

Potential Assessment of Spatial Correlation to Improve Maximum Distributed PV Hosting Capacity of Distribution Networks

Han Wu, Yue Yuan, Junpeng Zhu, Kejun Qian, and Yundai Xu

Abstract—Successful distributed photovoltaic (PV) planning now requires a hosting capacity assessment process that accounts for an appropriate model of PV output and its uncertainty. This paper explores how the PV hosting capacity of distribution networks can be increased by means of spatial correlation among distributed PV outputs. To achieve this, a novel PV hosting capacity assessment method is proposed to account for arbitrary geographically dispersed distributed PVs. In this method, the empirical relation between the spatial correlation coefficient and distance is fitted by historical data in one place and then applied to model the joint probability distribution of PV outputs at a neighboring location. To derive the PV hosting capacity at candidate locations, a stochastic PV hosting capacity assessment model that aims to maximize the PV hosting capacity under thermal and voltage constraints is proposed. Benders decomposition algorithm is also employed to reduce the computational cost associated with the numerous sampling scenarios. Finally, a rural 59-bus distribution network in Suzhou, China, is used to demonstrate the effectiveness of the proposed PV hosting capacity assessment methodology and the significant benefits obtained by increasing geographical distance.

Index Terms—Copula, mixed-integer cone programming, PV capacity assessment, spatial correlation, stochastic program.

I. INTRODUCTION

AS a kind of promising renewable energy sources to reduce the consumption of traditional fossil energy, the penetration of distributed photovoltaic (PV) has proliferated rapidly in recent years [1]. However, ever-increasing penetration of distributed PV raises the uncertainty of the distribution network and causes many emerging technical problems such as reverse power flow and overvoltage [2]-[5], which poses severe challenges to distribution network planners.

Therefore, the PV hosting capacity assessment task is now required to help distribution network planners understand the maximum PV hosting capacity that can be absorbed by existing distribution networks and ensure distribution networks continue to operate reliably [6].

The aim of hosting capacity assessment task is to derive the maximum hosting capacity (MHC) of a given distribution network without violating operation constraints. For now, various technologies have been developed to achieve MHC. In [7] and [8], a Monte Carlo simulation is employed to determine MHC. In [9] and [10], a sensitive analysis of feeder characteristics, distributed generation (DG) locations, and inverter features is presented. Besides simulation and sensitivity analysis, optimization method is also a mainstream approach to obtain precise MHC results under actual operation environment. In [11] and [12], an AC power flow based non-linear programming is employed to acquire precise MHC results. In [13], the robust optimization is employed to tackle the uncertainty of daily PV forecasting. Based on precise power flow functions and advanced optimization methods, many active management schemes like network reconfiguration [14], voltage and reactive power control [15], [16] have also been studied to access their impact to hosting capacity. Since inflexible operation rules usually lead to an uneconomic decision in planning, some works try to balance the DG curtailment risk and investment cost. For instance, [17] and [18] employ the conditional value at risk (CVaR) and stochastic dominance technique to gain a more efficient planning result.

Despite that the impact of network structure and management has been studied in determining MHC, a fundamental phenomenon that relates to the aggregated PV outputs has been ignored. Due to the variability of cloud patterns, solar irradiance usually fluctuates over both space and time, which leads to a correlation of solar irradiance probability distributions among dispersed sites [9]. Thus, the geographical dispersion of the PV stations may significantly affect the aggregated PV outputs [19]. Recently, this phenomenon has been observed and verified by many references at transmission level [20], [21]. In [20], a versatile probability model of multiple PV farms is proposed for probabilistic power flow. In [21], the impact of the smoothing effect on PV forecasting error is modeled and simulated. However, the discussions in [20] and [21] are all focused on operation problem, and the

Manuscript received: December 21, 2020; accepted: May 8, 2021. Date of CrossCheck: May 8, 2021. Date of online publication: June 7, 2021.

This work was supported in part by the National Key Research and Development Program of China (No. 2016YFB0900100) and in part by the National Natural Science Foundation of China (No. 51807051).

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

H. Wu (corresponding author), Y. Yuan, J. Zhu, and Y. Xu are with the College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China, and H. Wu is also with the Nanjing Institute of Technology, Nanjing 211167, China (e-mail: wuhanichina@vip.qq.com; yyuan@hhu.edu.cn; jzhu@hhu.edu.cn; hanxyd@126.com).

