Big Data Analysis for Effective Management of Power Distribution Network

Xi Chen,[1*] Yuan La,[1] Ji-Guang Zhao,[2] Wei Zhang,[2] and Ting-Cheng Chang[3]

[1]China Southern Power Grid Company Limited, No.106, Fengze East Road, Guangzhou 510000, China
[2]Digital Grid Research Institute of China Southern Power Grid,
No. 11, Zhongmian Road, Guangzhou 510000, China
[3]College of Information Engineering, Guangzhou Panyu Polytechnic, 1342 Liang Road, Guangzhou 510000, China

To find a way to manage power distribution networks efficiently, we researched the use of big data analysis and established a model with mathematical functions to assess the benefit, risk, and economy of the power supply in a power distribution network. The necessary data were collected from the sensors in the network and analyzed with an algorithm using the particle swarm optimization (PSO) method. The powers from wind and solar energy were adopted as distributed power generation (DG) sources. The result of this study showed that the position of the access of the DG to the network is important as it affects the benefit and risk of the power supply for the network. We tested three different connections of the DG to the network, which had a 10% difference in the maximum power supply in the network. Along with the appropriate position of the DG access, the consideration of the risk assessment and the risk-taking also had a significant effect on the efficient management of the network. The model with the power supply risk function (R-PS) required a fourfold higher power supply from the DG, yielding a higher power supply (11%) and overall benefit (44%) than those without the risk function. The degree of risk-taking also affected the management of the network as the result revealed that power supply management with high risk-taking needed less power from the DG (14%), less power supply (2%), and had one-third less overall benefit than those with low risk-taking. We expect the method and results in this study to provide a model for the effective management of a power distribution network with power from DG sources.

1. Introduction

Power grids are becoming 'intelligent' with the development of big data analysis technology. Recent distribution networks of energy are highly interactive and coupled with other technologies such as Internet of Things (IoT). This makes the structure of the distribution network more complicated, meaning that the efficiency and coverage must be combined with big data to ensure the safe and economic operation of the network. In the network, the allocation and capacity of distributed power generation (DG) are important from the perspective of ensuring the stable power supply of the network.[1] The DG must ensure the

---

*Corresponding author: e-mail: chenxi197405@163.com

security and stability of the power supply of the network. Real-time access of the DG requires the appropriate selection and position in the network, along with optimization for the economy of the network with multiple loads.[2] However, previous studies[1,2] did not consider the real-time characteristics of the DG and changes in the network power capability. Thus, Xu and An established an objective function to guide DG access from the perspective of network loss and benefits,[3] and Teng *et al.* evaluated the optimization of the reactive power in a distribution network with DG.[4] The methods with DG enable the effective operation and the reduced loss of power in the distribution network.[5] A conventional cost estimation considers the failure rate and reliability index of a distribution line with line switches.[6,7] Owing to the differences between a transmission network and a distribution network, the $K \, (N - 1 + 1)$ reliability criterion is usually used to estimate the cost of line switches. However, previous studies did not consider the cost of the switches in a distribution network with DG.

Researchers have proposed a DG grid connection based on network acceptance[8] and adjusted the level and distribution of loads in the network by adjusting a photovoltaic inverter.[9] Yang proposed a method of analyzing the capacity of DG acceptance based on the flexibility of the distribution network.[10] With the method, a hierarchical and multilevel analysis was performed, and data in distribution networks were processed to plan the related distribution network.[11] As the development of sensor technologies and wireless sensor networks is creating huge databases in power distribution networks with DG, several methods of evaluating the importance of big data in power grids have been proposed[12] and the benefit of DG to power distribution networks has been proved.[13] To plan a power grid economically, the positions of DG must be chosen to distribute power reasonably within the capacity of the grid. It is also necessary to determine how to use big data to integrate multilevel data and optimize the DG access to improve the economy of the power supply for structuring the network. This ensures the sufficient capacity of the power supply and the reliability of the distribution network.

The above studies all used various data analysis methods to provide more reasonable suggestions for the access of DG to the network. In practice, however, the location of DG access to a distribution network is determined by the power user, the output power of DG is random, and the importance of various types of DG is different for different measures of the power grid operation, such as power quality, node voltage, etc. In general, the impact of DG on a distribution network is comprehensive. A distribution network obtains data through various sensor elements and analyzes it through big data analysis technology. Therefore, comprehensive benefits can be used to reflect the impact of DG access, so as to guide DG access to the distribution network.

With this background, we propose in this paper a new evaluation method for distribution networks based on big data analysis. This method first guarantees the power supply capacity of the distribution network. In order to evaluate the risk degree and actual benefits of the DG to the distribution network, we use the structure and economic big data to carry out a hierarchical analysis of the distribution network, and then evaluate the economic benefits of the network. In addition to securing the capacity, to assess the degree of risk and practical benefits of the DG for a distribution network of a power supply, we used structural and economic big data for the hierarchical analysis of the network. Then, we assessed the economic benefits of the

distribution network. To do this, we evaluated the capacity of the power supply of the network, estimated the risk function and the maximum benefit, and determined the optimal size of the DG that is connected to the distribution network. The results ensure the safety and economic viability of the distribution network and provide guidelines for operating distribution networks.

2. Power Supply of Distribution Network

Power supply affluence refers to sufficient capacity to supply enough load for a power distribution network. Such affluence reflects the relationship between the reliability and economy of the distribution network. A sufficient power capacity is necessary to sustain the safety of a network. Figure 1 shows the adequacy of a distribution network within the boundaries in which the power load is available. Within the safety range between boundaries one and two, the larger the power generation (load), the greater the power supply affluence in the system. Figure 1 also implies that, depending on the initial conditions of the system and the agreed boundaries, the affluence has two aspects:

(1) When the power supply is between boundary one and the *x*-axis, the capacity of the power supply in the network, the network structure, and other factors have the ability to afford an increase in the power load. Thus, the larger the ratio of the power supply (generation) to the power load, the stronger the network is.

