power generation system and photovoltaic power generation system, and Node 4, Node 7, Node 8, Node 14, Node 18, Node 20 and Node 32 are selected installation nodes of DG. In this paper, the related parameters of DG grid connection are set as follows: the capacity of a single DG is set to 50 kW, the power factor of distributed generation is controlled at 0.85 with the participation of reactive power compensation equipment, the maximum active power installation capacity of DG at each selected installation node is 400kW, and the total penetration rate of DG access is 35%. It is explained here that whether it is a wind turbine generator or a photovoltaic generator, its single unit configuration shall be configured according to the above requirements.

The specific parameters involved in the objective function in the DG optimal configuration model are as follows: The

investment and operation and maintenance costs of wind turbine are respectively  $C_{DG1}^I=1$ ten thousand yuan / kW ,  $C_{DG1}^{OM}=0.33$ ten thousand yuan / kW; The investment and operation and maintenance costs of photovoltaic generator sets are respectively  $C_{DG2}^I=1.3$ ten thousand yuan / kW ,  $C_{DG2}^{OM}=0.2$ ten thousand yuan / kW; Electricity price  $C_e = 0.5$ yuan/kWh; Government subsidy expenses  $C_g = 0.25$ yuan/kWh; DG planning period is 20 years; The discount rate is 0.1.

Considering the environmental benefits brought by DG access to distribution network into the annual comprehensive cost, Table 1 is the emission data of traditional thermal power pollution gas involved in the environmental cost function.

**Table 1.** Emission data of traditional thermal power pollution gas

Types of polluted gases	CO	CO <sub>2</sub>	SO <sub>2</sub>	NO <sub>x</sub>
Emissions (g/kWh)	0.103	633	6.51	2.90
Environmental value (yuan/kg)	0.881	0.019	5.25	6.92
Penalty standard (yuan/kg)	0.138	0.078	0.831	1.71

The probability model parameters of wind speed and illumination intensity are obtained according to the annual wind and light data of a certain place. Because the wind speed and sunlight radiation intensity are different in each period of a year, different parameters are adopted for each month in this paper. See Table 2 for specific parameters. The relevant parameters of the fan are:  $v_{ci} = 3.5$  m/s,  $v_r = 12$  m/s and  $v_{co} = 20$  m/s; The relevant parameters of photovoltaic are:  $I_r = 1000$  W/m<sup>2</sup>

**Table 2.** Parameters of wind and solar probability model in each month

Month	Wind speed		Illumination intensity	
	$k$	$c$	$\alpha$	$\beta$
1	2.24	7.63	2.03	2.50
2	1.94	8.21	2.14	2.63
3	2.34	8.15	1.78	2.42
4	2.14	7.73	2.12	2.47
5	2.21	7.23	2.35	2.42
6	2.04	9.05	2.42	1.95
7	2.56	7.36	2.31	2.06
8	2.12	6.89	2.21	2.49
9	2.31	7.23	2.14	2.48
10	2.16	8.68	1.94	2.6
11	2.09	8.25	2.34	2.50
12	2.12	8.63	2.13	2.48

Then, 8760 hours of initial wind and solar output data are obtained by Latin super-force sampling proposed in chapter 3, and 8760 hours of load rate data are obtained by analyzing historical load data. The source load data are combined into a three-dimensional matrix to form an initial scene of source load. Finally, the improved K-means clustering algorithm proposed in chapter 3 is used to reduce the source load scene. After clustering, 10 typical planning scenes are obtained, and the scene reduction results are shown in Table 3:

**Table 3.** Typical Scenarios

Scene	Fan probability	Photovoltaic probability	Load probability	Scene probability
1	0.247	0.293	0.808	0.091
2	0.012	0.512	0.866	0.130
3	0.426	0.597	0.725	0.074
4	0.117	0.678	0.717	0.119
5	0.544	0.172	0.830	0.081
6	0.975	0.496	0.836	0.150
7	0.968	0.623	0.801	0.103
8	0.649	0.436	0.829	0.046
9	0.721	0.839	0.832	0.102
10	0.977	0.215	0.711	0.104

In this section, three different schemes are set first. Scheme 1: Do not install distributed power; Scheme 2: Install DG on the node to be selected, but do not consider the uncertainty of source load; Scheme 3: Install DG on the node to be selected, considering the influence of source load uncertainty. The improved particle swarm optimization algorithm proposed in chapter four is used to realize the optimization of DG optimal allocation model, and the results of DG planning are shown in Table 4. Note here: 4 (0, 5) in the table means that 0 fans and 5 photovoltaics are installed at Node 4.