K. Qian is with the Suzhou Power Supply Branch of State Grid Jiangsu Electric Power Company, Suzhou 215004, China (e-mail: kejun.qian@xjtlu.edu.cn).

DOI: 10.35833/MPCE.2020.000886



geographic scale in them is far larger than a distribution network. For distribution network, [22] and [23] study the joint probability distribution of spatially distributed clear-sky index and apply it in probabilistic load flow simulations. References [22] and [23] demonstrate that the probability distributions of vital operation indexes, i. e., node voltages and power losses, with spatial correlation differ substantially from those with a non-spatial approach. Thus, the result of distributed PV hosting capacity should also be impacted by the spatial correlation. It will result in incorrect capacity assessment and inappropriate PV installation scheme without consideration of this correlation [24]. However, there is few quantitative methods that consider the spatial correlation in present distributed PV hosting capacity assessment studies. From the perspective of promoting PV accommodation, it is urgent to build an efficient MHC assessment model that considers spatial correlation among adjacent distributed PV outputs in distribution networks.

To this end, this paper proposes a novel PV hosting capacity assessment method considering spatial correlation among adjacent distributed PV outputs. In this paper, the empirical relation between the distance and correlation coefficient of PV outputs is revealed by a set of real PV output data in China. Then, the Gaussian Copula is employed to model the joint probability distribution of distributed PV outputs at N adjacent stations. Compared with the variance-covariance-based method [25], the Copula-based joint probability distribution provides more information about the dependence relations. After that, a novel PV hosting capacity assessment model based on the mixed-integer second-order cone programming (MISOCP) is presented to compute the MHC with correlated PV output data samples. In the proposed model, the non-linear AC power flow is represented by a convex second-order cone programming (SOCP) model to ensure the accuracy of MHC results under thermal and voltage constraints. Finally, an increasingly tight linear cut is utilized to shrink the relaxation gap of the second-order cone relaxation (SOCR) [26], [27] of the power flow function, and a Benders decomposition algorithm is employed to reduce the whole computational cost. Major contributions of this paper are summarized as follows.

1) The potential of spatial correlation among distributed

PVs to improve the MHC is investigated by a real case in Suzhou, China. This provides a feasible option to increase the ability of distribution networks to host distributed PVs.

2) In order to consider the spatial correlation of arbitrary geographical spread distributed PVs in the MHC assessment progress, a novel MISOCP-based PV hosting capacity assessment model and its solving algorithm are proposed.

The rest of the paper is organized as follows. In Section II, the spatial correlation model of adjacent distributed PV outputs is presented. Section III describes the stochastic and chance-constrained PV hosting capacity assessment model based on MISOCP. Section IV presents the Benders decomposition algorithm and the cutting plane method. Numerical results are presented in Section V. Finally, conclusions are drawn in Section VI.

II. SPATIAL CORRELATION MODEL OF ADJACENT DISTRIBUTED PV OUTPUTS

In this section, the relation between the spatial correlation coefficient of PV outputs and the distances is revealed by a historical data set from a real case in Suzhou, China. Then, the Gaussian Copula is applied to build the joint probability distribution of arbitrary geographical spread distributed PVs based on the empirical relation between the spatial correlation coefficient and distance. Finally, a sampling method is presented to generate correlated PV output samples from the joint probability distribution.

A. Data Source

The data for this study are collected from 18 grid-connected distributed PV stations in Suzhou, China. Due to the subtropical monsoon climate, cumulus and broken clouds are often occurrences in Suzhou, which results in significant and frequent variations in the incoming PV outputs. The measured PV arrays are located at two districts in Suzhou, i. e., Zhangjiagang and Wujiang, with the stations grouped around 31.86°N , 120.55°E and 31.80°N , 120.54°E . These distributed PV stations spread over an area of approximately 10 km and 15 km and are numbered 1-10 and 11-18, respectively, as shown in Fig. 1. Measurements of PV output were taken and recorded at 5-min intervals at each station in the year 2018.