(2) When the power supply is between boundary two and the *x*-axis, the capacity of the network no longer handles the demanded load, transmission, and supply of power. In this case, affluence is not satisfied, and the amount of load (transaction) that must be disconnected from the network needs to be determined within the acceptable limit by the network.[14]

In Fig. 2, the optimal operating point of the network is assumed to be S_n, whose corresponding degree of affluence is D_n. The probability of the occurrence of S_n is defined as P_n. Under normal circumstances, it is usual to determine the maximum and minimum operating points of the network with deterministic criteria. Therefore, the final operating range of the network is

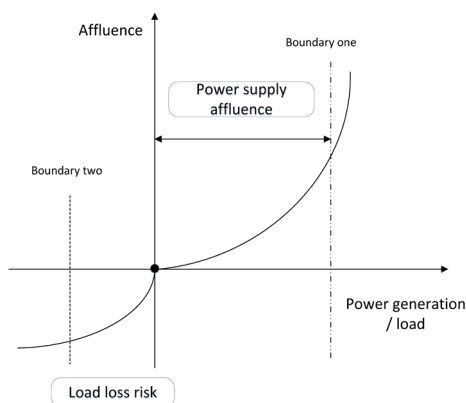$$\{\min(D_n), \max(D_n \times p_n)\}. \tag{1}$$

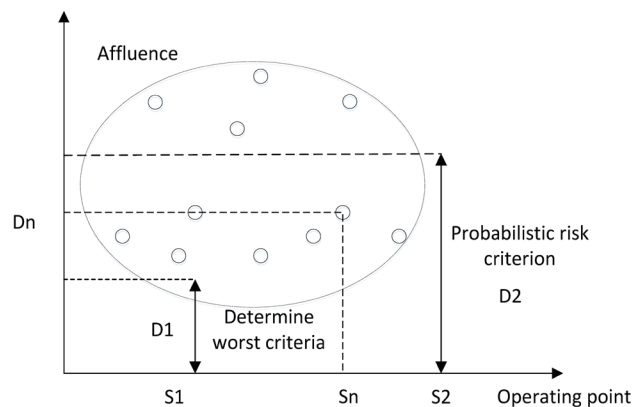Fig. 1.   Adequacy of a power distribution network.   Fig. 2.   Adequacy evaluation diagram.

To determine the working capacity of the system more accurately, the operating range must consider not only the adequacy of each operating point but also the occurrence probability of each operating point.[14] The adequacy of operating points is judged with the decision-making criteria for the most serious incident without considering the economy of the network operation and the occurrence probability of operating points. When the affluence is determined conservatively in general, the economy of the network operation is not good. However, if the probability evaluation takes into account the adequacy of the power distribution network and the possible scale of a certain operation point, the range of the economic network operation becomes more reasonable and practical.

3. Power Supply Capacity

A distribution network has an economic power supply capacity with an adequate power supply.[15] As DG cannot be determined when its volatility is large, the risk of the deficiency of the power in the network is high when DG is connected to the distribution network. Therefore, the power supply risk of DG must be considered to determine the capacity of the power supply in the distribution network. The power supply risk is represented by the probability of the power supply at different sizes. The risk probability of the power supply with connected DG is calculated as

$$f(s_A) = \frac{N(s < s_A)}{N_S},\tag{2}$$

where $N(s < s_A)$ is the number of times the power supply is generated at a certain level, N_S is the total number of samples, and $f(s_A)$ is the consequence of the degree of risk in practice. The greater the risk, the more serious the possible consequence (the loss of the DG power supply). The analysis of the risk requires several functions to be defined: a power supply benefit function (B-PS), a power supply risk function (R-PS), and an economic power supply function (E-PS). All functions of the distribution network with DG reflect the economic capacity of the power supply in the network. The following are the expressions for the three functions.

3.1 Power supply benefit function

With a larger power supply, the economic benefit increases as users obtain more available power. For cross-sectional transmitted power,[16] the B-PS is defined as

$$B_{ps}(x) = xT_{obv}C_{uit}f(x),\tag{3}$$

where T_{obv} is the observation time for calculating the transmitted power, C_{uit} is the unit of the economic benefit from the transmitted power, and $f(x)$ is the risk probability function of the power supply.

3.2   Power supply risk function

According to the principle of safe and stable operation of the network,[17] the estimation of the economic benefit of the power supply in a power grid considers various types of incidents that may occur during the transmission of the DG output.

First, the benefit function $y(x)$ can be expressed as three different equations.

$$y(x) \rightarrow \begin{cases} y_1(x) = \dfrac{1}{c_1}\left(1 - e^{-a_1(x+b_1)}\right) \\ \text{Conservativeutility function} \\ y_2(x) = \dfrac{1}{c_2}\left(e^{a_2(x+b_2)} - 1\right) \\ \text{Adventurous utility function} \\ y_3 = a_3 x + b_3 \\ \text{Neutral utility function} \end{cases} \tag{4}$$

Then, the R-PS is defined as

$$R_{ps}(x) = y(x)T_{res}L(x,\, s_A)g_{DG}(x - s_0), \tag{5}$$

where T_{res} is the duration of occurrence of event E_i, L is the economic cost caused by the event, $f(s_A)$ is the risk of occurrence of event A, g_{DG} is the average probability of the failure of the power distribution, and s_0 is the loss of DG in the back-up power supply of the network.

3.3   Economic power supply function

According to the relationship between the B-PS and the R-PS, the E-PS is defined as

$$E_{ps}(x) = B_{ps}(x) - R_{ps}(x), \tag{6}$$

where $E_{ps}(x)$ is the E-PS and $B_{ps}(x)$ and $R_{ps}(x)$ are the B-PS and R-PS, respectively. Equation (6) comprehensively represents the actual risk and benefit with the attitude of the decision-maker. The analysis of the theory combined with practical experience makes the comprehensive results more reasonable.

