Optimization of Wind Power Configuration in Distribution Network Based on Scenario Clustering and Power Flow Entropy

Ming-ze ZHANG1, Yi-chao HUANG1 and Min WANG2,*

1State Grid Shanghai Economic Research Institute 200120, Shanghai, China
2College of Energy & Electrical Engineering, Hohai University, 211100, Nanjing, China
*Corresponding author

Keywords: Distribution network, Fuzzy C mean clustering, Scenario analysis, Power flow entropy.

Abstract. When the wind power is connected to the distribution network, the access points and the access capacity of each point in the distribution network will also affect the stability of the distribution network. Therefore, based on the fuzzy C clustering of wind speed, this paper optimizes the location and capacity of wind power in the distribution network by using the proposed entropy index of power flow which reflects the power flow equilibrium degree of the system. Firstly, typical wind speed and output scenarios are selected through scenario clustering, and then the flow entropy index reflecting system stability is adopted to optimize the location and capacity of wind power. The improved IEEE33 was used to verify the method.

Introduction

With the development of energy technology, it is imperative to connect renewable energy such as wind power to distribution network. However, due to the inherent uncertainty of wind power, it is bound to affect the stability of distribution network. Therefore, how to plan the access location of wind power in distribution network and the access capacity of wind power in each access location is very important [1,2].

Since the randomness of wind power output is strong, the randomness of wind power generation needs to be analyzed before planning and optimization. A probability distribution can be used to describe the randomness of variables comprehensively. Since the randomness of different meteorological variables is not the same, they must correspond to different theoretical distribution models. The results show that the wind speed probability distribution curve can be fitted by a statistical model. Common fitting models include Rayleigh model, log-normal distribution model, Gamma distribution model and two-parameter Weibull distribution model [3]. However, in the process of calculating the probability distribution function, parameter estimation is difficult. At the same time, the accuracy of the probability function is difficult to guarantee because of too many uncertain factors. Scenario analysis can exactly solve the above problems. The scenario set describes the probability of the possible occurrence of uncertain events in the future. Since the probability measure of random events can be observed, the scenario set can comprehensively reflect the occurrence of full probability scenarios [4,5].

Too many scenarios will result in too much computation. Therefore, the scene needs to be clustered. Fuzzy C Mean (FCM) is a classification based Fuzzy clustering method. By describing the membership of samples to different categories, this algorithm can objectively cluster, which can overcome the defects of traditional clustering algorithm of either/or, and occupies an important position in Fuzzy clustering [6].

Since the access of wind power will inevitably affect the power flow distribution of the system, the power flow distribution of the system also affects the stability of the system. Therefore, this paper proposes to use the power flow entropy index to quantify the power flow distribution of the system, and to analyze the optimal node and capacity of system access to wind power [7,8].
To sum up, in the process of selecting the location of the wind power distribution network and determining the node capacity, this paper firstly adopts the method of fuzzy C clustering to analyze the uncertain wind speed scenario. After obtaining the typical scenario, based on the theory of entropy, the flow entropy of the system which can quantitatively describe the imbalance of power flow of system is taken as the optimization index to optimize the access location of wind power and the access capacity of each location.

**Clustering of Scenarios**

**Scenario Clustering based on Fuzzy C**

In practical engineering application, it is not difficult to find the development law of random events through many data statistical analysis. The continuous probability function is discretized into several samples by means of sampling discretization. Each sample is called a "scenario". The content of each scenario consists of two parts: the value of the random variable \( w \) and the probability \( p \) of the value. This method is called the scenario method for solving optimization problems involving random variables, so as to obtain the probability of random events in each possible situation. The greater the probability of the scenario, the greater the impact on the target event, thus forming the scenario analysis method.

In order to obtain enough scene sets, many samples are needed, which leads to an exponential increase in the size of the final scene set, resulting in "dimension disaster", which is not conducive to analysis and calculation. Scene reduction is to use a more concise set of scenes to represent the original set of scenes, so scene reduction is also known as scene simplification. Therefore, this paper adopts fuzzy c-means algorithm to cluster renewable energy scenarios. The objective function of the fuzzy c-means is:

\[
J(U,V) = \min \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^2 d_{ij}^2
\]

where, \( u_{ij} \) represents the membership degree of element \( i \) for element \( j \), and \( d_{ij} \) represents the distance between element \( i \) and element \( j \).