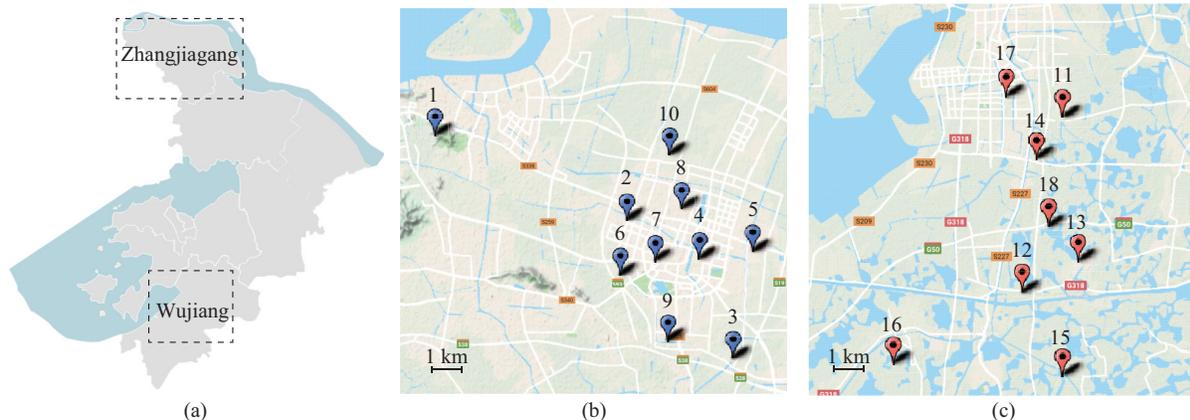


Fig. 1. Location of measured PV stations. (a) Location of two districts in Suzhou, China. (b) PV stations in Zhangjiagang. (c) PV stations in Wujiang.

In order to eliminate the useless zero PV outputs at night, only the data from 9 a.m. to 3 p.m. are used for each day. Furthermore, to avoid the impact of bad data, the time periods with zero PV outputs in the daytime are omitted in the test, which leads to a data set of 19149 data points for each location. Since we focus on the correlation and fluctuation of the PV outputs, the impact of the PV installation capacity can be omitted by normalizing the PV outputs by their capacities. Finally, the correlation matrix and distance matrix are presented in Figs. 2 and 3, respectively.

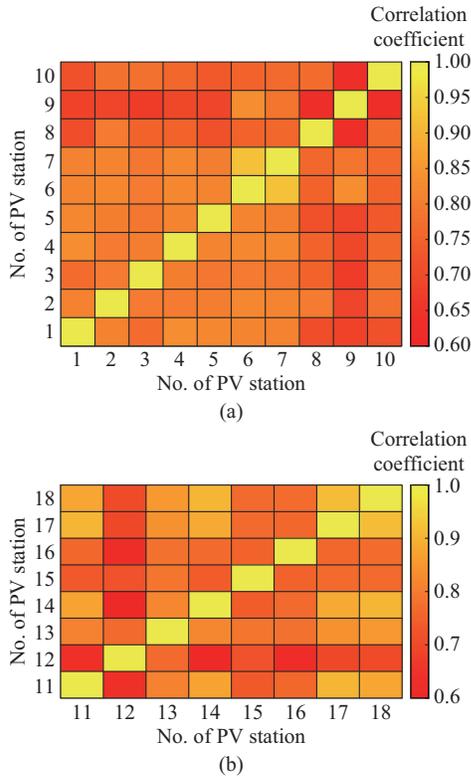


Fig. 2. Correlation coefficient matrix of PV outputs. (a) PV stations in Zhangjiagang. (b) PV stations in Wujiang.

By visual inspection, the correlation coefficient in Fig. 2 shows a strong relation to the distance in Fig. 3. That is, the shorter the distance between neighboring locations is, the higher the correlation will be.

B. Probabilistic Model and Definitions

The output at each PV station can be regarded as stochastic variables x_i , with probability distribution function $f_{x_i}(x_i)$ and cumulative distribution function $F_{x_i}(x_i)$. Since the distribution network usually covers a small geographical area, the PV stations are assumed to be affected by the same type of cloud patterns and cloud movements. Accordingly, all stochastic variables x_1, x_2, \dots, x_N can be modeled by the same probability distribution function.