4.   Optimized Implementation of DG with Big Data Analysis

4.1   Data integration

The data of the distribution network are obtained from various types of sensors.

(1) Grid structure data

The real-time data of a general power grid mainly come from the supervisory control and data acquisition (SCADA) system. It can provide time data; the operation data related to the detailed state of the power grid are collected and recorded in every minute, such as system frequency, total output, total load, and station operation conditions, including unit output, line power flow, bus voltage, and switch status. The sensor is mainly collected by the remote terminal unit (RTU) set at the side of the station. The interface can be an RS485 or RS232 asynchronous data interface. According to the demand for information, the rate is generally 1200 to 9600 bit/s.

(2) Telemeter reading data

The telemeter reading system (TMRS) is a subsystem for the automatic collection, remote transmission, storage, preprocessing, and statistical analysis of electric energy data. It supports the development of future smart grids, the grid connection of new energy, the operation assessment of the power market, the setting of electricity charges, and the calculation of economic compensation. It is the basis of the system operation.

The statistical power data mainly includes power supply, power generation, etc. The data of the TMRS come from the measurement of electric energy at the gateway points of power plants on-line and off-grid and their tie line, which are collected, stored, and processed in different periods, providing the basis for settlement and analysis.

(3) Demand-side data

The load control terminal is installed on the power user side. Owing to the large number of points, the overall business volume is relatively small. For economic considerations, public network communication systems are generally used to realize information transmission, and most of them adopt a wireless mode. In the early stage, power communication was applied for special wireless frequency point resources and directly controlled the load terminal to perform a load-shedding operation through a wireless channel. At present, the mainstream technology is to use the general packet radio service (GPRS) or code division multiple access (CDMA) to obtain the power consumption of users and carry out load-limiting operations.

A distribution network can use big data analysis technology to classify and mine existing data, make full use of the potential value of the data, and integrate different types of data to provide better convenience for power grid planning.

The distribution network uses big data integration to facilitate grid planning efficiently. One of the strengths of big data is real-time collection and analysis. At present, three steps are involved in big data analysis: analysis, integration, and application of the data.[18] Distribution networks with different structures collect a variety of data continuously. The amount of data is huge for the actual operation of a distribution network with DG. At the same time, the data of the network are messy and uncertain.

The integration of the big data from various sources correlates, integrates, synthesizes, and obtains accurate information through computational analysis.[19] It uses mathematical methods to transform the gathered information to comprehensive information.[20] Different levels of data integration are needed for different analysis, integration, and application of the data. For low-level integration, the methods of state estimation (filtering method), data association, and

classification recognition are used, while for high-level integration, reasoning and intelligent algorithms and artificial intelligence (AI) such as artificial neural networks (ANNs) and genetic algorithms (GAs) are used.

4.2 Data integration for DG access

The main idea of the data integration for optimizing the DG access to the power distribution network is shown in Fig. 3. First, the original data are collected and processed for a hierarchical analysis. Then, the classified data at different levels are integrated and finally the data are used. The categories of the data include the data of the power supply and the economic benefits of the distribution network. The power supply has the structural data of a power grid. The size of the data is determined by the structure of the power grid and the loads of the network. The economic benefit of the power grid is determined by the difference between the cost and income from the users of the power grid. We established the relationship between the expenditure and income through a risk function. We analyzed the correlation among the different types of data, and integrated the data for the DG access.

The repeat power flow (RPF) method is used to estimate the maximum power supply with DG connected to the power distribution network.[21] The input values for the method are the power distribution, power consumption, cost, and so on. The maximum power supply is obtained by the RPF, which is calculated from the risk function of the DG access to the power distribution network. According to the function of the economic benefits of the power supply, a particle swarm optimization (PSO) method is adopted to solve the economic power supply of DG and obtain the output power of the DG.
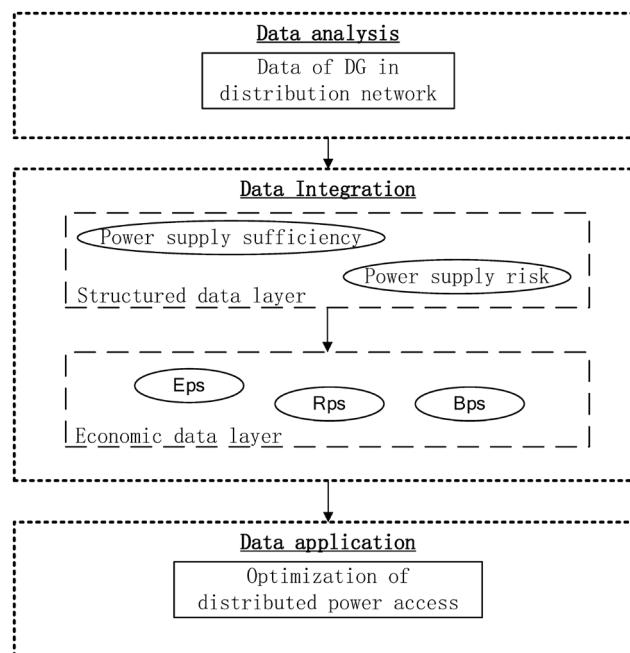

Fig. 3.   Optimized data integration for DG access.

The RPF method is often used to solve an increasing load power flow, that is, by gradually increasing the power generation and load of the system, the power flow equation is solved repeatedly until the constraint condition exceeds the limit.

