Article

Distributed Generation Planning in Active Distribution Networks Based on Multi-Scene Analysis

Sitong Lv¹, Jianguo Li^{1,*}, Yongxin Guo² and Zhong Shi¹

¹ College of Electrical Engineering, Shanghai Dianji University, Shanghai 20136, China; sitonglv987@hotmail.com (S.L.); shizhongsd@126.com (Z.S.)

² Training Center of Jilin Province, State grid Corporation of China, Jilin 130062, China; Nsjiane888@163.com

*Correspondence: lijiang@sdju.edu.cn; Tel.: +86-199-2126-6168

Abstract: In recent years, distributed generation technology has developed rapidly. Renewable energy, represented by wind energy and solar energy, has been widely studied and utilized. In order to give full play to the advantages of Distributed Generation (DG) and meet the challenges after power grid access, Active Distribution Network (ADN) is considered as the future development direction of traditional distribution network because of its ability of active management. Nowadays, multi-scenario analysis is widely used in the research of optimal allocation of distributed power supply in active distribution network. Aiming at the problems that may arise when using multi-scenario analysis to plan DG with uncertainties in large-scale scenarios, a scenario reduction method based on improved clustering algorithm is proposed. The validity of the scene reduction method is tested, and the feasibility of the method is verified. At present, there are few studies on the optimal allocation of DG in ADN under fault state. In this paper, comprehensive safety indicators are introduced. Considering the timing characteristics of DG and the influence of active management mode, a bi-level programming model is established, which aims at minimizing the investment of annual life cycle and the removal of active power. The bi-level model is a complex mixed integer non-linear programming model. A hybrid algorithm combining cuckoo search algorithm and primal dual interior point method is used to solve the model. Finally, through the simulation of the IEEE-33 node system, the superiority of the scenario reduction method and the comprehensive security index used in this paper to optimize the configuration of DG in ADN is verified.

Keywords: active distribution network; distributed generation; multi-scene analysis; Scene reduction; improved clustering algorithm; bi-level programming; comprehensive security index

1. Introduction

With the rapid development of global economy in the 21st century, the demand for energy in various countries is also increasing. According to the current consumption rate, the oil stock can only be used for 40 years, the natural gas village stock can only be used for 60 years, and the coal stock can be used for 200 years. The energy crisis caused by over-exploitation of non-renewable energy restricts the rapid development of economy. The excessive use of traditional energy has also caused serious environmental pollution. Under such circumstances, it is necessary to improve the development and utilization of existing non-renewable energy resources, develop and utilize new environmentally friendly energy resources, and make necessary supplements and innovations to the existing energy system. Therefore, Distributed Generation (DG) has received extensive attention and

support [1-3]. It is of great significance to vigorously develop distributed energy and give full play to the role of DG in the power grid [4-6].

In the early stage of DG planning, it was assumed that the output of DG was constant [7, 8], regardless of its uncertainty and timing characteristics. For wind power and photovoltaics, their output is affected by many factors, including geographical location and climate environment. Without considering its timing characteristics, DG can not fully play its role in balancing adjustment and dynamic complementarity, which will also lead to inaccurate DG access location and capacity in the planning process.

At this stage, a lot of research work [9-11] has been done on the output of distributed generation considering uncertainties, including DG location and capacity, demand side response, network reconfiguration, power and voltage quality, etc. A control algorithm based on improved amplitude adaptive notch filter (AANF) is proposed for generation management of different energy sources in autonomous micro-grid [9]. The voltage-controlled oscillator (VCO) less phase-locked loop (PLL)-based control of voltage source converter was presented to improve power quality [10]. A new attempt of utilizing the sunflower optimization (SFO) algorithm in solving the problem of optimal power flow (OPF) in the field of power systems was introduced in order to optimize the generating units' fuel cost under the system constraints [11].

Scenario analysis has been widely used in the optimal configuration of DG at this stage [12-14]. Based on multi-scenario technology, multi-stage planning considering DG reactive power output constraints is studied [12]. An active distribution network expansion planning model under multi-stage and multi-load scenarios is proposed. The model considers the application of new generation distributed generation and the construction of feeder in planning layer, the utilization of distributed generation in operation layer and the reconfiguration of distribution network including micro-grid [13]. A multi-objective, multi-level model is proposed for active distribution system expansion planning with high-penetration renewable energy sources (RESs) and energy storage systems (ESSs) was proposed. The multi-scenario tools and K-means clustering are adopted to deal with the uncertainties and capture the time-variable nature of RESs and load demand [14]. In the process of scenario analysis, all historical data are used for multi-scenario analysis. Although the result of analysis is comprehensive, it will lead to a sharp increase in computational complexity and difficulty in solving problems. Therefore, it is necessary to reduce large-scale scenarios. Multi-scenario analysis chooses some scenarios according to the uncertainty of DG and load through specific rules, and then carries on the characteristic analysis according to the selected scenarios, and completes the optimization planning of DG. This method considers the uncertainty sufficiently, but how to select an effective scene is the difficulty of this method. DG programming model based on multi-scenario analysis is a typical mixed integer non-linear programming model, and its solution is a typical NP-hard problem. Intelligent algorithm and its improved algorithm are usually used to solve the problem. This kind of algorithm works well in small-scale systems, but when the scale of the problem increases, its computational efficiency and convergence are affected.

There are two kinds of DG programming methods considering uncertainty in ADN: single-level planning model and bi-level planning model. The single-level planning method of DG in ADN is simple in model and fast in solving. It has been widely used in the planning process of DG. The main disadvantage of the single-level programming model is that it can not deal with discrete variables, so the discrete constraints can not be considered in the planning process, resulting in low accuracy of the model. The bi-level planning model separates planning from operation, and solves the planning and operation problems respectively by using the relevant iterative strategy.

In the existing literature, the optimal allocation of DG in active distribution network involves consideration of energy storage configuration and reactive power compensation, active operation strategy, demand side management and network reconfiguration, and life cycle theory [15-17]. A novel selection strategy for restricted operation scenarios based on the shadow price is proposed to reduce the complexity of scenario selection and ease the computational burden substantially in virtue of the decoupling of ADN operation and planning [15]. A methodology for distribution

networks reconfiguration by controlling number, sharing, size, and location of DG units is proposed [16]. A local voltage control strategy of DGs with reactive power optimization based on a kriging metamodel is proposed [17]. The above models take into account the impact of active management mode on DG planning and operation, but for the failure state, the ADN active control to maintain the safe and stable operation of islands is not considered.