Then, the Pearson correlation coefficient ρ_{ij} is utilized to represent the linear correlation between two PV stations:

$$\rho_{ij} = \frac{Cov(X_i, X_j)}{\sigma_i \sigma_j} \tag{1}$$

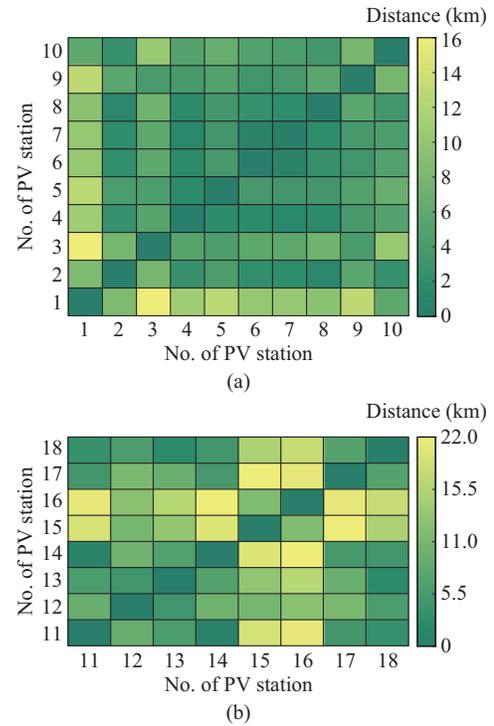


Fig. 3. Distance matrix of PV outputs. (a) PV stations in Zhangjiagang. (b) PV stations in Wujiang.

where X_i and X_j are the sample sets of x_i and x_j , respectively; $Cov(X_i, X_j)$ is the covariance between X_i and X_j ; and σ_i and σ_j are the corresponding standard deviations. The relation between the spatial correlation coefficient and distance can be described by:

$$\rho_{ij} = f(\xi_{ij}) \tag{2}$$

where ξ_{ij} is the distance between two PV stations.

Equation (2) presents a general form of the relation between the spatial correlation coefficient and distance. The specific form of (2) should be determined by the historical local PV outputs and the real distance among PV stations. In this paper, we take the data set of Suzhou as an example and present the fitting result in Fig. 4.

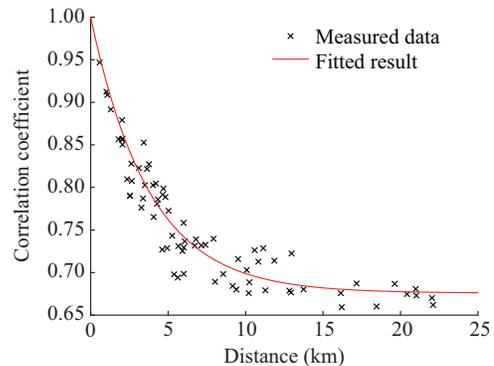


Fig. 4. Empirical correlation coefficient of PV outputs versus distance for all pairs of locations calculated from historical data in Suzhou, China.

Figure 4 shows that there exists a relation between the geographical distance and the correlation coefficient of PV

outputs. The relation is similar to that between geographical distance and the correlation coefficient of solar irradiance. In our case, when the geographical distance comes to 5 km, the correlation coefficient of PV outputs decreases to 0.75, which may have a strong impact on the power flow of distribution network and the MHC.

To further reveal the relation between the spatial correlation coefficient and distance and model correlations for arbitrary network configurations, (2) is fitted to the PV output data in Suzhou as:

$$\rho_{ij} = 0.3241e^{-0.2647r_{ij}} + 0.6759 \quad (3)$$

Equation (3) is an exponential decay formulation, which is consistent with the empirical relation between the clear-sky index and distance in [22] and [23]. However, the decreasing rate of the curve is lower than that in [22] and [23], which results from the climate differences and the efficiency of PV modules.

C. Copula-based Correlation Modeling

The last subsection reveals the relation between the geographical distance and correlation coefficient of PV outputs; however, such relation cannot directly apply to the MHC evaluation due to the stochastic nature of PV outputs. A common approach to tackle the correlation is through sample-based stochastic programming, which uses large-scale data samples to represent the correlation among random variables.