Equation (7) represents the objective function of maximum power supply. The constraint on $g(x_n, u_n, z_n)$ in Eq. (8) is the equality constraint for the load flow, and that on $h(x_n, u_n, z_n)$ in Eq. (8) is the inequality constraint for the node voltage and node power.

$$\max f(x, u, z) \tag{7}$$

$$\text{s.t.} \begin{cases} g(x_n, u_n, z_n) = 0 \\ h(x_n, u_n, z_n) \geq 0 \end{cases} \tag{8}$$

The relationship of the variables during the iteration is expressed as

$$u_n = u_{n-1} + l_n d_n, \tag{9}$$

where n is the number of iterations ($n = 1, 2, ..., N$), N is the maximum number of iterations, d_n is the amount of unit growth at the nth iteration (the adjustment direction), and l_n is the adjustment step size at the *n*th iteration.

The iterative trajectory of the power flow is shown in Fig. 4. With the initial state of the power flow S_0, the step size is changed one by one, and the result is obtained from the trajectory of the line M, which is an arc.

PSO is a method of obtaining optimal values. In a D-dimensional search space, there are N particles that form a community in which $i = 1, 2, ..., N$. The particle X_i and the "moving" speed V_i of the ith particle are respectively represented as two D-dimensional vectors as follows:

$$X_i = (x_{i1}, x_{i2}, ..., x_{iD}) \tag{10}$$

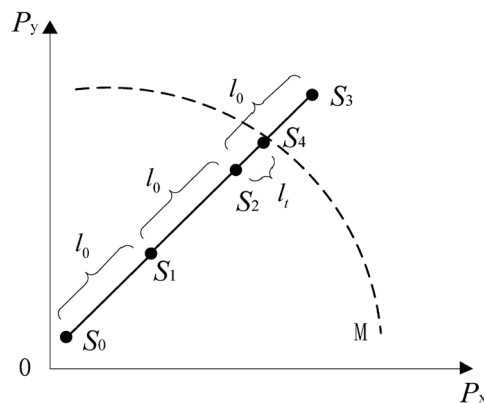$$V_i = (v_{i1}, v_{i2}, ..., v_{iD}) \tag{11}$$

Fig. 4.   Trajectory in RPF iterative method.

The optimal position found by each particle is called the individual extreme value, which is expressed as

$$p_{best} = (p_{i1}, p_{i2}, ..., p_{iD}). \tag{12}$$

By comparing the best positions for each individual value, the extreme value is recorded as

$$g_{best} = (g_{i1}, g_{i2}, ..., g_{iD}). \tag{13}$$

When these two optimal values are found, the particles update their speeds and positions according to the following equations:

$$v_{id} = w \times v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{gd}), \tag{14}$$

$$x_{id} = x_{id} + v_{id}. \tag{15}$$

With these equations, PSO optimizes the maximum power supply. That is, PSO treats the value of the maximum power supply as a particle that belongs to a community and obtains the economic power supply at a certain DG penetration rate according to the principle of PSO. The result evaluates the economic power supply of the distribution network. The specific model is as follows.

The objective function is

$$\max B_{ps}(x) = E_{ps}(x) - \alpha R_{ps}(x). \tag{16}$$

The constraints are

$$x_{min} \le x \le x_{max}, \tag{17}$$

$$x \le S_A, \tag{18}$$

where x is the maximum power supply, and x_{min} and x_{max} are the lower and upper limits of the maximum power supply range y, respectively.

4.3 Program for optimizing DG access

The probability model mainly considers wind power and photovoltaic power as the sources of DG and the load size in a power distribution network.

Solar cell

The output power of a solar cell is directly related to the light intensity.[22] The following equation defines the relationship between the output power of a solar cell and the light intensity:

$$P_R = rA\eta, \tag{19}$$

where P_R is the output power of the solar cell, r is the actual light intensity, η is the conversion efficiency of the solar cell array, and A is the actual area of receiving light on the solar cell array. Since the probability distribution of the light intensity is close to the β distribution, the output power of the solar cell also follows the β distribution and is defined as

$$f(P_R) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{P_R}{P_{R,\max}} \right)^{\alpha - 1} \left(\left(1 - \frac{P_R}{P_{R,\max}} \right)^{\beta - 1} \right), \tag{20}$$

where α and β are the shape parameters of the β distribution and $P_{R,max}$ is the maximum output power of the solar cell. The power curve of the solar cell output is shown in Fig. 5(a).

Wind turbine

Similar to the output power of a solar cell, as the output power of a wind turbine is mainly affected by the wind speed, the probability distribution of the wind speed mainly follows the two-parameter Weibull distribution,[23] and its probability density is defined as

$$f(v) = \left(\frac{k}{c} \right) \left(\frac{v}{c} \right)^{k-1} \exp\left[-\left(\frac{v}{c} \right)^k \right], \tag{21}$$

where k and c are the shape parameter and scale parameter of the Weibull distribution, respectively, and v is the wind speed of the wind turbine.

According to the relationship between the wind speed and the output power of the wind turbine, the actual output of the wind turbine is obtained as

$$P_W = \begin{cases} 0 & v \le v_{c1} \\ k_1 v + k_2 & v_{c1} \le v \le v_r \\ P_r & v_r \le v \le v_{co} \\ 0 & v_{co} \le v \end{cases}, \tag{22}$$
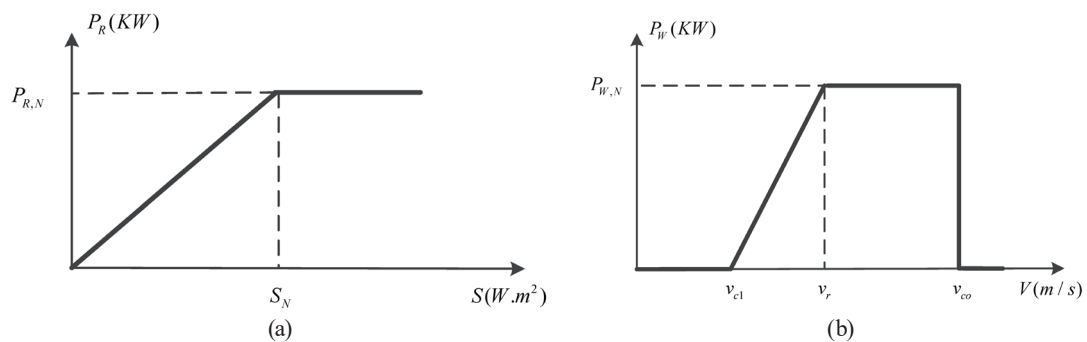

Fig. 5.   Power supply from DG: (a) solar cell, (b) wind turbine.

where $k_1 = P_r/(v_r - v_{c1})$, $k_2 = -k_1 v c_1$, P_W is the output power of the wind turbine, v_{c1}, v_r, and v_{co} are the cut-in wind speed, rated wind speed, and cut-out wind speed, respectively, and P_r is the rated power of the wind turbine. The power curve of the wind turbine output is shown in Fig. 5(b).

