2024 6th Asia Energy and Electrical Engineering Symposium (AEEES) | 979-8-3503-7347-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/AEEES61147.2024.10545016

Distributed PV Hosting Capacity Assessment of Distribution Network Based on Global Probabilistic Voltage Sensitivity

Mingze Zhang
State Grid Shanghai Economic Research Institute
Shanghai, China
casey.zhang1985@163.com

Shenghui Hua
State Grid Shanghai Economic Research Institute
Shanghai, China
huashenghuish@163.com

Yuan Ji
State Grid Shanghai Economic Research Institute
Shanghai, China
jiyuan1126@sina.com

Xincong Li
State Grid Shanghai Economic Research Institute
Shanghai, China
362067003@qq.com

Minhao Xia
State Grid Shanghai Economic Research Institute
Shanghai, China
13564759351@163.com

Rui Wang
School of Electrical and Power Engineering, Hohai University
Nanjing, China
221606030095@hhu.edu.cn

Jian Wang*
School of Electrical and Power Engineering, Hohai University
Nanjing, China
eewangjian@hhu.edu.cn
*corresponding author

***Abstract*—The scale of distributed PV integration into the distribution network is increasingly expanding recently. It is of significant importance to accurately assess the maximum capacity of distributed PV for ensuring the safe operation of the distribution network and PV planning. Therefore, an distributed PV hosting capacity assessment method of distribution network based on global probabilistic voltage sensitivity is proposed. Firstly, a voltage sensitivity model for the distribution network is constructed based on linear power flow model. Subsequently, a Gaussian mixture model is employed to establish a probability distribution model for the prediction errors of distributed PV. The probability analytical expressions for power fluctuation and voltage fluctuation are derived through the affine invariance of the Gaussian mixture distribution. Then, the probability distribution of voltage exceedance and voltage violation criteria are studied to calculate the hosting capacity of distributed PV in the distribution network. Finally, the adaptability and effectiveness of the proposed method are verified using an IEEE 33-node distribution network.**

Keywords—probabilistic voltage sensitivity, Gaussian mixture model, voltage violation, PV hosting capacity

I. INTRODUCTION

With the rapid development of renewable energy and the transformation of power systems, distributed PV, as a clean and renewable energy source, is playing an increasingly important role. However, the output of distributed PV exhibits significant intermittency and volatility. The high proportion of distributed PV integration makes the operational scenario of the distribution network more complex, bringing about various adverse effects[1]. Therefore, it is essential to accurately assess the hosting capacity of distributed PV in the distribution network to ensure the safe operation of the grid.

The current research methods for assessing the hosting capacity of distributed PV in distribution networks can be broadly categorized into two approaches: (1) optimization-based assessment methods and (2) trial-and-error verification assessment methods.

(1) Optimization-based assessment method: This approach aims to maximize the PV integration capacity, using voltage deviation, voltage fluctuation, and line flow as constraint conditions. It establishes a model for assessing the hosting capacity of distributed PV and employs suitable optimization algorithms for solving. For instance, in [2], transformer capacity, voltage fluctuation, and short-circuit current are considered as constraints to establish a model for the maximum integration capacity of distributed PV in a medium-voltage distribution network. Reference [3] further considers constraints such as network losses, voltage fluctuation, and short-circuit capacity, establishing a model for evaluating the PV accommodation capacity in the distribution network with the objective of maximizing PV integration capacity and minimizing network losses. Reference [4], building upon the analysis of the temporal characteristics of distributed PV and load, considers constraints such as transformer tap settings, reactive power compensation devices, and PV output to establish a model for the maximum integration capacity of distributed PV in a distribution network.

(2) Trial-and-error verification assessment method: This method involves continuously adjusting the integration capacity of distributed PV to simulate various scenarios of PV integration. Subsequently, power flow calculations are performed to evaluate distribution network performance indicators until one or more assessment criteria reach a critical threshold. At this point, the integration capacity of distributed PV becomes the hosting capacity. Reference [5] employs the trial-and-error verification method by incrementally increasing PV grid

capacity until voltage exceeds limits, thus determining the grid's hosting capacity. Reference [6] utilizes a software simulation platform to establish a distribution network simulation model, progressively adding PV capacity in different locations to determine the maximum capacity of PV integration into the grid under voltage deviation and harmonic constraints. Reference [7] calculate the additional capacity of distributed PV under different reverse load ratios, conducting verifications such as short-circuit current, voltage deviation, or harmonic verification. If verification fails, the additional PV capacity is gradually reduced to determine the hosting capacity of distributed PV in the distribution network. Reference [8] proposes a spatiotemporal probabilistic voltage sensitivity analysis to calculate the PV hosting capacity of a distribution network. The method involves incrementally increasing the PV penetration level, identifying the penetration level at which the first voltage limit is exceeded as the hosting capacity, although this method entails a large computational workload during the derivation of probabilistic voltage sensitivity expressions.