When sampling correlated PV output data from N different sites, the joint probability distribution of all PV outputs is required. A convenient and common approach to obtain the joint probability distribution is through the Copula function. Sklar's theorem states that an N -dimensional joint probability distribution function can be decomposed into several marginal distribution functions and a Copula function, where the Copula function describes the dependence relation among variables. Therefore, the joint cumulative distribution function $F(x_1, x_2, \dots, x_N)$ is expressed by the marginal distribution function $F_{x_i}(x_i)$ with a Copula function $C(\cdot)$ [28]:

$$F(x_1, x_2, \dots, x_N) = C(F_{x_1}(x_1), F_{x_2}(x_2), \dots, F_{x_N}(x_N)) \quad (4)$$

The Copula function $C(\cdot)$ is a multivariate probability distribution function using uniform marginals of realizations of $F_{x_1}(x_1), F_{x_2}(x_2), \dots, F_{x_N}(x_N)$ as the input. To explain the Copula approach more specifically, we employ the inverse probability integral transform, i.e., $x_i = F_{x_i}^{-1}(u_i)$, and reformulate (4) as:

$$F_{x_1, x_2, \dots, x_N}(F_{x_1}^{-1}(u_1), F_{x_2}^{-1}(u_2), \dots, F_{x_N}^{-1}(u_N)) = C(u_1, u_2, \dots, u_N) \quad (5)$$

where u_i is the cumulative distribution value of variable x_i .

When employing the widely used Gaussian Copula function, the joint probability distribution function can be expressed as:

$$C(u_1, u_2, \dots, u_N) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_N), \rho_{ij}) \quad (6)$$

where $\Phi_{\Sigma}(\cdot)$ and $\Phi^{-1}(\cdot)$ are the joint probability distribution function and its inverse cumulative distribution function of the N -dimensional standard Gaussian distribution with the zero mean and correlation matrix Σ , respectively.

Sampling from the joint probability distribution is also

straightforward, and the procedure is as follows: ① draw samples from N -dimensional Gaussian distribution with the correlation matrix Σ given by the empirical spatial correlation model; ② feed the sampled values through Gaussian distribution to obtain the probability value; ③ feed the probability value to the inverse marginal distribution function $F_{x_i}^{-1}(x_i)$ and get the realizations of each PV output. Further description of the sampling process can be found in [29] and [30].

III. MATHEMATICAL FORMULATION OF PV HOSTING CAPACITY ASSESSMENT MODEL

In order to consider the stochastic and correlated PV outputs while deterring MHC, the PV hosting capacity assessment problem is formulated through the scenario-based stochastic programming, where a set of correlated discrete scenarios sampled from the aforementioned Gaussian Copula model are adopted to represent their uncertainties. In this section, we present the stochastic PV hosting capacity assessment model based on the sampling results of the Copula model, followed by its chance-constrained variant.

A. Stochastic PV Hosting Capacity Assessment Model

This stochastic model can find a proper PV hosting capacity that satisfies all constraints in all possible scenarios, which leads to a conservative hosting capacity result.

1) Objective Function

$$\max \sum_{i \in \Psi_{PV}} c_i^{PV} \quad (7)$$

where c_i^{PV} is the installed PV capacity at bus i ; and Ψ_{PV} is a set of candidate buses that are potentially connected to PV stations. In (7), the objective is to derive an optimal sizing and location plan to maximize the total PV hosting capacity.

2) PV Capacity Constraints

$$0 \leq c_i^{PV} \leq c_i^{PV, \max} \quad \forall i \in \Psi_{PV} \quad (8)$$

where $c_i^{PV, \max}$ is the maximum PV installation capacity at bus i .

Constraint (8) represents the limit of the maximum PV installation capacity of candidate PV installation locations, which is associated with the local environment. For simplicity, we assume the distribution network operator (DNO) has sufficient information, e.g., the maximum available area of the PV station, about the PV candidate location after a careful investigation. Thus, the maximum PV installation capacity is defined ahead.

3) PV Output Constraints

$$P_{i,s}^{PV} = \eta_{i,s}^{PV} c_i^{PV} \quad \forall s \in S, \forall i \in \Psi_{PV} \quad (9)$$

$$Q_{i,s}^{PV} = P_{i,s}^{PV} \tan \phi_{i,s} \quad \forall s \in S, \forall i \in \Psi_{PV} \quad (10)$$

where $P_{i,s}^{PV}$ and $Q_{i,s}^{PV}$ are the active and reactive PV outputs at bus i in scenario s , respectively; S is the set of all scenarios; $\eta_{i,s}^{PV}$ is the predicted efficiency factor of PV output at bus i in scenario s ; and $\phi_{i,s}$ is the power factor angle of PV output at bus i in scenario s . Equations (9) and (10) represent the expected PV output at bus i in scenario s .