**Table 4.** DG planning results of different schemes

Program	DG planning results	Cost (ten thousand yuan)
1	-	989.2
2	4(0,5),8(6,0),14(2,0),32(0,2)	896.14
3	14(6,0),18(4,0),20(3,0),32(4,1)	783.04

According to the analysis in Table 5-4, the comparison scheme 1 does not consider the installation of distributed generation, and the scheme 2 and scheme 3 rationally configure the grid-connected position and capacity of distributed generation, so that after DG is connected to the grid, even after generating certain investment and operation and maintenance costs, the annual comprehensive cost is still lower than that when DG is not connected. This is because DG access is equivalent to reducing the load demand of distribution network nodes, reducing the power transmitted on the line, and then reducing the power purchase cost and network loss cost. Comparing scheme 2 and scheme 3 in the table, it can be seen that after considering the uncertainty, the optimal configuration result of distributed generation is more reasonable. DG merging nodes are roughly distributed at high load points near the end of the line, which reduces the network loss caused by the increase of transmission capacity. At the same time, injecting certain reactive power into the end load improves the voltage level at the end of the line.

The specific costs incurred by the three schemes are compared, and the specific results are shown in Table 5.

**Table 5.** Planning results of different programs

Cost (ten thousand yuan)	Option 1	Option 2	Option 3
$C_I$	0	100.43	107.47
$C_{OM}$	0	84.84	62.57
$C_P$	936.18	775.68	640.30
$C_L$	53.02	40.07	39.19
$C_S$	0	78.75	49.35
$C_E$	0	26.13	17.14
Total cost	989.2	896.14	783.04

By comparing the costs incurred under different schemes in table 5, the following conclusions can be drawn:

(1) After the distributed generation is connected to the grid, the active power loss cost of the distribution network and the power purchase cost from the superior power grid are greatly reduced. This is because after the distributed generation is connected, a certain amount of power is injected into the merged node, which reduces the load demand of the node and its connected nodes on the superior power grid, and also reduces the transmission power of the line. This shows that after the distributed generation is connected, the dependence on the superior power grid is reduced, and the wind-solar grid connection realizes certain self-compensation.

(2) The comprehensive cost of considering uncertainty is lower than that of not considering certainty, and it is more in line with the actual situation, which is consistent with the conclusion that considering uncertainty is more practical in the previous analysis.

At the same time, the influence of different planning results on the voltage distribution of distribution network is analyzed, as shown in Figure 5.2. It can also be seen from the figure that the scheme obtained by considering uncertainty has better optimization results for distribution network voltage distribution.

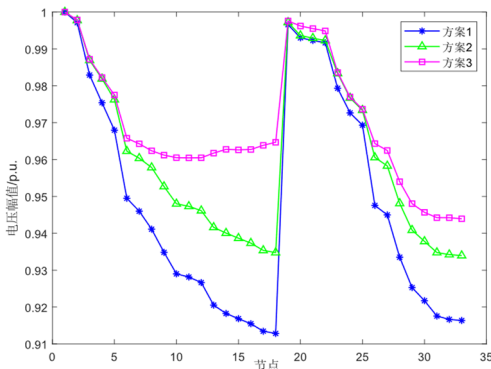


Fig 2. Voltage distribution of distribution network under different planning schemes

In order to prove the superiority and effectiveness of this algorithm in solving DG optimal allocation model, this algorithm, basic particle swarm optimization algorithm and chaotic particle swarm optimization algorithm are simulated on scheme 3 respectively. See chapter 4 for the parameter setting of the algorithm. The obtained DG planning results are shown in table 6, and the convergence characteristics of different algorithms are shown in figure 3.

Table 6. Planning results of different algorithms

	DG planning results	Total cost (10,000 yuan/year)
Basic PSO	14(5,1),18(4,0),32(0,2)	793.42
Chaotic PSO	14(7,0),18(2,1),20(4,0)	786.91
Improved PSO	14(6,0),18(4,0),20(3,0),32(4,1)	783.04

From Table 6, we can get the results of DG planning under different optimization algorithms and the required annual comprehensive cost. The improved particle swarm optimization algorithm proposed in this paper finally gets the least annual comprehensive cost. It is also easy to see from the convergence results of each algorithm curve in Fig. 3 that

the elementary particle swarm optimization falls into local optimum prematurely in iteration, and is not suitable for solving nonlinear DG optimal allocation problem with constraints; Chaotic particle swarm optimization algorithm is slow in convergence speed and does not meet the practical application requirements; The improved particle swarm optimization algorithm proposed in this paper not only effectively avoids the convergence of the algorithm to the local optimum, but also guarantees the calculation speed and accuracy.