Constraint conditions:

\[
u_{ij} \in [0,1] \quad 1 \leq j \leq n, 1 \leq i \leq c
\]

\[
\sum_{i=1}^{c} u_{ij} = 1
\]

when \( \forall i \in j, u_{ij} = 0 \),

\[
v_i = \frac{\sum_{j=1}^{n} u_{ij}^m \cdot x_j}{\sum_{j=1}^{n} u_{ij}^n}
\]

where \( v_i \) represents the \( i \)th clustering center.

**Evaluation Index of Clustering Algorithm**

Silhouette coefficient is an evaluation method of clustering effect. It was first proposed by Peter J. Rousseeuw in 1986. It can be used to evaluate the influence of different algorithms or different operation modes of algorithms on clustering results based on the same original data. The size of
Silhouette coefficient can reflect the degree of subordination of each sample to the clustering center. The larger the value is, the higher the membership degree is. Its expression is shown below.

\[ S(i) = \frac{\min(b) - a}{\max(a, \min(b))} \]  

(4)

where \( a \) represents the normalized distance between the \( i \)th point and the points of the same category; \( b \) represents the normalized distance from the points of different categories. The value range of contour value \( S(i) \) is \([-1, 1]\). The larger \( S(i) \) is, the more reasonable the classification is.

**Power Flow Entropy**

**Introduction to Entropy Theory**

German physicist R Clausius first proposed the concept of "entropy", which was used to describe the chaotic degree of material movement. He clarified that the size of entropy was only related to the start and stop state of the object under study. Therefore, the definition of entropy and the relationship between entropy change could be used to judge the degree of brittle evolution of the system.

When a given system has multiple operating states and the probability of state \( X_i \) is \( P(X_i) \) \((i=1,2,\ldots,n)\), the entropy value of the system is defined as:[8]

\[ H = -C \sum_{i=1}^{n} P(X_i) \ln P(X_i) \]  

(5)

where, \( C \) is a constant and \( n \) is the total number of possible states of the system. The higher the entropy of the system, the lower the order and the lower the stability of the system. Conversely, the higher the degree of order of the system, the higher the degree of stability.

**Power Flow Entropy**

Since entropy can measure the degree of chaos and disorder of the system, we can use entropy to describe the equilibrium degree of power flow distribution in the power system, and then study the robustness of the system.

Load ratio of branch can reflect the state characteristics of a single branch. By analyzing the load rate distribution of each branch, the robustness of the system can be reflected.

If the maximum transmission capacity of the \( i \)th line is \( F_{i,\text{max}} \) and the actual transmission capacity of line is \( F_{i,\text{act}} \), then the load rate \( \rho_i \) of line is:

\[ \rho_i = \frac{F_{i,\text{act}}}{F_{i,\text{max}}} \]  

(6)

\( U \) is the arithmetic constant sequence of branch load rate after segmentation, \( U = [U_1, U_2, U_3, \ldots, U_m] \), and \( m \) is the number of segments divided. \( l_k \) is used to represent the number of lines whose load rate is in the interval \((U_k, U_{k+1}]\). The probability of the number of lines in different load rate intervals is obtained as follows:

\[ P(k) = \frac{l_k}{\sum_{k=1}^{m} l_k} \]  

(7)

where, \( P(k) \) is the proportion of the number of lines that are in \((U_k, U_{k+1}]\) to the total number of lines. Then, according to formula (1) to formula (3), the power flow entropy of the system is obtained as:
\[ H = -C \sum_{1}^{n-1} P(k) \ln P(k) \] (8)

The higher the power flow entropy, the higher the inconsistency of the system line load rate. Too scattered line load rate will increase the possibility of system instability after disturbance. Therefore, when the distribution network needs to be reconstructed, the power flow entropy index can be taken as the index to optimize the reconstructed network.