Although the actual PV output can be impacted by the local temperature, tilt angle, and azimuth of the PV panel, for high-

lighting the impact of smoothing effect, we assume the PV outputs in different scenarios can be represented by a product of the efficiency coefficient of a PV panel $\eta_{s,i}^{PV}$ and PV capacity.

4) Power Balance Constraints

$$\sum_{k \in \delta(j)} P_{jk,s} - \sum_{i \in \pi(j)} (P_{ij,s} - \tilde{I}_{ij,s} r_{ij}) = P_{j,s}^{PV} - P_{j,s}^L \quad \forall s \in S, \forall j \in \Psi_n \quad (11)$$

$$\sum_{k \in \delta(j)} Q_{jk,s} - \sum_{i \in \pi(j)} (Q_{ij,s} - \tilde{I}_{ij,s} x_{ij}) + b_j \tilde{V}_{j,s} = Q_{j,s}^{PV} - Q_{j,s}^L \quad \forall s \in S, \forall j \in \Psi_n \quad (12)$$

where Ψ_n is the set of all buses; $\delta(j)$ and $\pi(j)$ are the sets of upstream and downstream buses of bus j , respectively; $P_{ij,s}$ and $Q_{ij,s}$ are the active and reactive power flows of branch ij in scenario s , respectively; $\tilde{I}_{ij,s}$ is the squared current of branch ij in scenario s ; r_{ij} and x_{ij} are the resistance and reactance of branch ij , respectively; $P_{j,s}^L$ and $Q_{j,s}^L$ are the active load demand at bus j and the reactive load demand at bus i in scenario s , respectively; and $\tilde{V}_{j,s}$ is the squared voltage magnitude at bus j in scenario s .

Equations (11) and (12) denote the active and reactive power injections at each bus in scenario s . The left-hand side of (11) and (12) represents the power flow from bus i to its directly connected bus j , and the right-hand side represents the power injection of bus i .

5) Power Flow Constraints

$$\tilde{V}_{j,s} = \tilde{V}_{i,s} - 2(P_{ij,s} r_{ij} + Q_{ij,s} x_{ij}) + \tilde{I}_{ij,s} (r_{ij}^2 + x_{ij}^2) \quad \forall s \in S, \forall ij \in \Phi_b \quad (13)$$

$$\left\| \begin{bmatrix} 2P_{ij,s} \\ 2Q_{ij,s} \\ \tilde{I}_{ij,s} - \tilde{V}_{i,s} \end{bmatrix} \right\|_2 \leq \tilde{I}_{ij,s} + \tilde{V}_{i,s} \quad \forall s \in S, \forall ij \in \Phi_b \quad (14)$$

where Φ_b is the set of all branches.

Equation (13) describes the voltage drop along branch ij , and (14) is the SOCR formula of the AC power flow model of a radial distribution network.

6) Bus Voltage Constraints

$$\tilde{V}_{sub,s} = V_{sub}^2 \quad \forall s \in S \quad (15)$$

$$(V_j^{\min})^2 \leq \tilde{V}_{j,s} \leq (V_j^{\max})^2 \quad \forall s \in S, \forall j \in \Psi_n \quad (16)$$

$$\tilde{I}_{ij,s} \leq (I_{ij}^{\max})^2 \quad \forall s \in S, \forall ij \in \Phi_b \quad (17)$$

where I_{ij}^{\max} is the upper limit of the current of branch ij ; $\tilde{V}_{sub,s}$ is the squared voltage magnitude at substation bus in scenario s ; V_{sub} is the voltage magnitude at substation bus; and V_j^{\min} and V_j^{\max} are the lower and upper limits of the squared voltage magnitude at bus j , respectively.

Equation (15) limits the voltage at the low-voltage level of the substation. The security constraint of the distribution network is represented by (16) and (17).

Comprising (7)-(17), the stochastic PV hosting capacity assessment model is in the formulation of SOCP. Note that in the stochastic model, the constraint will be satisfied in all the scenarios, i.e., no PV curtailment occurs in any circumstances.