The main work of this paper includes:

1. Two types of DGs are modeled indefinitely. The characteristics of annual, seasonal, continuous multi-day and typical day scenarios are analyzed. The uncertainty and time series characteristics of two types of DGs are analyzed.
2. In order to fully reflect the timing characteristics of two types of intermittent DGs, and to avoid the difficulty of computation caused by large-scale data in multi-scenario analysis, a scenario reduction method based on improved clustering algorithm is proposed, and its effectiveness is verified by using wind power generation scenarios.
3. In view of the fact that ADN is in a state of failure when optimizing the allocation of DG in ADN at the present stage, considering the influence of active management mode, a comprehensive safety index is introduced, and a bi-level programming model aiming at the minimum annual life cycle investment and the minimum amount of active power removal is established. Aiming at the complex mixed integer non-linear programming model, a hybrid algorithm of cuckoo search algorithm and primal dual interior point method is used to solve the problem.
4. Scene reduction method based on improved clustering algorithm proposed in this paper is used to reduce two kinds of DGs and validity test is carried out. Through the simulation of the IEEE-33 node system, the planning schemes considering the active management mode and not considering the active management mode under the constraints of the comprehensive security indicators, the planning schemes considering the constraints of the comprehensive security indicators and not considering the constraints of the comprehensive security indicators under the active management mode, and the progressing of the planning schemes under the three scenarios set are presented. Corresponding comparisons were made.

2. Study on the Characteristics of Typical Intermittent Distributed Generation

2.1. Uncertain Modeling of Typical Intermittent Distributed Generation

2.1.1. Uncertainty Model of Wind Power Generation

Wind energy resource is one of the most abundant and mature intermittent distributed power sources. Wind power is affected by many factors, which can be roughly divided into atmospheric characteristics, terrain characteristics, wind power, behavior index, other indexes and geographical conditions. Wind power has more influence and is greatly influenced by the change of natural environment, which leads to the obvious uncertainty of wind power output.

In this paper, Weibull distribution with two parameters, which has the best application effect in engineering practice, is adopted. Its probability density function [18, 19] is:

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

where k is the shape parameters, c is the scale parameters, and v is the wind speed.

Scale parameter c and shape parameter k can be determined by equation 2.

$$k = \left(\frac{\sigma_w}{E_w}\right)^{-1.086}, c = \frac{\bar{v}}{\Gamma(1 + k^{-1})} \quad (2)$$

where σ_w is the variance of v ; $\bar{v}$ is the average value of v ; E_w is the generation capacity; Γ is the gamma function.

When the cut-in wind speed is reached, the fan starts to exert its power. With the increase of wind speed, the output of fan will also increase. When the wind speed is too high, in order to protect the fan, the fan equipment will be automatically removed. Therefore, the output power of the fan can generally be expressed by a piecewise function, as shown in equation 3:

$$P_{WTG}(v) = \begin{cases} 0, & 0 \leq v \leq v_{ci} \\ \frac{P_r(v - v_{ci})}{v_r - v_{ci}}, & v_{ci} < v \leq v_{cr} \\ P_r, & v_{cr} < v \leq v_{co} \\ 0, & v > v_{co} \end{cases} \quad (3)$$

where v_{ci} is the cut-in wind speed, v_{cr} is the rated wind speed, v_{co} is the cut-out wind speed, and P_r is the rated active power.

2.1.2. Uncertainty Model of Photovoltaic Power Generation

Solar energy is the most abundant of all renewable energy sources. Photovoltaic power generation has remarkable flexibility, and its installation is simple and flexible. It is an important form and component of distributed power generation. At the same time, with the continuous development of photovoltaic power generation and the increase of investment, the cost of photovoltaic power generation has decreased significantly in recent years. The continuous increase of grid-connected photovoltaic power generation has also brought many impacts on the current grid scheduling and control, and this impact will continue to increase with the increase of grid-connected photovoltaic power generation. Similar to wind power, photovoltaic power generation has obvious randomness and uncertainty.

In this paper, Beta distribution in probability model is used to describe the uncertainty of illumination intensity. Its probability density function [20, 21] is shown in equation 4.

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

where S is the illumination intensity; S_{max} is the maximum illumination intensity; α and β is the two parameters corresponding to Beta distribution.

α and β can be calculated by the expected μ and variance σ^2 of illumination intensity over a certain period of time, as shown in equation 5 and 6:

$$\alpha = \mu \left[\frac{\mu(1 - \mu)}{\sigma^2} - 1 \right] \quad (5)$$

$$\beta = (1 - \mu) \left[\frac{\mu(1 - \mu)}{\sigma^2} - 1 \right] \quad (6)$$

The output power of photovoltaic power generation equipment will gradually increase with the increase of illumination intensity until it reaches the rated power. The relationship between output power and illumination intensity can be expressed by equation 7.

$$P_{PVG} = \begin{cases} P_{PVG,r} S / S_r, & S \leq S_r \\ P_{PVG,r}, & S > S_r \end{cases} \quad (7)$$

2.2. Characteristic Analysis of Typical Intermittent Distributed Generation Scene

Wind power output is affected by wind speed uncertainties, while photovoltaic output is mainly affected by illumination intensity uncertainties. Therefore, both wind power output and photovoltaic output have obvious volatility and randomness. This section will analyze the scene characteristics of two kinds of DGs and study their inherent characteristics.

2.2.1. Characteristic Analysis of Wind power Output Scene

According to the uncertain modeling results of wind speed, the corresponding wind power output curve can be obtained. Figure 1 is the annual wind power output variation curve of a certain area. In order to observe its variation more intuitively, the fitting curve is made. Figure 2 is the average wind power output variation curve in different seasons, Figure 3 is the continuous

172 multi-day wind power output variation curve drawn randomly, and Figure 4 is the typical wind
173 power output curve.