The aforementioned optimization-based assessment methods usually consider specific PV integration scenarios, where the PV installation locations are relatively fixed. Inspired by trial-and-error verification methods and recognizing that voltage is a primary factor in determining PV hosting capacity[9], voltage sensitivity analysis can help identify voltage violation situations to calculate the PV integration capacity of the distribution system. In this paper, we propose a method for calculating the hosting capacity of distributed PV in distribution networks based on probabilistic voltage sensitivity. Firstly, the voltage sensitivity model of the distribution network is derived based on a power flow model. Next, a Gaussian mixture model is employed to fit the prediction errors of distributed PV generation. Leveraging the properties of the Gaussian mixture distribution, the probabilistic voltage sensitivity model of distributed PV generation to node voltages and the probability distribution function of voltage fluctuation are obtained. Subsequently, the voltage violation criteria in the distribution network are studied. For each increment in the distributed PV integration capacity, *N* scenarios of PV integration are randomly generated. By evaluating the voltage violation criteria in each scenario, the maximum distributed PV integration capacity without voltage violation is determined as the sought-after hosting capacity of PV in the distribution network. Finally, the effectiveness of the proposed method is analyzed through case studies.

II. Voltage Sensitivity Equation

The voltage sensitivity equation derived is based on a radial distribution network with *n* nodes, as illustrated in Fig. 1.

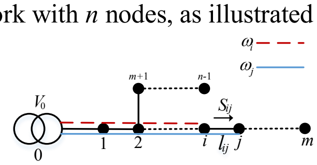

Fig. 1. Radial distribution network topology with *n* nodes.

In Fig. 1, taking the balanced node as the root node and denoting the voltage at this node as V_0. ω_i and ω_j are the sets of branches from nodes i and j to the root node, respectively. l_{ij} is the branch between node i and j. Z_{ij} is the impedance of the branch between node i and node j, and its value is $R_{ij}+X_{ij}$. S_{ij} is the current flow between node i and node j, and its value is $P_{ij}+Q_{ij}$. So, based on the *distflow* power flow model, the voltage relationship between node i and node j can be expressed as (1):

$$V_j = V_i - \frac{P_{ij}R_{ij}+Q_{ij}X_{ij}}{V_j} \tag{1}$$

On this basis, assuming that the impact of line power loss is negligible and that the voltage magnitude at each node is approximately equal to V_0, we can obtain (2)-(4):

$$V_j = V_0 + \frac{P_{j,eq}R_{ij}+Q_{j,eq}X_{ij}}{V_0} \tag{2}$$

$$P_{i,eq} = P_j + \sum\nolimits_{k\in\phi_{down}(j)} P_k \tag{3}$$

$$Q_{i,eq} = Q_j + \sum\nolimits_{k\in\phi_{down}(j)} Q_k \tag{4}$$

Where $P_{i,eq}$ and $Q_{i,eq}$ represent the equivalent active and reactive power injections at node j, respectively; $P_k = P_k^g - P_k^c$ and $Q_k = Q_k^g - Q_k^c$ represent the injected active and reactive power at node k, respectively; $\phi_{down}(j)$ denotes the set of downstream nodes from node.