Stochastic load model

   It has been proposed that the normal distribution can approximate the uncertainty of the load.[24] The parameters μ_p, σ_p, μ_q, and σ_q of loads P and Q are included in the following probability density functions.

$$f(P) = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left(-\frac{(P-\mu_p)^2}{2\sigma_p^2}\right) \tag{23}$$

$$f(Q) = \frac{1}{\sqrt{2\pi}\sigma_q} \exp\left(-\frac{(P-\mu_q)^2}{2\sigma_q^2}\right) \tag{24}$$

   The specific implementation steps of the model are as follows:
(1)   According to the probability density curve of the DG and the load, the number of samples N is determined, and the uncertainty factors are sampled using the multilevel coordinate search (MCS) method to determine the initial value of the network load S_0, the DG processing size S_{DG}, and the load growth value S_d.
(2)   The load growth rate k as the search step in the program is determined and the convergence accuracy ε is provided.
(3)   By using $S = S_0 + kS_d$ as the new load value to perform the load flow calculation, it is determined whether the load flow meets the constraint in Eq. (8). If there is no out-of-range flow, then step (3) is repeated. In the case of an out-of-range flow, step (4) is executed.
(4) $S = S_0 - kS_d$ is calculated.
(5)   It is determined whether k is less than ε. If this is the case, step (8) is executed. If it is greater than ε, step (6) is executed.
(6)   The step size is decreased to increase the load growth rate h, and $h = h/2$ (initially $h_0 = k/2$). $S = S + hS_d$ is used as the new load value for the power flow calculation to determine whether the flow is out of range.
(7)   If there is no out-of-range flow, step (5) is repeated. If there is an out-of-range flow, $S = S - hS_d$ is calculated first, then step (5) is repeated to continue the calculation.
(8)   With the maximum power supply S_{max} as the output, steps (1) to (7) are repeated.
(9)   After the completion of the calculation in the range of the maximum power (A, B), the probability distribution function $f(s)$ of the maximum power is yielded.
(10) The particle swarm is initialized with a group size N, particle dimension D, inertia weight w, acceleration factors c_1 and c_2, particle positions x_i, and particle velocities v_i. The risk function $f(s)$ is substituted into the target function and recorded as a fitness function.

(11) It is determined whether the range of x_i and v_i meet more conditions, and the adaptive value $F_{it}(i)$ for each particle is calculated.

(12) For each particle, its fitness value $F_{it}(i)$ is compared with the individual extreme value $p_{best}(i)$. If $F_{it}(i) > p_{best}(i)$, $p_{best}(i)$ is replaced with $F_{it}(i)$.

(13) For each particle, its fitness value $F_{it}(i)$ is compared with the global extreme value g_{best}. If $F_{it}(i) > g_{best}(i)$, $F_{it}(i)$ is used instead of g_{best}.

(14) The velocity v_i and the position x_i of the particles are updated.

(15) With the satisfaction of the end condition, go to step (16). Otherwise, return to step (11).

(16) The maximum value of the fitness function is the final output, which includes the economic power supply and the corresponding value of x_i.

5.   Results

We took a power distribution network for the simulation with the structure shown in Fig. 6. In this structure, the original load is 3800 + 2469j kVA and the voltage is 12.66 kV.[25]

The rated power of the photovoltaic cell is 200 kW and the shape coefficient of its probability density function is 0.75. The rated power of the wind turbine is 300 kW and the shape parameter and scale parameter of the probability density function of the wind speed are 2.49 and 8.96, respectively. v_{c1}, v_r, and v_{co} are 2.6, 12, and 20 m/s, respectively. The K ($N - 1 + 1$) criterion of the line is not considered.[26]

5.1   Model validation

To validate the model in this study, we set the sampling number as 1000 and used a combination of Monte Carlo simulation and RPF. The maximum and minimum power supply of the DG at different positions are shown in Table 1.
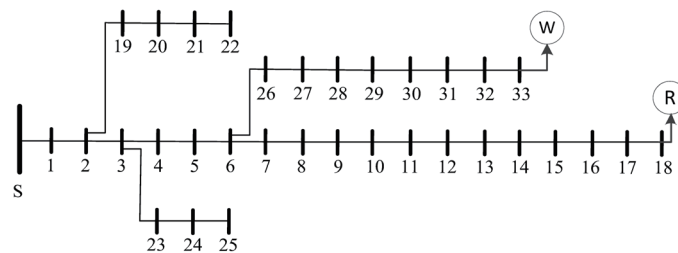

Fig. 6.   33-node distribution network structure. (W: wind turbine, R: solar cell, S: slack bus)

Table 1
Power supply range at different DG positions.