174 **Figure 1.** Annual wind power output curve

175 **Figure 2.** Wind power output curve in different seasons

176 **Figure 3.** Continuous multi-day wind power output variation curve

Figure 4. Typical wind power output curve

The characteristics of the above wind power output curves are analyzed and summarized as follows:

1. Random. As shown in Figure 1, the fitting curve of the annual wind power output curve can be clearly observed. For the hourly statistics of wind power output, the output at each moment shows obvious uncertainty.
2. Intermittence. The wind speed has obvious intermittence, and is affected by the cutting-in speed and cutting-out speed of wind turbines, so there are some points in the curve where the output of wind turbines is zero. The point where the output of these fans is zero may be due either to the failure to reach the cut-in wind speed or to the fact that the fan has been removed because the cut-out wind speed has been reached, thus the output of wind power is not continuous.
3. Seasonal variation characteristics. As shown in Figure 2, wind power output has a certain seasonal variation characteristics. Wind power output is relatively large in autumn and relatively small in summer, and the difference is relatively obvious. At the same time, the output curve of each season is quite different from the typical solar output curve, and the typical solar output curve can not well reflect the characteristics of output variation in each season.
4. Time series characteristics and similarity. It can be seen from the continuous multi-day variation curve and typical sunrise curve that the wind power output has obvious time series characteristics and similarity. The output of wind power is large at night and small at daytime, which has good peak regulation characteristics. The sunrise curve in continuous time has certain similarity, which means that it can reduce the scene effectively.

From the above analysis, it can be seen that the wind power output has obvious time series characteristics, and the wind power output has obvious differences with different seasons and different periods of the day. Typical daily method can not adequately express all the information contained in the annual output curve of wind power. At the same time, its contribution has some similarities, which means that it can reduce the necessary scene.

2.2.2 Characteristic Analysis of Photovoltaic Output Scene

According to the uncertain modeling results of illumination intensity in the previous section, the corresponding photovoltaic output curve can be obtained. Figure 5 is the annual photovoltaic output variation curve and its fitting curve of a region. Figure 6 is the photovoltaic output curve of different seasons. Figure 7 is the continuous multi-day photovoltaic output variation curve. Figure 8 is the typical solar output variation curve.

Figure 5. Annual wind power output curve

Figure 6. Photovoltaic output curve in different seasons

Figure 7. Continuous multi-day photovoltaic output curve

Figure 8. Typical photovoltaic output curve

The characteristics of the above photovoltaic output curves are analyzed and summarized as follows:

1. Periodicity. From the solar photovoltaic output curve, it can be clearly observed that the photovoltaic output has obvious periodicity, rising from the morning until about noon to reach the maximum photovoltaic output, and declining in the afternoon. This periodicity is evident in all curves.
2. Seasonal variation characteristics. Photovoltaic output is mainly affected by light intensity. Generally speaking, the photovoltaic output is higher in summer than in winter because of the highest illumination intensity.
3. Time series characteristics and similarity. From the continuous multi-day curve and typical daily curve, it can be seen that the photovoltaic output changes obviously with time. On sunny days, the continuous multi-day curves sampled randomly have obvious similarities. This also means that the effort scenario can be reduced.

From the above analysis, it can be seen that the photovoltaic output has obvious time-series characteristics, and the photovoltaic output has obvious differences with different seasons and intra-day periods. Typical daily method can not adequately express all the information contained in the annual output curve of photovoltaic. At the same time, its contribution also has certain similarity, which means that it can reduce the scene.

3. Scene reduction method based on improved clustering algorithm

Cluster analysis is a common method for scene analysis in DG planning process. Cluster analysis grouped the same or similar scenarios in DG and load output scenarios, and obtained the classes of similar elements. Clustering algorithm has been widely used in data analysis. For different sets, different classes are needed, so the clustering algorithm has been improved from the corresponding aspects in the specific application at this stage. In this section, we will introduce the common clustering algorithms. In view of the shortcomings of clustering algorithms and the set of scenarios used in this paper, we propose a kind of improved clustering algorithm.

3.1. Improved clustering algorithm

According to the scenario characteristics of two kinds of DGs, it can be seen that the number of scenario sets in the whole year is large and has certain similarities. Through the introduction of related scene reduction methods, this paper chooses clustering algorithm to reduce the annual scene set. Intra-class similarity and inter-class difference are the criteria for evaluating clustering algorithm. In order to test the clustering results effectively, this paper chooses BWP index to test the clustering results to judge the reliability of clustering scenarios. This index can also give the optimal number of clustering that traditional clustering algorithm can not give.

Let $K = \{X, R\}$ be the clustering space, $X = \{x_1, x_2, \dots, x_n\}$, Assuming that n objects are eventually clustered into class c , the minimum distance $b(j, i)$ of the sample i in class j is the minimum average distance from the sample to all other types of samples, such as equation 8:

$$b(j, q) = \min_{1 \leq k \leq c, k \neq j} \left(\frac{1}{n_k} \sum_{p=1}^{n_k} \|x_p^{(k)} - x_i^{(j)}\|^2 \right) \quad (8)$$

Where $x_i^{(j)}$ is the sample i in class j ; $x_p^{(k)}$ is the sample p in class k ; n_k is the number of samples in class k ; and $\|\cdot\|^2$ is the square Euclidean distance.

The intra-class distance $w(j, i)$ of the sample i in class j is the average distance from the sample to all other samples in class j , such as equation 9:

$$w(i, j) = \frac{1}{n_j - 1} \sum_{q=1, q \neq i}^{n_j} \|x_q^{(j)} - x_i^{(j)}\|^2 \quad (9)$$

Where $x_q^{(j)}$ is the sample q in class j , and $q \neq i$, n_j is the number of samples in class j .
 $baw(j, i)$ is the sum of the minimum class-to-class distance and the intra-class distance of the sample:

$$baw(j, i) = b(j, i) + w(j, i) \quad (10)$$

$bsw(j, i)$ is the difference between the minimum class-to-class distance and the intra-class distance of the sample:

$$bsw(j, i) = b(j, i) - w(j, i) \quad (11)$$

The index $BWP(j, i)$ of the sample i in class j is the ratio of the clustering distance to the clustering distance of the sample:

$$BWP(j, i) = \frac{bsw(j, i)}{baw(j, i)} = \frac{b(j, i) - w(j, i)}{b(j, i) + w(j, i)} \quad (12)$$

According to the definition of BWP index, the bigger the value of BWP index is, the better the clustering result is. The average value of BWP index can reflect the quality of clustering results. When the average value of BWP index is the largest, k is the optimal clustering number. $avgBWP(k)$ is used to represent the average value of BWP indices of all samples when data set D is clustered into k class, and k_{opt} is used to represent the optimal clustering number.