Define $r_{j,k}$ and $x_{j,k}$ as the shared branch resistance and reactance from node j to the root node, and from node k to the root node, as given in (5). These are derived based on the grid's topology structure and the impedance of the lines:

$$r_{j,k}+x_{j,k} = \begin{cases} R_{0,j}+jX_{0,j}, if\ k\in\phi_{down}(j)\ or\ k=j \\ R_{0,k}+jX_{0,k}, if\ j\in\phi_{down}(k) \\ R_{0,m}+jX_{0,m}, if\ k\notin\phi_{down}(j)\ or\ j\notin\phi_{down}(k) \end{cases}$$
$$R_{0,n}+jX_{0,n} = \sum\nolimits_{l\in\omega_n}(R_l+jX_l), n=j,k,m \tag{5}$$

Where m is the first node in the upstream nodes of nodes j and k where an intersection occurs; $\phi_{down}(j)$ and $\phi_{down}(k)$ represent the sets of downstream nodes for nodes j and k, respectively; ω_n is the set of branches from node n to the root node.

Furthermore, by incorporating (5) into (2) and extending it to all branches in the distribution network, we obtain (6):

$$V_j = V_0 + \sum_{k=1}^{n-1}\left(P_k\frac{r_{j,k}}{V_0}+Q_k\frac{x_{j,k}}{V_0}\right) \tag{6}$$

Substituting $P_k = P_k^g - P_k^c$, $Q_k = Q_k^g - Q_k^c$ into (6), we obtain (7):

$$\begin{aligned} V_j &= V_0 + \left[\sum_{k=1}^{n-1}\left((P_k^g-P_k^c)\frac{r_{j,k}}{V_0}+(Q_k^g-Q_k^c)\frac{x_{j,k}}{V_0}\right)\right] \\ &= V_0 + \left[\sum_{k=1}^{n-1}\left(P_k^g\frac{r_{j,k}}{V_0}+Q_k^g\frac{x_{j,k}}{V_0}\right)-\sum_{k=1}^{n-1}\left(P_k^c\frac{r_{j,k}}{V_0}+Q_k^c\frac{x_{j,k}}{V_0}\right)\right] \end{aligned} \tag{7}$$

Where $r_{i,k}/V_0$ and $x_{i,k}/V_0$ can be respectively considered as the sensitivity of the active power P_k at node k to the voltage V_j at node j, and the sensitivity of the reactive power Q_k at node k to the voltage V_j at node j. Transforming (7) into matrix form, we obtain (8):

$$V_j = V_0 + \left[\left(\boldsymbol{P}^g\boldsymbol{J}_j^r+\boldsymbol{Q}^g\boldsymbol{J}_j^x\right)-\left(\boldsymbol{P}^c\boldsymbol{J}_j^r+\boldsymbol{Q}^c\boldsymbol{J}_j^x\right)\right] \tag{8}$$

Where $P^g = \left[P_1^g, \ldots, P_j^g, \ldots, P_{n-1}^g\right]^{\mathrm{T}}$ is the vector of injected active power at nodes; $Q^g = \left[Q_1^g, \ldots, Q_j^g, \ldots, Q_{n-1}^g\right]^{\mathrm{T}}$ is the vector of injected reactive power at nodes; $P^c = \left[P_1^c, \ldots, P_j^c, \ldots, P_{n-1}^c\right]^{\mathrm{T}}$ is the vector of active power consumption of loads at nodes; $Q^c = \left[Q_1^c, \ldots, Q_j^c, \ldots, Q_{n-1}^c\right]^{\mathrm{T}}$ is the vector of reactive power consumption of loads at nodes; $J_j^r = \left[r_{j,1}/V_0, \ldots, r_{j,k}/V_0, \ldots, r_{j,n-1}/V_0\right]$ is the active-voltage sensitivity vector at node j ; $J_j^x = \left[x_{j,1}/V_0, \ldots, x_{j,k}/V_0, \ldots, x_{j,n-1}/V_0\right]$ is the reactive-voltage sensitivity vector at node j .

From (8), if there is a power fluctuation at a certain node in the distribution network, the voltage at node j will undergo corresponding changes. At this point, the voltage is denoted as $V_{j,new}$, and the voltage fluctuation is denoted as ΔV_j, with the value expressed as shown in (9):