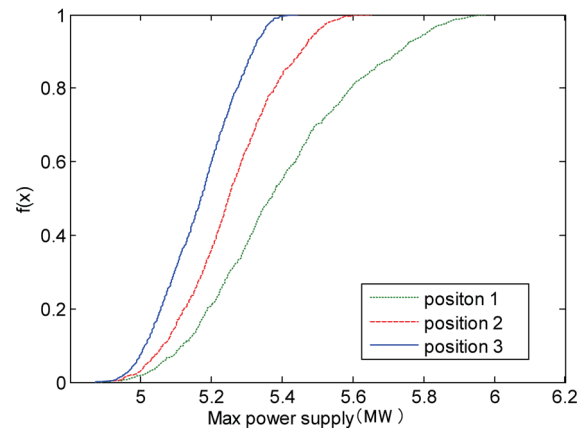| Position	Photovoltaic location	Wind turbine position	Minimum power supply (MW)	Maximum power supply (MW)
1	Node 18	Node 33	4.8869	5.9757
2	Node 11	Node 33	4.8735	5.6510
3	Node 11	Node 24	4.8735	5.4450

Fig. 7.   (Color online) Grid power supply risk.

The power risk of the grid is shown in Fig. 7, which is obtained from the risk function [Eq. (4)]. In Fig. 7, $f(x)$ shows the change in power supply risk when DG is connected. The slopes of the lines at lower and higher ends of the $f(x)$ values are gentler than the slope at intermediate values. If the power supply of the distribution network is in the intermediate range of $f(x)$, a change in the supply has a considerable impact on the power supply risk. At the same time, as Fig. 7 shows, DG at position 3 has the largest power supply change. To determine the power supply risk, we used the PSO algorithm with the following parameters: number of iterations $N = 1000$, size of initial population = 24, weight coefficient $w = 0.85$, learning factors c_1 and $c_2 = 2.13$, maximum iteration number $N_{max} = 1500$. We took the test time of $T_{obv} = 1$ h and the economic benefit per power unit of $C_{uit} = 3000$ yuan/kW·h. When the duration of the risk was $T_{res} = 0.25$ h, the average economic cost for the risk was $L = 700$ yuan/kW·h with a DG failure probability of 0.15 given by the R-PS.[27]

The B-PS at position 3 with and without the risk function is shown in Fig. 8. The effect of the risk function on the power supply and the economic benefit are clear, as can also be seen in Table 2.

From Fig. 8 and Table 2, the following are observed.

(1) Without the risk function, the benefit of the network does not fluctuate, but the benefit decreases with increasing power supply. With the risk function, the benefit increases with the peak power supply, then decreases. A power supply of larger than 5.34 MW has similar values and a similar tendency for the benefit function with and without the risk function. When the power from the DG has a strong influence on the power grid at position 3, the benefit becomes minimal with decreasing power supply. Therefore, the maximum benefit is more reliable with the risk function.

(2) Without the risk function, the power from the DG is much lower with a higher power supply and the benefit is lower. Thus, the risk function makes the benefit function of the network more real and reliable.

(3) The economic power supply with the risk function is 12% higher than that without the function, but the overall benefit is 31% lower.
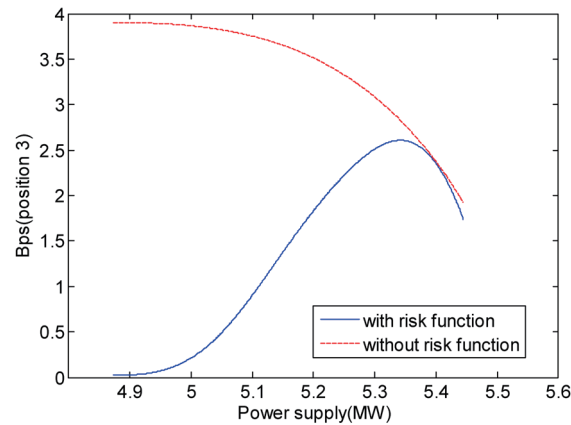
Fig. 8.   (Color online) B-PS at position 3 with and without risk function.

Table 2
Effect of risk function on economic power supply and overall benefits.

Risk function	Power from DG [wind turbine / solar cell (MW)]	Economic power supply (MW)	Overall benefit (10000 yuan)
Considered	0.412 / 0.114	5.342	2.608
Not considered	0.023 / 0.097	4.768	3.767

5.2   Discussion

Figure 9 show the changes in E-PS, power supply risk function R-PS, and B-PS with the power supply at positions 1, 2, and 3. The power input from the DG at different positions produces different E-PS, R-PS, and B-PS of the distribution network. The E-PS of position 1 is almost 0, whereas that of position 2 is slightly higher than that of position 1. The B-PS of position 3 shows the largest value. Increasing the power supply increases the benefit, but after the power supply reaches between 5.3 and 5.4 MW, the benefit decreases.

The E-PS and B-PS at different positions of the DG access are shown in Table 3. The overall benefit is calculated as follows. As the cost of the power is 3000 yuan/kW·h, assuming that the duration of the risk is 0.25 h, the average cost for the risk is 700 yuan/kW·h. For the failure probability of 0.15 for the power supply from the DG, the final calculation result is defined as the overall benefit. Table 3 reveals that the DG access position affects the power supply range, the power from the DG, the economic power supply, and the overall benefits of the network. Positions 1 and 3 have similar power from the DG but position 2 has significantly different power. The economic power supply varies by 2–5% among the three positions, and the overall benefit of the distribution network at the three positions varies considerably: position 2 has 3.13 times the benefit of position 1 and position 3 has 4.78 times the benefit of position 1. This implies that the design of the DG access to the power distribution network is important for the efficient management of the network.