$$avgBWP(k) = \frac{1}{n} \sum_{j=1}^k \sum_{i=1}^{n_j} BWP(j, i) \quad (13)$$

$$k_{opt} = argmax_{2 \leq k \leq n} \{avgBWP(k)\} \quad (14)$$

BWP index is used to improve the maximum and minimum distance K-means algorithm, and the best clustering result is determined according to the BWP value. The improved algorithm steps are as follows:

1. Choosing a center according to the maximum and minimum distance criterion described above
2. Clustering according to K-means clustering method based on maximum and minimum distance
3. Calculate the BWP value of the clustering result and turn to step 2
4. Comparing the BWP value of clustering results, the K value of clustering results is the best clustering number when the BWP value is maximum
5. Clustering results corresponding to the maximum output BWP value

It should be pointed out that when the scene reduction method based on improved clustering algorithm is used to reduce the specific scene, the reduced scene with larger BWP value can be selected according to the actual scene reduction requirement rather than the maximum value. Choosing the reduced set corresponding to the high K value can make use of the time series characteristic of retaining the original scene set to a greater extent.

3.2. Validity test

3.2.1 Scene Reduction Process Based on Improved Clustering Algorithm

Scene reduction is the process of classifying and merging objects to be clustered. According to the results of past research and the analysis of scenario characteristics in Chapter 2, the wind farm scenic set is divided into four scenarios in spring, summer, autumn and winter. The photovoltaic scenic set is divided into 12 scenarios in spring, summer, autumn, winter and three weather types: cloudy, sunny and rainy. When scene reduction of DG is carried out, scene reduction is carried out with day as the basic unit of clustering. Assuming that the total number of individual scenarios $n(1, 2, \dots, N)$ is N . A single scenario has T -period scenario data. The data contained in all scenarios can be represented by matrix $N * T$. By improving the clustering algorithm to merge the same kind of scene into K scenes, the reduced scene data can be represented by $K * T$. Typical scenes

obtained after reduction have the same temporal characteristics as the original scenes, so as to ensure the temporal characteristics of the scenes before reduction. Scenario reduction of two types of DGs and loads is carried out by using the above method. This paper takes the wind power generation scenario as an example to test the effectiveness of the improved clustering algorithm proposed.

The scene reduction process based on improved clustering algorithm is shown in Figure 9.

Figure 9. Reduced flow chart

3.2.2. Validity Test of Wind Power output

In this paper, wind power generation scenarios are taken as an example to verify the effectiveness of the reduced scenarios. According to the four seasons of spring, summer, autumn and winter, all scenes are divided. The improved clustering algorithm proposed in this paper is used to reduce the partitioned scenes, and the BWP value of clustering results is calculated to select the optimal result.

Figure 10 shows the change curve of BWP value with K value after clustering.

Figure 10. BWP Value Change Curve

According to the change curve of BWP value, the BWP value is the largest in spring, autumn and winter scenarios when $k = 2$, and in summer, when $k = 3$, the BWP value is the largest. But in order to reflect the temporal characteristics of the original scene to the greatest extent, this paper chooses the case where the K value is relatively large and the number of scenes is relatively large. Taking spring as an example, when $k = 2$ and $K = 5$, the K value is larger. Choosing $k = 5$ here, the scenarios to be reduced are divided into five categories. The output curves of these five scenarios are given below, as shown in Figure 11.

Figure 11. Spring Output Curve at $k=5$

According to the law of large numbers, the corresponding probability of each scene at $k=5$ is shown in Table 1.

Table 1. The probability of typical temperatures throughout one year

Season	Scene reduction				
	1	2	3	4	5
Spring	0.57	0.21	0.07	0.12	0.03

In order to reflect the relationship between the reduced scene and the original scene more clearly, scene No. 5 is selected, and two scenes are randomly selected from the reduced scene set for comparative analysis, as shown in Figure 12.

Figure 12. Scene reduction contrast graph

According to the reduced output, nearly half of the wind power output in Figure 11 is as shown in Scenario 1. In the other scenarios, the wind power output shows obvious peak reversal characteristics. In the second chapter of this paper, the intra-day variation characteristics of wind power output obtained from scenario characteristics analysis are better reflected in different reduced scenarios. Through the verification of BWP value, and from Figure 12, we can see that the reduced scene obtained by the improved clustering algorithm in this paper has better coincidence with the original scene and can better reflect the temporal characteristics of the original scene.

3.2.3. Scene Reduction for Two Kinds of Intermittent DG

The scene reduction method based on improved clustering algorithm is used to reduce the output curves of two kinds of DGs, and the validity test is carried out. The results of scene reduction are given directly here.

$$\delta_i(t) = P_{wav}(t)/P_T \quad (15)$$

After reducing and normalizing the wind generator scenes divided by seasons, the typical daily scenes are shown in Figure 13.

Figure 13. Wind power curve after scene reduction

After reducing and normalizing the photovoltaic power generation scenarios divided by season and weather, the typical daily scenarios are shown in Figure 14 to Figure 17, respectively.

336 **Figure 14.** Photovoltaic curve after scene reduction in spring

337 **Figure 15.** Photovoltaic curve after scene reduction in summer

338 **Figure 16.** Photovoltaic curve after scene reduction in autumn

339 **Figure 17.** Photovoltaic curve after scene reduction in winter

340 **4. Bi-level Programming Model for Distributed Generation Optimal Configuration**

341 In the existing DG planning model, the influence of active management mode on distribution
342 network planning is generally considered. However, the ADN active control is generally not
343 considered in the model to maintain the safe and stable operation of islands under the failure state.
344 This paper establishes an upper-level programming model aiming at minimizing the annual

comprehensive cost. In the upper-level planning, ADN operation strategy in fault scenario is considered, and the comprehensive safety index is introduced and converted into upper-level constraints. To minimize the amount of active power cut-off, this paper adopts the following three kinds of active management measures [22].