$$\begin{aligned} \Delta V_j &= V_{j,new} - V_j \\ &= \left[\left(\boldsymbol{J}_j^r \boldsymbol{P}_{new}^g + \boldsymbol{J}_j^x \boldsymbol{Q}_{new}^g\right) - (\boldsymbol{J}_j^r \boldsymbol{P}_{new}^c + \boldsymbol{J}_j^x \boldsymbol{Q}_{new}^c)\right] \\ &\quad - \left[\left(\boldsymbol{J}_j^r \boldsymbol{P}^g + J_j^x \boldsymbol{Q}^g\right) - \left(J_j^r \boldsymbol{P}^c + \boldsymbol{J}_j^x \boldsymbol{Q}^c\right)\right] \\ &= \left[\boldsymbol{J}_j^r (\boldsymbol{P}_{new}^g - \boldsymbol{P}^g) + \boldsymbol{J}_j^x (\boldsymbol{Q}_{new}^g - \boldsymbol{Q}^g)\right] \\ &\quad - \left[\boldsymbol{J}_j^r (\boldsymbol{Q}_{new}^g - \boldsymbol{Q}^g) + \boldsymbol{J}_j^x (\boldsymbol{Q}_{new}^g - \boldsymbol{Q}^g)\right] \end{aligned} \tag{9}$$

Where let $\Delta \boldsymbol{P}_g = \boldsymbol{P}_{new}^g - \boldsymbol{P}^g$, $\Delta \boldsymbol{Q}_g = \boldsymbol{Q}_{new}^g - \boldsymbol{Q}^g$, $\Delta \boldsymbol{P}_c = \boldsymbol{P}_{new}^c - \boldsymbol{P}^c$, $\Delta \boldsymbol{Q}_c = \boldsymbol{Q}_{new}^c - \boldsymbol{Q}^c$, $\Delta \boldsymbol{P}_g$, $\Delta \boldsymbol{Q}_g$, $\Delta \boldsymbol{P}_c$, $\Delta \boldsymbol{Q}_c$ are the vectors of active and reactive power fluctuations from distributed PV generation at nodes and the vectors of active and reactive power fluctuations from loads at nodes, respectively. Then, (9) can be simplified to (10):

$$\Delta V_j = \left(\boldsymbol{J}_j^r \Delta \boldsymbol{P}^g + \boldsymbol{J}_j^x \Delta \boldsymbol{Q}^g\right) - \left(\boldsymbol{J}_j^r \Delta \boldsymbol{P}^c + \boldsymbol{J}_j^x \Delta \boldsymbol{Q}^c\right) \tag{10}$$

Based on this, further assumptions are made: 1) The active power of distributed PV generation is random, and distributed PV generation generally injects only active power according to the power factor, while reactive power can be ignored; 2) Load fluctuations are ignored. Then, (10) can be simplified to:

$$\begin{aligned} \Delta V_j &= \boldsymbol{J}_j^r \Delta \boldsymbol{P}^g \\ \Delta \boldsymbol{P}^g &= \left[\Delta \boldsymbol{P}_1^g, \ldots, \Delta \boldsymbol{P}_j^g, \ldots, \Delta \boldsymbol{P}_{n-1}^g\right] \end{aligned} \tag{11}$$

Thus, a voltage sensitivity equation based on *distflow* is formulated, which enables the tracking of the sensitivity of node power fluctuations to the voltages at various nodes. In reality, the power of distributed PV generation can be expressed as the sum of the predicted value and the prediction error, as shown in (12):

$$P_{i,actual}^g = P_{i,forecast}^g + P_{i,error}^g \tag{12}$$

Where $P_{i,actual}^g$ is the actual output of distributed PV generation at node i ; $P_{i,forecast}^g$ is the predicted output of distributed PV generation at node i ; $P_{i,error}^g$ is the prediction error of distributed PV generation output at node i . Substituting (12) into (11), the predicted voltage fluctuation at node j can be obtained as follows (13) (14):

$$\Delta V_{j,forecast} = \left(\boldsymbol{J}_j^r\right)^{\mathrm{T}} P_{error}^g \tag{13}$$

$$\boldsymbol{P}_{error}^g = \left[P_{1,error}^g, \ldots, P_{j,error}^g, \ldots, P_{n-1,error}^g\right]^{\mathrm{T}} \tag{14}$$

III. Method for Distributed PV Hosting Capacity Assessment of Distribution Network Based on Probabilistic Voltage Sensitivity