Both the access position and output supply of the DG affect the power supply capacity of the network. However, to ensure the power supply capacity of the distribution network, it is

(a) position 1

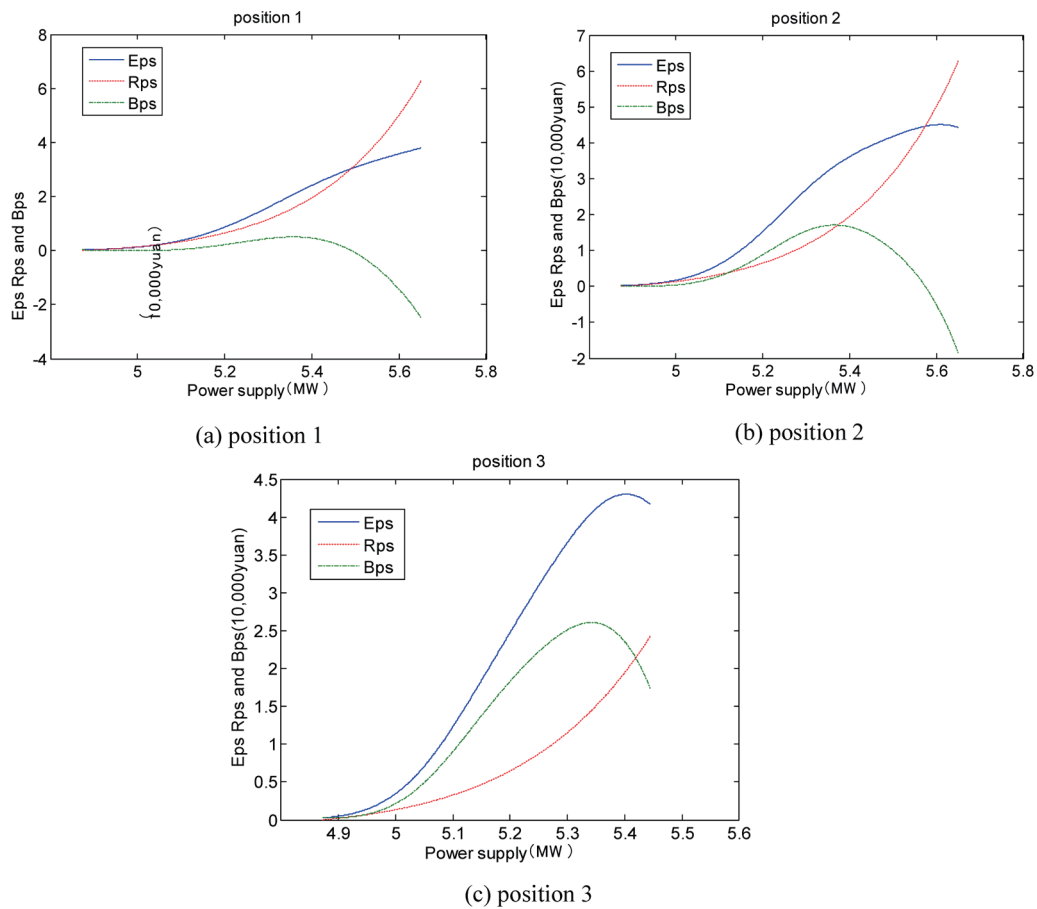(b) position 2

(c) position 3

Fig. 9.   (Color online) Changes in E-PS, B-PS, and R-PS with power supply in the power distribution network.

Table 3
E-PS and B-PS for different positions of DG.

DG position	Power supply range (MW)	Power from DG [wind turbine / solar cell (MW)]	Economic power supply (MW)	Overall benefit (10000 yuan)
Position 1	4.886–5.975	0.207 / 0.346	5.4912	0.5457
Position 2	4.874–5.651	0.191 / 0.137	5.3663	1.7079
Position 3	4.872–5.445	0.412 / 0.114	5.1121	2.6085

necessary to find a reasonable position for the DG. We used big data analysis to compare the economy of the distribution network connected to the DG. Figure 10 shows the effect of the attitude of decision-makers toward the use of DG access in position 3 on B-PS. The E-PS and B-PS for different attitudes are shown in Table 4.

The degree of the risk-taking affects the power supply from the DG, the economic power supply, and the overall benefit. Higher risk-taking requires a 14% lower power supply from the DG than lower risk-taking. The value of the economic power supply is 1.8% lower for high risk-taking than for low risk-taking. The difference in the overall benefit is significant as the
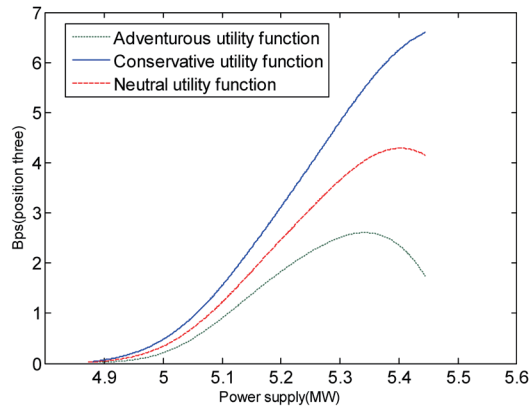
Fig. 10.   (Color online) B-PS function for different attitudes of decision-makers at position 3.

Table 4
E-PS and B-PS for different degrees of risk-taking at position 3.

Degree of risk-taking	Power from DG [wind turbine / solar cell (MW)]	Economic power supply (MW)	Overall benefit (10000 yuan)
Low	0.107 / 0.948	5.441	6.573
Middle	0.097 / 0.766	5.396	4.287
High	0.034 / 0.892	5.343	2.608

management with low risk-taking has 2.5 times larger benefit than that with high risk-taking. The degree of risk-taking has more effect on the overall benefit than the position of the DG access to the network. Thus, the level of risk taken for the management is important for the management of the power distribution network.

The main novelty of this study is the proposal of a new evaluation method for distribution networks based on big data analysis. This method first guarantees the power supply capacity of the distribution network itself in order to evaluate the risk of DG to the distribution network. We used structural and economic big data to analyze the distribution network hierarchy, and then evaluated the economic benefits of the network.