1. Distributed generator output control
2. Switching of reactive power compensation
3. Adjustment of on load transformer

4.1. Upper level programming model

The upper-level programming model considers DG layout and installation capacity planning. The objective function is to minimize the annual life cycle investment cost.

$$\min C_1 = C_I + C_{OM} + C_P + C_{AM} + C_L \quad (16)$$

The specific expressions of each cost are as follows:

1. DG Equivalent Investment Annual Cost

$$C_I = \left(\sum_{i=1}^{N_{bus}} C_{WTG,i} S_{WTG,i} + \sum_{i=1}^{N_{bus}} C_{PVG,i} S_{PVG,i} \right) \frac{r(1+r)^y}{(1+r)^y - 1} \quad (17)$$

Where N_{bus} is the number of nodes in the distribution network, r is the discount rate, y is the life span of DG for 20 years, $C_{WTG,i}$ and $C_{PVG,i}$ are fixed investment costs of unit capacity of wind power and PV installed at the i node respectively, $S_{WTG,i}$ and $S_{PVG,i}$ are the rated capacity of wind power and PV installed at the i node respectively.

2. DG annual operation and maintenance costs

$$C_{OM} = \sum_{n=1}^{12} p_n \times 365 \left(\sum_{t=1}^{24} \left(\sum_{i=1}^{N_{bus}} C_{WTG,i} E_{WTG,in}(t) + \sum_{i=1}^{N_{bus}} C_{PVG,i} E_{PVG,in}(t) \right) \right) \quad (18)$$

Where P_n is the scenario probability of the n scenario, $C_{WTG,i}$ and $C_{PVG,i}$ are the operation and maintenance costs of the wind power and the photovoltaic unit electricity received by the i node, $E_{WTG,in}(t)$ and $E_{PVG,in}(t)$ are the wind power received by the i node and the photovoltaic unit electricity generated during the t period of the n typical day respectively.

3. Annual Electricity Purchase Cost

$$C_P = \sum_{n=1}^{12} P_n \times 365 \left(\sum_{t=1}^{24} E_{nt} P_t \right) \quad (19)$$

Where E_{nt} is the t time of n typical days to buy electricity from a higher power grid. P_t is the unit cost of operators purchasing electricity from a higher power grid.

4. DG annual active management cost

$$C_{AM} = \sum_{n=1}^{12} p_n \times 365 \left(\sum_{t=1}^{24} \left(\sum_{i=1}^{N_{bus}} C_{AWTG,i} E_{WTG,in}(t) + \sum_{i=1}^{N_{bus}} C_{APVG,i} E_{PVG,in}(t) \right) \right) \quad (20)$$

Where $C_{AWTG,i}$ and $C_{APVG,i}$ are the active management costs of the wind power and the photovoltaic at the i node respectively.

5. Network Loss Cost

$$C_L = \sum_{n=1}^{12} P_n \times 365 \left(\sum_{t=1}^{24} Q_{ntL} P_{ntL} \right) \quad (21)$$

Where Q_{ntL} is the net loss of the n typical day t period, P_{ntL} is the unit network loss cost.

The constraints are:

1. DG Installation Capacity Limitation

$$\begin{aligned} 0 &\leq R_{WTGi} \leq R_{WTGmax} \\ 0 &\leq R_{PVGi} \leq R_{PVGmax} \end{aligned} \quad (22)$$

Where R_{WTGi} and R_{PVGi} are the wind capacity and PV capacity node i respectively. R_{WTGmax} and R_{PVGmax} correspond to the maximum access capacity of DG respectively.

2. Capacity Limitation of DG Total Installation

$$\sum_{i=1}^{N_{WTG}} R_{WTGi} + \sum_{i=1}^{N_{PVG}} R_{PVGi} \leq R_{DGmax} \quad (23)$$

Where R_{DGmax} is the maximum installed capacity.

3. Constraints of Comprehensive Safety Indicators

$$C_{csi} = \frac{1}{2} \left(\frac{1}{N_{TS}} \sum_{n=1}^N \sum_{t=1}^T \sum_{s=1}^S C_{n,t,s} + \min \{ C_{n,t,s} \} \right) \quad (24)$$

Where $C_{n,t,s}$ is the index of safe power supply rate of branch s of the n scenario at t period.

$$C_{n,t,s} = 1 - \frac{\sum_{t=1}^{t_T} \sum_{y \in \varphi_{ntf}} \gamma_y S_{n,t,y} \Delta T_{n,t,y}}{\sum_{t=1}^{t_T} \sum_{i \in \varphi_{ntl}} \gamma_y S_{n,t,y} \Delta T_{n,t}} \quad (25)$$

Where t_T is the period of system failure elimination. φ_{ntf} is the n scenario of t period outage load set. γ_y is the grade factor of Class y load. $S_{n,t,y}$ is the Capacity of Class y Load. $\Delta T_{n,t,y}$ is the outage time of y -load in the n th scenario at t period. φ_{ntl} is t period load set for the n scenario.

4.2. Lower level programming model

The lower level planning model mainly considers the operation constraints related to the operation of distribution network. At this stage, DG access to power grid costs higher. In order to maximize the utilization of DG, the lower level objective is to minimize the amount of active power cut-off of DG, and its expression is as follows:

$$\min C_2 = \sum_{n=1}^{12} P_n \times 365 \left(\sum_{t=1}^{24} P_{cnt} \right) \quad (26)$$

The constraints are:

1. Node Power Balance Constraints

$$\begin{aligned} P_{ci,i,t,n} + P_{co,i,t,n} - P_{WTG,i,t,n} - P_{PVG,i,t,n} = \\ U_{i,t,n} \sum_{j=1}^{N_{bus}} U_{j,t,n} (G_{ij} \cos \theta_{t,n,ij} + B_{ij} \sin \theta_{t,n,ij}) \end{aligned} \quad (27)$$

$$\begin{aligned} Q_{ci,i,t,n} + Q_{co,i,t,n} - Q_{WTG,i,t,n} - Q_{PVG,i,t,n} - Q_{c,i,t,n} = \\ U_{i,t,n} \sum_{j=1}^{N_{bus}} U_{j,t,n} (G_{ij} \sin \theta_{t,n,ij} - B_{ij} \cos \theta_{t,n,ij}) \end{aligned} \quad (28)$$

Where $P_{WTG,i,t,n}$ and $P_{PVG,i,t,n}$ are the active output of the t time of the n scenario, respectively. $P_{ci,i,t,n}$ and $P_{co,i,t,n}$ are the active power of residential and commercial loads at the first time of t in the first n scenario, respectively. $Q_{WTG,i,t,n}$ and $Q_{PVG,i,t,n}$ are the reactive power of DG at the t time of the n scenario, respectively. $Q_{ci,i,t,n}$ and $Q_{co,i,t,n}$ are the reactive power of resident load reactive power and commercial load at the t time of the n scenario, respectively. $Q_{c,i,t,n}$ is reactive power supplied by reactive power compensation device. $U_{i,t,n}$ and $U_{j,t,n}$ are the voltage amplitude of node i

and the voltage amplitude of node j at the t time node of the n scenario, respectively. $\theta_{t,n,i,j}$ is the phase difference between node i and node j of t in the n scenario, respectively.