**Example**

Since the volatility of renewable energy is strong, the mean and variance of wind speed are used as the indexes of wind speed scenario clustering in this paper. Sampling data of wind speed in a certain area were selected with a sampling interval of 15 minutes. The clustering results of typical scenarios are shown in Figure 1.

![Figure 1](image1.png)

Figure 1. The Silhouette coefficient value of wind speed clustering results.

According to the contour value coefficient, it can be found that the daily wind speed curves cluster into three types and the clustering effect is better. The probability of clustering center curves in three scenarios is 25.71%, 15.71% and 58.58%, respectively. The scene with a large proportion is selected as the main influencing factor for analysis. Its clustering centers are as follows:

![Figure 2](image2.png)

Figure 2. The clustering center of wind speed.

After the typical scenario of wind speed is selected, the wind speed is converted into power according to the output model of the wind turbine, and the location and capacity of the wind turbine are set by the IEEE33 node model. The network structure diagram is shown in Figure 3. The total wind power capacity is 1000kW, and the unit wind turbine capacity is 20kW.

![Figure 3](image3.png)

Figure 3. IEEE33 node network structure diagram.
The entropy index can reflect the balance degree of the power flow in the system and then reflect the stability margin of the system. Therefore, the power flow entropy index used to optimize the location and capacity of the wind power in the distribution network. According to the particle swarm optimization algorithm, the final optimization results are shown as follows.

| Location | Number of wind turbines | Power flow entropy |
|----------|-------------------------|--------------------|
| 2        | 10                      | 0.1312             |
| 19       | 20                      |                    |
| 25       | 20                      |                    |

**Summary**

Due to the instability of wind power, it will affect the power flow of distribution network during the connection process. In this paper, based on the fuzzy C scenario clustering of wind speed, the main scenarios affecting the planning and optimization of wind turbines are selected to optimize the location selection and capacity selection of wind turbines in the distribution network. The power flow entropy, which reflects the power flow equilibrium degree in the distribution network, is used as the main optimization index. Finally, the optimization results are obtained.

**Acknowledgement**

This work is supported by Science and Technology Project Funding of State Grid Corporation Limited (Contract No: 52096017000C)

**References**

[1] Kang Longyun, Guo Hongxia, Wu Jie, et al. Characteristics of distributed generation system and related research issues caused by connecting it to power system, J. Power System Technology. 11(2010)51-59.

[2] Ma Huimeng, Li Be, Li Jianlin, et al. Analysis on factors sensitive to capacity configuration of battery energy storage system suitable for centralized renewable energy sources, J. Power System Technology, 38(2014) 328-334.

[3] Wu Weibing, Zhou Songlin, Xing Shicha. Correlative Wind Speed Simulation Based on Copula Theory and Time Series Reconstruction Technology, J. Proceedings of the CSU-EPSA, 31(2019)66-72.

[4] Peng Chunhua, Yu Rong, Sun Huijuan. An Identification Model for Self-organized Criticality of Power Grids Based on Power Flow Entropy, J. Electric Power Automation Equipment, 10(2015)58-65.

[5] Yang Shuqiang, Guo Li, Liu Jiaoyang, Zhao Zongzheng, Pan Jing, Rong Xiuting. Photovoltaic Access Planning Based on Scenario Compression and Voltage Regulation Strategy, J. Automation of Electric Power Systems ,15 (2018) 31-38+215-218.

[6] Dong Zihui, Lin Lingxue. Dynamic Reconfiguration of Distribution Network Based on Improved Fuzzy C-means Clustering of Time Division, J. Power System Technology, 7(2019)2290-2305.

[7] Li Meicheng, Mei Wenming, Liu Yongqiang, Zhang Lingkang, Song Dunwen, Zhao Chong. Accurate Identification Method of Brittle Branches in Power Grid Based on Improved Load Flow Entropy Indexes, J. Power System Technology, 3(2019)1026-1033.

[8] Cao Yijia, Wang Guangzeng, Cao Lihua, Ding Lijie. An Identification Model for Self-organized Criticality of Power Grids Based on Power Flow Entropy, J. Automation of Electric Power Systems, 7(2011)1-6.