This article employs a Gaussian mixture model (GMM) to characterize the uncertainty in the prediction errors of distributed PV generation output. Let x be a random variable vector, and $f(x)$ be the joint probability distribution of x . $f(x)$ can be approximated as a GMM distribution and expressed as follows:

$$f(x) = \sum_{k=1}^{K} \omega_k N(x|\mu_k, \textstyle\sum_k) \tag{15}$$

Where $N(x|\mu_k, \sum_k)$ represents a multivariate Gaussian probability distribution, μ_k is the mean, and $\sum_k$ is the covariance matrix. ω_k is the weight factor of the k -th component of the GMM, satisfying $\sum_{k=1}^{K} \omega_k = 1, 0 < \omega_k \le 1$; K is the number of Gaussian components. This article uses the Bayesian Information Criterion (BIC) to find the optimal number of components and then employs the Expectation-Maximization (EM) algorithm[10] for parameter estimation. EM is an iterative algorithm used for parameter estimation in probabilistic models with missing data or latent variables.

Additionally, GMM possesses the property of affine invariance[11]. If a random variable ξ follows a GMM in n dimensions, then $Y = C^{\mathrm{T}}\xi + c$ also follows a GMM. Here, C is a constant matrix of dimensions $n \times m$, c is a constant. The probability density functions (PDF) of ξ and Y are given by (16-17):

$$PDF_{\boldsymbol{\xi}}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \alpha_i N\left(\boldsymbol{\theta} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\right) \tag{16}$$

$$PDF_Y(\eta) = \sum_{i=1}^{n} \alpha_i N\left(\eta \middle| \boldsymbol{A}^{\mathrm{T}} \boldsymbol{\mu}_i + c, \boldsymbol{A}^{\mathrm{T}} \boldsymbol{\Sigma}_i \boldsymbol{A}\right) \tag{17}$$

Therefore, assuming that the random variable P_{error}^g in (13) follows a GMM distribution, it can be expressed as:

$$\boldsymbol{P}_{error}^g \sim \sum_{k=1}^{K} \omega_k N\left(\boldsymbol{p}_{error}^g \middle| \boldsymbol{\mu}_k^g, \boldsymbol{\Sigma}_k^g\right) \tag{18}$$

Furthermore, based on the affine invariance property of GMM, it can be inferred that the voltage sensitivity of node j will also follow a GMM distribution:

$$\Delta V_j = \left(\boldsymbol{J}_j^r\right)^{\mathrm{T}} P_{error}^g \sim \sum_{k=1}^{K} \omega_k N\left[\left(\boldsymbol{J}_j^r\right)^{\mathrm{T}} \mu_k^g, \left(\boldsymbol{J}_j^r\right)^{\mathrm{T}} \Sigma_k^g \boldsymbol{J}_j^r\right] \tag{19}$$

Considering the predicted data of distributed PV generation as deterministic input for distribution network operation scheduling, the voltage at node j denoted as $V_{j,p}^t$, when there is fluctuation in distributed PV generation, can be expressed as:

$$V_{j,f}^t = V_{j,p}^t + \Delta V_j^t = V_{j,p}^t + \left(\boldsymbol{J}_j^r\right)^{\mathrm{T}} \boldsymbol{P}_{error}^{g,t} \tag{20}$$

Where $V_{j,p}^t$ is the voltage without fluctuation in distributed PV generation, and $P_{error}^{g,t}$ is the prediction error of distributed PV generation. Therefore, the random variable $\Delta V_{j,f}^t$ for voltage violation is given by:

$$\Delta V_{j,f}^t = V_{j,f}^t - V_N = V_{j,p}^t + \left(\boldsymbol{J}_j^r\right)^{\mathrm{T}} \boldsymbol{P}_{error}^{g,t} - V_N \tag{21}$$

Where V_N is the rated voltage of the system. Based on (19) and (21), the probability density function and cumulative distribution function of the voltage violation risk for node j are given by (22) and (23), respectively:

$$f_{\Delta V_j^t}(v)=\sum_{k=1}^{K}\omega_k\frac{e^{-\frac{\left(v-\mu_k^t\right)^2}{2\Sigma_k}}}{\left(2\pi\right)^{\frac{1}{2}}\left|\Sigma_k\right|^{\frac{1}{2}}} \tag{22}$$

$$F_{\Delta V_j^t}(v)=\sum_{k=1}^{K}\omega_k\Phi\left(\frac{v-\mu_k^t}{\sqrt{\left|\Sigma_k\right|}}\right) \tag{23}$$

Where $\mu_k^t=\left(J_j^r\right)^{\mathrm{T}}\mu_k^g+V_{j,p}^t-V_N$, $\Phi(\cdot)$ is the cumulative distribution function of the standard Gaussian distribution.