2. Node Power Balance Constraints

$$U_{imin} \leq U_i \leq U_{imax} \quad (29)$$

Where U_i is node voltage. U_{imin} and U_{imax} are the minimum voltage values and maximum voltage values allowed by node i respectively.

3. Branch power constraints

$$S_i \leq S_{imax} \quad (30)$$

Where S_i is the apparent power of branch L . S_{imax} is the limit of branch transmission capacity.

4. DG output control constraints

$$P_{imin} \leq P_i \leq P_{imax} \quad (31)$$

Where P_{imin} and P_{imax} are the minimum active power output of node i and the maximum active power output of distributed generation respectively.

5. Constraints of reactive power compensation

$$Q_{imin} \leq Q_i \leq Q_{imax} \quad (32)$$

Where Q_{imin} and Q_{imax} are the minimum value of reactive power compensation device of node i and the maximum value of reactive power compensation device.

6. Regulation constraints of on-load tap-changer

$$T_{imin} \leq T_i \leq T_{imax} \quad (33)$$

Where T_i is the tap position of transformer i . T_{imin} and T_{imax} are the tap values of transformer i and the maximum tap value of i respectively.

5. Bi-level Programming Model Solving Algorithms

The solution of mixed non-integer programming problem is a common problem in the process of optimal allocation of distributed power supply, and its essence is NP-hard problem. At present, heuristic algorithm and deterministic algorithm are the main solving methods. Particle swarm optimization (PSO), genetic algorithm (GA) and related improved algorithms are widely used in DG programming. In this paper, cuckoo search algorithm is used to solve the upper model, and the lower model is solved by the original dual interior point method.

5.1. Cuckoo Search Algorithms

Cuckoo search algorithm was first proposed in 2009 by Professor Xin-She Yang and S. Deb of Cambridge University. CS algorithm can efficiently search the optimal solution of the problem by simulating the parasitic brooding of some species of cuckoo. At the same time, CS also uses the relevant Levy flight search mechanism.

In the process of cuckoo reproduction, the nest location of cuckoo's offspring is uncertain. In the process of simulating its search for bird's nest, three principles need to be recognized:

1. Cuckoos lay only one egg at a time of reproduction, and then they choose their nests arbitrarily for hatching and rearing.
2. The most suitable nest will be extended to the next generation of reproduction in a randomly selected set of options.
3. The total number of nests available N is a fixed value, and the probability that the original owner of the nest has $P_a \in [0,1]$ can identify a non-self-laid bird's egg.

Based on these three principles, the path and location of cuckoo nest selection are determined by equation 34:

$$x_i^{e+1} = x_i^e + \alpha * L(\lambda), i = 1, 2, \dots, n$$

(34)

438 Where, x_i^e is the position of the i nest in the selection of the e generation; $*$ is point-to-point
439 multiplication; α is the step size control in the process of choosing nests for cuckoos, which obeys
440 the normal distribution; $L(\lambda)$ is the path through which Levy searches bird's nest arbitrarily, and
441 $L(d, \lambda) \sim s^{-\lambda} (1 < \lambda \leq 3)$, d is the random step obtained by Levy's flight.

442 *5.2. Bi-level Programming Model Solving Process*

443 The detailed flow chart of solving multi-objective bi-level programming model by CS algorithm
444 and PDIPM method is shown in Figure 18. Among them, G is the number of iterations.

Figure 18. Bi-level planning model solving process

6. Examples and Analysis of planning Results

6.1. Introduction of Examples

Figure 19. IEEE-33 node distribution system.

The system voltage is 12.66 kV, total active load is 3.715 MW, total reactive load is 2.300 MW, Weibull distribution parameter $k = 2.30$, $C = 8.92$, wind power access cost 6500 yuan/kW, operation and maintenance cost 0.3 yuan/ kW·h, environmental protection subsidy 0.1 yuan/ kW·h, rated lighting intensity of photovoltaic generator is 1 kW/m², shape parameter B of beta distribution is 0.85, photovoltaic access The cost is 10,000 yuan/kW, the operation and maintenance cost is 0.2 yuan/ kW·h, the rated capacity of a single distributed power supply is 125 kW, the equipment life is 20 years, and the discount rate is 0.1. The repair time of N-1 fault is 4 hours, and the comprehensive safety index value is set to 0.5. Wind power installation nodes are 5, 7, 11, 12. The photovoltaic installation nodes are 20 and 23.

6.2. Result analysis

On the premise of considering comprehensive safety constraints, the installation types, capacity and cost of distributed generators are compared when the active management mode and the active management mode are taken into account in the lower level planning model. The results are shown in Table 2. Active power removal of lower target in each season is shown in Table 3.

Table 2. Optimal allocation schemes with and without active management.

Pattern	D G	Installation node	Installation capacity /kW	Investment cost/¥	Operation cost /¥	Electricity purchase cost/¥	Network loss cost/¥
active management mode		5	125	1964700	771500	1711400	99800
	W	7	125				
	G	11	250				
		12	250				
	P	20	125				
Non active management mode	V	23	250	1554500	1126400	3212000	96500
		5	125				
	W	7	250				
	G	11	250				
		12	375				
	P	20	125				
	V	23	250				

The installation capacity of DG configuration scheme is 1125 kW when considering active management mode, which is less than 1375 kW when not considering active management mode. On the premise that the comprehensive security index of the system is taken into account to ensure the safe and stable operation of the system, the active management mode can reduce the amount of DG access. This is because the active management method can actively cooperate with the DG operation according to the actual operation of the system, the tap-in of on-load voltage regulator and the switching of reactive power compensation, so that the system can achieve the optimal operation state at each moment. At this stage, DG access costs are higher, and reducing DG access can reduce the annual life cycle investment costs.