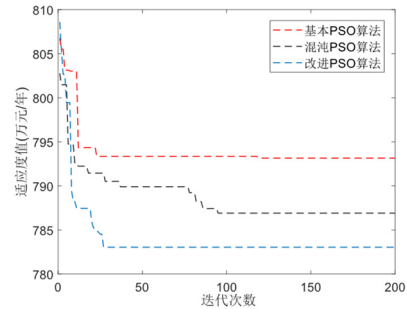


Fig 3. Comparison of convergence characteristics of algorithms

## 7. Conclusion

In this paper, multi-scenario analysis is applied to solve the problem of source load uncertainty in order to reduce the complexity of solving this kind of problem, aiming at the problem of distributed generation output and load forecasting uncertainty in the research of distributed generation optimal configuration. Firstly, the source charge probability model is constructed. Secondly, Latin hypercube sampling is used to get the initial planning scene of source load, and the improved K-means clustering algorithm is used to reduce the scene; Finally, the improved particle swarm optimization algorithm is used to solve the distributed generation grid-connected optimal configuration model. The simulation results show that the economy of the grid-connected distributed generation is optimal under the condition of meeting the operation requirements of distribution network, and considering the uncertainty of source load is more in line with the actual planning requirements. At the same time, by comparing with different algorithms, we can get the superiority of the improved particle swarm optimization algorithm in solving the optimal configuration problem of distributed generation.

## References

- [1] Georgilakis P S, Hatzigrygiou N D. Optimal distributed generation placement in power distribution networks: models, methods, and future research[J]. IEEE Transactions on Power Systems, 2013,28(3): 3420-3428.
- [2] Bernards R,Morren J, Slootweg H. Development and Implementation of Statistical Models for Estimating Diversified Adoption of Energy Transition Technologies[J]. IEEE Transactions on Sustainable Energy, 2018, 9(4):1540-1554.
- [3] D'Adamo, Christian, Jupe, Samuel, Abbey, Chad. Global survey on planning and operation of active distribution networks - Update of CIGRE C6.11 working group activities [P]. Electricity Distribution - Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on,2009.
- [4] ZHANG Shenxi,CHENG Haozhong,XING Haijun, et al.Review of DG planning considering uncertainties for distribution network[J].Electric Power Automation Equipment, 2016, 36( 8) : 1-9.

- [5] Lin Ma, Liu Jianpeng. Multi-objective and multi-type distributed generation planning considering time sequence characteristics and environmental benefits [J]. *Power System Protection and Control*, 2016, 44 (19): 32-40.
- [6] Jiang Huizhe, Zhang Chunyan. Study on location and capacity of distributed units considering renewable energy [J]. *Renewable Energy*, 2021, 39 (09): 1248-1254.
- [7] Chen Jiajun, Jiang Tiezheng, Zhou Yong, et al. Constant capacity and location of distributed generation considering voltage sag [J]. *Power Grid Technology*, 2014, 38 (08): 2244-2249.
- [8] Zhang S, Cheng H, Ke L, et al. Optimal siting and sizing of intermittent distributed generators in distribution system [J]. *IEEE Transactions on Electrical and Electronic Engineering*, 2015, 10(6).
- [9] Mola J H, Barforoshi T, Firouzjaee J A. Distributed Generation Expansion Planning Considering Load Growth Uncertainty: A Novel Multi-Period Stochastic Model [J]. *International Journal of Engineering, Transactions B: Applications*, 2018(3).
- [10] JAMSHID A, TAHER N, RASOUL A A, et al. Scenario-based dynamic economic emission dispatch considering load and wind power uncertainties [J]. *Electrical Power and Energy Systems*, 2013(47):351-367.
- [11] Fu Wenjie, Wang Yuxi, Shen Hongtao, Tao Peng, Wang Shaolin, Li Kangping, Ge Xinxin, Wang Fei. Baseline load estimation method for residential users based on Latin hypercube sampling and scenario subtraction [J]. *Power Grid Technology*, 2022, 46 (06): 2298-2307.
- [12] Chen Z P, Yan Z. Scenario Tree Reduction Methods Through Clustering Nodes [J]. *Computers & Chemical Engineering*, 2018, 109: 96-111.
- [13] Yao Huangjin, Lei Xia, Fu Xinquan, Hu Yi. Clustering analysis of daily load curve based on improved adaptive density peak algorithm [J]. *Power System Protection and Control*, 2022, 50 (03): 121-130.
- [14] YU Tingfang, LIN Zhongda. Application of float genetic algorithms-partially solved combined with punishing function in power plant units commitment problem [J]. *Proceedings of the CSEE*, 2009, 29(2):107-112.
- [15] BARAN M E, WU F E. Network reconfiguration in distribution systems for loss reduction and load balancing [J]. *IEEE Transactions on Power Delivery*, 1989, 4(2):1401-1407.