A threshold λ is set. When the cumulative probability of voltage violation at node j exceeding 0.05 p.u. of the rated voltage surpasses the threshold λ [12], it can be considered as a voltage violation :

$$1-F_{\Delta V_j^t}(0.05)=1-\sum_{k=1}^{K}\omega_k\Phi\left(\frac{0.05-\mu_k^t}{\sqrt{\left|\Sigma_k\right|}}\right)>\lambda \tag{24}$$

In conclusion, the distributed PV hosting capacity assessment method for distribution networks based on probabilistic sensitivity is obtained. The steps are as follows, and the flowchart is illustrated in Fig. 2 below:

(1) Set the range of distributed PV access capacity to the distribution network.

(2) For each PV access capacity, randomly generate N scenarios of PV access. For each scenario, use the probability sensitivity method to calculate the probability distribution of voltage violation at a specific capacity. Evaluate whether each voltage violates the criteria according to (24).

(3) If node voltages do not violate the criteria, continue to increase the PV access capacity.

(4) Repeat the above steps until the maximum PV access capacity without voltage violations is determined, which is considered as the PV hosting capacity.

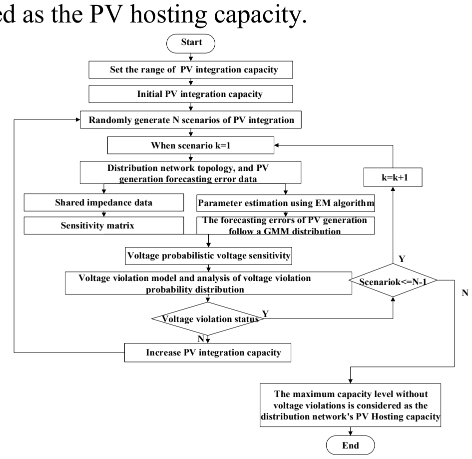

Fig. 2. Distribution network distributed PV hosting capacity assessment process based on probabilistic voltage sensitivity flowchart.

IV. Case Study

The model derived in this paper was validated using the IEEE 33-node distribution network system. The system's structural diagram is shown in Fig. 3. The system's base voltage and base capacity are 12.66kV and 10MVA, respectively. The voltage magnitude at the balanced node is 1.00 p.u. , with a total active load of 3.535MW and a reactive load of 2.3Mvar. Typical PV and load prediction data are presented in Fig. 4.

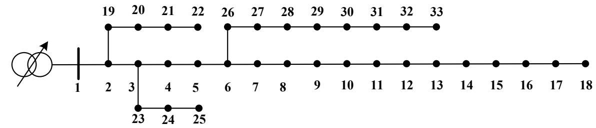

Fig. 3. IEEE 33-node distribution system.

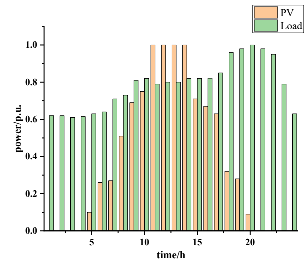

Fig. 4. Typical daily forecasting operational data.

A. Verification of the Voltage Sensitivity Model Simulation Results

Verification of the accuracy of the voltage sensitivity model will be conducted by comparing the simulated results with the actual voltage fluctuations obtained through power flow calculations in the IEEE 33-node distribution network. The simulation results are illustrated in Fig. 5 and Fig. 6.

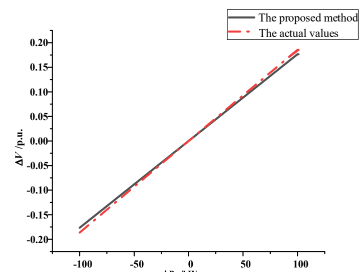

Fig. 5. Voltage fluctuation at node 25 when there is a power change at node 25.

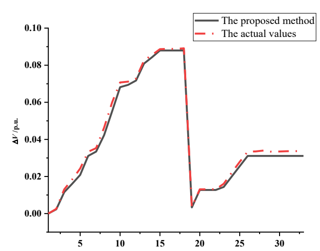

Fig. 6. All node voltage fluctuations caused by an increase of 1 kW at nodes 10, 15, 20, and 25.