Table 3. The excision amount of active output with and without active management.

	DG removal volume in different seasons (MW·h)			
	Spring	Summer	Autumn	Winter
Active management mode	8.21	3.24	12.15	22.17
Non active management mode	22.50	15.63	33.21	41.54

According to Table 3, the amount of active power removal in different seasons is smaller when subjective management mode is taken into account than when active management mode is not taken into account. In summer when the load level is high, the DG output connected to the distribution network can be basically absorbed. Therefore, the introduction of active management measures improves the utilization rate of distributed energy, and plays a more obvious role in improving the phenomenon of "wind abandonment" and "light abandonment".

On the premise of considering active management, the installation type, capacity and cost of distributed generators with comprehensive safety index and without comprehensive safety index are compared. The results are shown in Table 4.

Table 4. Optimal allocation schemes with and without comprehensive safety index.

Pattern	DG	Installation node	Installation capacity /kW	Investment cost/¥	Operation cost /¥	Electricity purchase cost/¥	Network loss cost/¥	Comprehensive Safety Indicators
With Comprehensive Safety Indicators	W G P V	5	125	1964700	771500	1711400	99800	0.50
		7	125					
		11	250					
		12	250					
		20	125					
Without Comprehensive Safety Indicators	W G P V	23	250	1056400	863500	6001700	112400	0.36
		5	125					
		7	375					
		11	250					
		12	125					

Without considering the constraints of comprehensive security indicators, the access capacity of DG is smaller than that of the scheme considering comprehensive security indicators. This is because

the upper model in this paper aims at minimizing the annual cost. Without considering the security constraints, the scheme will tend to have fewer DGs with higher access costs.

The scheme with comprehensive security index constraints, whose index value is 0.50, and the scheme without comprehensive security index constraints, whose index value is 0.36. When the system is not connected to DG. Without considering the constraints of comprehensive security indicators, the access of DG has little effect on improving the island operation capability of ADN under fault condition. At present, the main function of DG access to distribution network is to improve the operation of the system. Therefore, considering the comprehensive safety indicators, it has obvious effect on improving the operation status of the system in the case of failure in the planning process.

On the premise of considering both comprehensive security indicators and active management mode, the economic and technical comparison of DG optimal allocation under the annual time series scenarios, typical day scenarios and reduced scenarios is made. Full-time scenarios are selected as benchmarks to test the planning schemes under the other two scenarios. The annual average of each period is selected as the data of typical day scenes. A summary of the three plans is shown in Table 5.

Table 5. Comparison of three kinds of scene set planning schemes.

Pattern	DG	Installation node	Installation capacity /kW	Annual life cycle investment cost/¥	Active power excision/ MW·h	Computing time/s
Annual Sequence Scene	WG	5	125	5123600	48.26	1803
		7	125			
		11	375			
		12	250			
	PV	20	125			
		23	250			
Typical Day Scene	WG	5	125	6145600	31.48	8
		7	125			
		11	250			
		12	125			
	PV	20	125			
		23	125			
Scene Reduction	WG	5	125	4547400	45.77	60
		7	125			
		11	250			
		12	250			
	PV	20	125			
		23	250			

From the comparison results of three scenario planning schemes, it can be concluded that: Most directly, when using reduced scene set for DG planning, the computational time is reduced from about 1803 seconds to only about 60 seconds, and the computational efficiency has been significantly improved. Although the time spent in DG planning using reduced scene sets is slightly longer than that using typical day scenes, the calculation accuracy is higher. In this paper, DG with fixed capacity is used to access the corresponding nodes. Under this assumption, the DG access capacity in reduced scenarios is the same as that in annual sequential scenarios, and the typical daily scenario with average value is smaller than the other two scenarios. Compared with the typical Japanese method, the annual comprehensive cost and the amount of effective removal of reduced scenes are closer to the annual time series scenes. In DG planning process, the access capacity and location of the reduced scene are close to that of the year-round sequential scene. In

summary, the scene reduction method based on improved clustering algorithm proposed in this paper has better retention effect for the original scene time series data, and the economic and technical indicators can basically accurately reflect the sequence status of the scene before reduction.

7. Conclusions

The main work of this paper is to use scenario reduction multi-scenario analysis method to optimize the allocation of distributed power in active distribution network. The main problem of this paper is how to take account of both computational efficiency and accuracy in DG planning process by using multi-scenario analysis method, and how to fully consider the active management mode of ADN and the operation status under failure state. Based on uncertainty modeling and scenario reduction, a DG bi-level programming model for active distribution network is constructed. Through the simulation and verification of the IEEE-33 node system, the specific contents are as follows:

1. The uncertain modeling of two kinds of DGs is introduced. The scene characteristics of DGs are analyzed. The analysis shows that both kinds of DGs have obvious uncertainties and time series characteristics. Therefore, the typical scenes composed of average method or peak-valley difference maximum sunrise force can not effectively reflect the characteristics of DG, and the two kinds of DG's output scenes are similar and can be effectively reduced.
2. Aiming at the shortcomings of clustering algorithm for large-scale scenarios, a scene reduction method based on improved clustering algorithm is proposed and applied to DG planning process. The effectiveness of this method is verified through wind power generation scenarios.
3. Through specific examples, the planning schemes considering active management mode and not considering active management mode under the constraints of comprehensive safety indicators, the planning schemes considering comprehensive safety indicators constraints and not considering comprehensive safety indicators constraints under the active management mode, and the planning schemes under three scenarios are compared accordingly. The results show that the active management mode can reduce DG access and annual comprehensive cost while maintaining the stable operation of the system; the introduction of comprehensive security indicators can improve ADN operation capability in the DG planning stage; the planning schemes under three scenarios set show that using the scenario reduction method proposed in this paper. The scene constructed by this method has a good effect on preserving the original scene, and can take account of both computational efficiency and accuracy.

Author Contributions: All authors contributed to the research. Conceptualization, Sitong Lv; Data curation, Yongxin Guo; Formal analysis, Sitong Lv and Zhong Shi; Project administration, Jianguo Li; Writing – original draft, Sitong Lv; Writing – review & editing, Sitong Lv.

Funding: This work is supported by the Scientific Research Projects of Shanghai Science and Technology Commission (Grant Nos.17DZ1201200).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A. Parameters of 33-bus distribution network.