6.   Conclusions

In this study, we considered the impact and role of DG for a power distribution network on the power supply capacity and economy of the network. Through the integration of big data that were obtained from the sensors in the network, we proposed a method of evaluating the economic power supply capacity, enabling us to provide guidelines for connecting DG to the power grid. The method included three functions to evaluate the power supply benefit, power supply risk, and economic power supply. We used an algorithm with a stochastic load model and the power from a wind turbine and a solar cell as the DG. The results showed that the risk of the power supply with DG access to the distribution network was different at different positions of the connection. The risk function affected the power supply from the DG and

the economic power supply of the network. Without the risk function, the overall benefit was higher, but the power supply from the DG and the economic power supply were lower than those with the risk function. The position of the connection of the DG to the power distribution network also affected the benefit, risk, and economy of the power supply in the network. The decision on how much risk to take in the power supply management is also important for the economic operation of the power distribution network. Depending on the position and the degree of risk-taking, the overall benefit can vary by a factor of 3–5. With appropriate use, the power from the DG ensures the affluent power supply capacity of the distribution network to maximize the overall economic benefits of the distribution network. For the management of the power distribution network, big data analysis can play an important role in establishing guidelines for optimizing the use of DG.

References

1 J. Zhang, X. Li, Y. Gao, H. Wang, and Y. Yang: Power Supply **36** (2019) 35. http://doi.org/10.3969/j.issn.1006-6357.2016.04.006
2 E. Ali and Q. Yang: IEEE Trans. Sustainable Energy **10** (2019) 1813. http://doi.org/10.1109/TSTE.2018.2873370
3 J. Xu and Y. An: Electr. Eng. Technol. **482** (2018) 86. http://doi.org/10.3969/j.issn.1002-1388.2018.20.036
4 D. Teng, H. Teng, and X. Liu: Electrometer **56** (2018) 13. http://doi.org/10.3969/j.issn.1001-1390.2018.24.010
5 J. Li, Z. Gu, and Z. Lu: North China Electric Power **6** (2011) 500. http://doi.org/10.16308/j.cnki.issn1003-9171.2008.02.015
6 T. Wang, P. Wang, and M. Fan: 2000 Proc. CSEE (CSEE, 2000) 54–58. http://doi.org/10.3321/j.issn:0258-8013.2000.05.014
7 W. Lu and Z. Guo: 2000 Proc. CSEE. (CSEE, 2000) 86–91. http://doi.org/10.3321/j.issn:0258-8013.2003.08.017
8 P. Gao, L. Wang, L. Li, B. Xu, and J. Zhou: Elect. Power Autom. Equip. **2** (2019) 539. http://doi.org/10.16081/j.issn.1006-6047.2019.04.028
9 J. Li, L. Li, and C. Gao: 2013 Proc. the CSU-EPSA (CSEE, 2013) 13–21. http://doi.org/10.3969/j.issn.1003-8930.2013.05.003
10 L. Yang: Research and Application of Smart Grid Big Data Fusion Method (Northeast Electric Power Univ., Doctoral diss, 2018). http://doi.org/cnki:cdmd:2.1016.276383
11 Q. Meng: Rural Power Mgt. **5** (2020) 44. http://doi.org/10.3969/j.issn.1672-2450.2020.05.020
12 T. Zhao, Y. Zhang, and D. Zhang: IEEE Power Syst. Technol. **38** (2014) 12. http://doi.org/10.13335/j.1000-3673.pst.2014.12.006
13 B. Hong, B. Huang, L. Ge, Q. Li, D. Ge, and Y. He: Power Autom. Equip. **39** (2019) 26. http://doi.org/10.16081/j.issn.1006-6047.2019.02.004
14 J. Li, L. Li, C. Gao, Y. Yang, W. Guo, and F. Wen: Proc. CSU-EPSA. **25** (2013) 13. http://doi.org/10.3969/j.issn.1003-8930.2013.05.003
15 Z. Wang, S. Wu, and W. Liu: 2014 Proc. Innovative Smart Grid Technologies-Asia (ISGT Asia, 2014) 354–359.
16 F. Xiao and J. D. Mccalley: IEEE Trans. Power Syst. **24** (2009) 78. http://doi.org/10.1109/TPWRS.2008.2004823
17 J. Zhao: Value Eng. **31** (2014) 25. http://doi.org/10.3969/j.issn.1006-4311.2014.31.012
18 S. Li and S. Zhao: Electr. Eng. **54** (2018) 35. http://doi.org/10.19768/j.cnki.dgjs.2019.24.050
19 L. Wang, H. Li, and P Zhang: Power Syst. Technol. **37** (2014) 358. http://doi.org/10.13335/j.1000-3673.pst.2014.03.034
20 S. Xiao, S. Liu, M. Song, N. Ang, and H. Zhang: Multibody Syst. Dyn. **48** (2020) 451. http://doi.org/10.1007/s11044-019-09718-9
21 Z. Li, X. Chen, H. Liu, and K. Yu: Electr. Autom. Electr. Power Syst. **33** (2009) 36. https://doi.org/10.3321/j.issn:1000-1026.2009.06.008
22 N. Tang, M. Xia, W. Xiao, and Y. Zhong: Autom. Electr. Power Syst. **39** (2015) 45. https://doi.org/10.7500/AEPS20140330005
23 Y. Li: Wind Speed Model with Correlation and Its Application in Reliability Evaluation of Power Generation System (Chongqing Univ., Doctoral diss., 2012).
24 Y. Yuan and L. Dong: Mod. Electr. Power. **24** (2007) 5. https://doi.org/10.3969/j.issn.1007-2322.2007.04.002

25 Y. Ou and C. Singh: IEE Proc. Generation, Transmission and Distribution. **150** (2003) 239. https://doi.org/10.1049/ip-gtd:20030024
26 J. Zhang, G. Wei, L. Zhu, J. Chi, and J. Chu: Autom. Electr. Power Syst. **40** (2016) 64. https://doi.org/10.7500/AEPS20150615003
27 S. Zhang and H. Chen: IEEE Trans. Power Syst. **226** (2018) 743. https://doi.org/10.1109/TPWRS.2013.2245924