Fig. 5 shows the voltage fluctuations at node 25 when the power fluctuation range at node 25 is -100kW to +100kW, as calculated by two different methods. It can be observed that within a small range of power fluctuations, the results from both methods are approximately consistent with minimal error.

Additionally, Fig. 6 addresses voltage fluctuations caused by power changes at multiple nodes. From Fig. 6, it can be seen that the results from both methods are approximately consistent. The degree of voltage fluctuation at a node is greater when the length of the shared path, or impedance, between the power-changing node and the voltage-fluctuating node is larger. The voltage fluctuations at nodes 25 to 33 remain constant because these nodes share the same path, meaning they have the same shared impedance.

In conclusion, it is evident that the proposed model demonstrates good accuracy, effectively reflecting the impact of node power changes on voltage. Furthermore, the complexity of the established model is relatively small, making it easy to compute, meeting practical application needs.

B. Simulation Results Verification of Distributed PV Hosting Capacity Assessment in Distribution Networks Based on Probabilistic Voltage Sensitivity

In order to assess the performance of the proposed method in the assessment of PV hosting capacity in distribution networks, a comparison was made with the baseline method using Monte Carlo Newton Raphson power flow. The introduced method for calculating distributed PV integration capacity based on probabilistic voltage sensitivity remains effective for nodes experiencing arbitrary power variations.

PV integration capacity into the distribution network is set in the range of 50 kW to 5000 kW, with a step size of 50 kW. First, under the initial PV integration capacity, a set of scenarios with random PV integrations was generated. For each scenario, the distribution parameters of the prediction error were established to form the error margin distribution function for PV generation[13,14], based on which prediction error data was generated to construct the GMM of PV errors. Next, the number of Gaussian components was determined using the BIC, and parameter estimation was performed using the EM algorithm. The voltage sensitivity matrix was then applied to the PV prediction error GMM through affine operations, resulting in the probability distribution of voltage fluctuations at each node. Then utilizing the forecast data for PV and load, as well as voltage violation indicators defined earlier, a threshold was set at 0.01. The PV integration capacity was incrementally increased, and the aforementioned steps were repeated to calculate the maximum PV integration capacity at which node voltages remain within permissible limits, representing the PV hosting capacity. The results of PV hosting capacity, calculated using both the proposed method and the Monte Carlo Newton Raphson power flow method, are presented in Table 1 below. The number of nodes prone to voltage violations at each integration capacity is illustrated in Fig. 7.

TABLE I. HOSTING CAPACITY OF IEEE 33-NODE DISTRIBUTION NETWORK

Test Network	Hosting Capacity in Distribution Networks (MW)	
	Probabilistic voltage sensitivity	***Monte Carlo Newton Raphson power flow***
IEEE 33	2.65	2.65

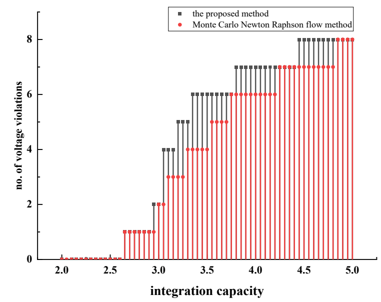

Fig. 7. The number of nodes susceptible to voltage violations in the 33-node system as the PV integration capacity into the distribution network increases.

In addition, the calculation results of hosting capacity will vary when different thresholds are set. Table II below provides the hosting capacity calculation results for both methods under varying thresholds in the IEEE 33-node distribution network.

TABLE II. HOSTING CAPACITY OF IEEE 33-NODE DISTRIBUTION NETWORK

Threshold	Hosting Capacity in Distribution Networks (MW)	
	Probabilistic voltage sensitivity	***Monte Carlo Newton Raphson power flow***
0.01	2.65	2.65
0.02	2.90	2.95
0.03	3.15	3.25
0.04	3.30	3.35
0.05	3.55	3.65

From Tables I and II, the proposed method demonstrates accuracy in estimating the hosting capacity of distributed PVs in distribution networks. Additionally, as the threshold setting increases, the hosting capacity calculation results for the distribution network also increase. Apart from improved accuracy, the proposed method also exhibits a significant advantage in terms of computation time. When the threshold is set to 0.05, the execution times for both methods in calculating the PV hosting capacity in the IEEE 33-node distribution network are provided in Table III.