Line data			
Access Rd		Branch impedance/ Ω	
First spot number	End point number	Resistance	Reactance
1	2	0.922	0.047
2	3	0.493	0.2511
3	4	0.366	0.1864
4	5	0.3811	0.1941

5	6	0.819	0.707
6	7	0.1872	0.6188
7	8	0.7114	0.2351
8	9	1.03	0.74
9	10	1.044	0.74
10	11	0.1966	0.065
11	12	0.3744	0.1238
12	13	1.468	1.155
13	14	0.5416	0.7129
14	15	0.591	0.526
15	16	0.7463	0.545
16	17	1.289	1.721
17	18	0.732	0.574
2	19	0.164	0.1565
19	20	1.5042	1.3554
20	21	0.4095	0.4784
21	22	0.7089	0.9373
3	23	0.4512	0.3083
23	24	0.898	0.7091
24	25	0.896	0.7011
6	26	0.203	0.1034
26	27	0.2842	0.1447
27	28	1.059	0.9337
28	29	0.8042	0.7006
29	30	0.5075	0.2585
30	31	0.9744	0.963
31	32	0.3105	0.3619
32	33	0.341	0.5302

References

1. Mehigan, L.; Deane, J.P.; Gallachoir, B.P.O.; Bertsch, V. A review of the role of distributed generation (DG) in future electricity systems. *Energy* **2018**, *163*, 822-836

2. Anaya, K.L.; Pollitt, M.G. Going smarter in the connection of distributed generation. *Energy Policy* **2017**, *105*, 608-617.

3. Allan, G.; Eromenko, I.; Gilmartin, M.; Kockar, M.; Mcgregor, P. The economics of distributed energy generation: A literature review. *Renew. Sustain. Energy Rev.* **2015**, *42*, 543-556.

4. Shi, Z.H.; Liang, H.; Huang, S.J.; Dinavahi, V. Multistage robust energy management for microgrids considering uncertainty. *IET. Gener. Transm. Dis.* **2019**, *13*, 1906-1913

5. Chen, G.; Lewis, F.L.; Feng, N.; Song, Y.D. Distributed optimal active power control of multiple generation systems. *IEEE Trans. Ind Electron.* **2015**, *62*, 7079-7090.

6. Tan, Y.S.; Wang, Z.Y. Incorporating unbalanced operation constrains of three-phase distributed generation . *IEEE Trans. Power Syst.* **2019**, *34*, 2449-2452.

7. Ziari, I.; Ledwich, G.; Ghosh,A. Optimal allocation and sizing of DGs in distribution network. In Proceeding of the IEEE Power and Energy Society General Meeting, Minneapolis , MN, USA, 25-29, July 2010; pp. 1-8

8. Othman, M.M.; El-Khattam, W.; Hegazy, Y.G.; Abdelaziz, A.Y. Optimal placement and sizing of distributed generators in unbalanced distribution systems using supervised big bang-big crunch method. *IEEE Trans. Power Syst.* **2015**, *30*, 911-919.

9. Kewat, s.; Singh, B. Modified amplitude adaptive control algorithm for power quality improvement in multiple distributed generation system. *IET. Power.Eletrons.* **2019**, *12*, 2321-2329.

10. Giri, A.K.; Arya, S.R.; Maurya, R.; Babu, B.C. VCO-less PLL control-based voltage-source converter for power quality improvement in distributed generation system. *IET. Electr. Power. App.* **2019**, *13*, 1114-1124.

11. Shaheen, M.A.M.; Hasanien, H.M.; Mekhamer, S.F.; Talaat, H.E.A. Optimal power flow of power systems including distributed generation units using sunflower optimization algorithm. *IEEE Access*. **2019**, *7*, 109289-109300
12. Zou, K.; Agalgaonkar, A.P.; Muttaqi, K.M. Distribution System Planning With Incorporating DG Reactive Capability and System Uncertainties. *IEEE Trans. Sustain. Energy*. **2012**, *3*, 112-123.
13. Shen, X.W.; Shahidehpour, M.; Zhu, S.Z.; Han, Y.D.; Zheng, J.H. Multi-stage planning of active distribution networks considering the co-optimization of operation strategies. *IEEE Trans. Smart Grid*. **2018**, *9*, 1425-1433.
14. Li, R.; Wang, W.; Xia, M.C. Cooperative planning of active distribution system with renewable energy sources and energy storage systems. *IEEE Access*. **2017**, *6*, 5916-5926.
15. Zhao, L.; Huang, Y.X.; Dai, Q.D.; et, al. Multistage active distribution network planning with restricted operation scenario selection. *IEEE Access*. **2019**, *7*, 121067-121080.
16. Shafik, M.B.; Chen, H.K.; Rashed, G.I.; et, al. Adequate topology for efficient energy resources utilization of active distribution networks equipped with soft open points. *IEEE Access*. **2019**, *7*, 99003-99016.
17. Li, P.; Zhang, C.C.; Fu, X.P.; et, al. Determination of local voltage control strategy of distributed generators in active distribution networks based on kriging metamodel. *IEEE Access*. **2019**, *7*, 34438-34450.
18. Du, E.S.; Zhang, N.; Kang, C.Q.; Kroposki, B.; Huang, H.; Miao, M.; Xia, Q. Managing wind power uncertainty through strategic reserve purchasing. *IEEE Trans. Power Syst.* **2017**, *32*, 2547-2559.
19. Wang, Y.; Zhou, Z.; Botterud, A.; Zhang, K.F. Optimal wind power uncertainty intervals for electricity market operation. *IEEE Trans. Sustain. Energy*. **2018**, *9*, 199-210.
20. Kaloudas, C.G.; Ochoa, L.F.; Marshall, B.; Majiothia, S.; Fletcher, I. Assessing the feature trends of reactive power demand of distribution networks. *IEEE Trans. Power Syst.* **2017**, *32*, 4278-4228.
21. Wan, C.; Lin, J.; Song, Y.H.; Xu, Z.; Yang, G.Y. Probabilistic forecasting of photovoltaic generation: An efficient statistical approach. *IEEE Trans. Power Syst.* **2017**, *32*, 2471-2472.
22. Zhang, J.T.; Cheng, H.Z.; Wang, C. Technical and economic impacts of active management on distribution network. *Int. J. Elec. Power*. **2009**, *31*, 130-138.