TABLE III. COMPARISON OF EXECUTION TIMES FOR THE TWO METHODS IN CALCULATING PV HOSTING CAPACITY

Test Network	Execution Times (min)	
	Probabilistic voltage sensitivity	***Monte Carlo Newton Raphson power flow***
IEEE 33	17.03	45.17

In Table III, it can be inferred that in the IEEE 33-node test network, the proposed method executes faster than the Monte Carlo Newton Raphson power flow method. As the network scale increases, this speed difference is likely to further widen.

V. Conclusion

This paper proposes a distributed PV hosting capacity calculation method based on probabilistic voltage sensitivity, which is effective for nodes experiencing arbitrary power variations. In comparison to Monte Carlo Newton Raphson method, the proposed approach reduces computational complexity and time. It provides robust support for distribution network planning and operation, holding significant importance for enhancing the distribution system's hosting capacity and promoting the integration of distributed PV.

References

[1] DONG Yichao, WANG Shouxiang, YAN Bingke. Review on evaluation methods and improvement techniques of DG hosting capacity in distribution network [J]. Power Grid Technology, 2019, 43(07): 2258-2266.

[2] Koutroumpezis G N, Safigianni A S. Optimum allocation of the maximum possible distributed generation penetration in a distribution network [J]. Electric Power Systems Research, 2010, 80(12): 1421-1427.

[3] LIANG Haiping, WANG Cui, WANG Zhengping, XING Zhikun ZHENG Lianyue,et al. Evaluation of distributed PV integration capacity based on improved FPA algorithm [J]. Renewable Energy Resources, 2019, 37(02): 190-198.

[4] LUAN Weijie, JIANG Xianwei, ZHANG Jietan, CHENG Haozhong, SUN Shihang, et al. Consumptive ability analysis for distributed PV generation considering active management [J]. Electric Power Construction 2016, 37(01): 137-143.

[5] LI Zhenkun, BAO Xinyu, SHAO Yuying, PENG Peng, WANG Wenjun. Studying accommodation ability of distributed PV considering various voltage regulation measures [J]. Power System Protection and Control, 2018(8): 10-16.

[6] HUANG Wei, WU Junyong, LU Siqi, HAO Liangliang. Maximum PV penetration considering voltage deviation and harmonic constraints [J]. Power System Protection and Control, 2016, 44(6): 49-55.

[7] LIANG Zhifeng, XIA Junrong, SUN Mengmeng, ZHOU Chang. Data driven assessment of distributed PV hosting capacity in distribution network [J]. Power Grid Technology, 2020, 44(07): 2430-2439.

[8] S Munikoti, M Abujubbeh, K Jhala,B Natarajan. A novel framework for hosting capacity analysis with spatio-temporal probabilistic voltage sensitivity analysis[J]. International Journal of Electrical Power & Energy Systems, 2022, 134: 107426.

[9] MSS Abad, J MA, D Zhang, AS Ahmadyar, H Marzooghi. Probabilistic assessment of hosting capacity in radial distribution systems [J]. IEEE Transactions on Sustainable Energy, 2018, 9(4): 1935-1947.

[10] K Li, Z Ma, D Robinson, J Ma. Identification of typical building daily electricity usage profiles using Gaussian mixture model-based clustering and hierarchical clustering[J]. Applied Energy, 2018, 231: 331-342.

[11] Y. Yang, W. Wu, B. Wang, M. Li. Analytical reformulation for Stochastic unit commitment considering wind power uncertainty with Gaussian mixture model[J]. IEEE Transactions on Power Systems, 2020, 35(4): 2769-2782.

[12] S Munikoti, B Natarajan, K Jhala, K Lai. Probabilistic voltage sensitivity analysis to quantify impact of high PV penetration on unbalanced distribution system [J]. IEEE Transactions on Power Systems, 2021, 36(4): 3080-3092.

[13] BM HoPVe, D Lew, M Milligan, E Gomez-Lazaro, XG Larsen, et al. Wind power forecasting error distributions: An international comparison [C]//11th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants. 2012.

[14] Y. Chen, J. Wen, S. Cheng. Probabilistic load flow method based on Nataf transformation and latin hypercube sampling[J]. IEEE Transactions on Sustainable Energy, 2013, 4(2): 294-301.