B. Chance-constrained SOCP Model

The stochastic PV hosting capacity assessment model represents the correlated PV outputs by adopting a large number of possible scenarios. Nevertheless, the assumption of full accommodation in all scenarios leads to a conservative and uneconomic PV integration. In fact, with the help of advanced distribution automation and active distribution management, DNO can control the PV output in the extreme scenarios and curtail the PV output if necessary. Thus, the extreme scenario in the generated sample set can be ignored for achieving economic PV hosting capacity. To simulate the PV curtailment, we adopt the chance-constrained programming and build a chance-constrained PV hosting capacity assessment model by replacing the deterministic power balance constraints (11) and (12) with the joint chance constraint (18).

represents the correlated PV outputs by adopting a large number of possible scenarios. Nevertheless, the assumption of full accommodation in all scenarios leads to a conservative and uneconomic PV integration. In fact, with the help of advanced distribution automation and active distribution management, DNO can control the PV output in the extreme scenarios and curtail the PV output if necessary. Thus, the extreme scenario in the generated sample set can be ignored for achieving economic PV hosting capacity. To simulate the PV curtailment, we adopt the chance-constrained programming and build a chance-constrained PV hosting capacity assessment model by replacing the deterministic power balance constraints (11) and (12) with the joint chance constraint (18).

$$\left\{ \begin{array}{l} \Pr \left\{ \sum_{k \in \delta(j)} P_{jk,s} - \sum_{i \in \pi(j)} (P_{ij,s} - \tilde{I}_{ij,s} r_{ij}) = P_{j,s}^{PV} - P_{j,s}^L \right\} \geq 1 - \delta \\ \quad \forall s \in S, \forall i \in \Psi_n \\ \Pr \left\{ \sum_{k \in \delta(j)} Q_{jk,s} - \sum_{i \in \pi(j)} (Q_{ij,s} - \tilde{I}_{ij,s} x_{ij}) + b_j \tilde{V}_{j,s} = Q_{j,s}^{PV} - Q_{j,s}^L \right\} \geq 1 - \delta \\ \quad \forall s \in S, \forall i \in \Psi_n \end{array} \right. \quad (18)$$

where δ is the curtailment probability. The joint chance constraint (18) guarantees that the constraint violation probability is less than a predefined risk level, i.e., the power balance constraints (11) and (12) will not exceed in $(1-\delta) \times 100\%$ scenarios. The risk level δ provides a natural and direct way to reflect the risk preferences of the DNO. Generally, a high-risk level leads to a high PV penetration level, vice versa. Therefore, by adjusting the risk level, the DNO can obtain a trade-off between the curtailment risk and PV penetration level.

Although the chance-constrained programming naturally represents the operation risk in the model, the introduction of (18) also results in an obstacle in solving the capacity assessment problem [31]. Reference [29] suggests that by extracting a finite set of scenarios from the original distribution and then imposing chance constraints on those sampled scenarios, the chance-constrained optimization can be solved by an MISOCP solver. Specifically, a binary variable w_s is introduced for each scenario, then (18) can be reformulated through Big- M algorithm as:

$$\left\{ \begin{array}{l} -Mw_s \leq \sum_{k \in \delta(j)} P_{jk,s} - \sum_{i \in \pi(j)} (P_{ij,s} - \tilde{I}_{ij,s} r_{ij}) - P_{j,s}^{PV} + P_{j,s}^L \\ \quad \forall s \in S, \forall ij \in \Phi_b \\ \sum_{k \in \delta(j)} P_{jk,s} - \sum_{i \in \pi(j)} (P_{ij,s} - \tilde{I}_{ij,s} r_{ij}) - P_{j,s}^{PV} + P_{j,s}^L \leq Mw_s \\ \quad \forall s \in S, \forall ij \in \Phi_b \end{array} \right. \quad (19)$$

$$\left\{ \begin{array}{l} -Mw_s \leq \sum_{k \in \delta(j)} Q_{jk,s} - \sum_{i \in \pi(j)} (Q_{ij,s} - \tilde{I}_{ij,s} x_{ij}) + b_j \tilde{V}_{j,s} - Q_{j,s}^{PV} + Q_{j,s}^L \\ \quad \forall s \in S, \forall ij \in \Phi_b \\ \sum_{k \in \delta(j)} Q_{jk,s} - \sum_{i \in \pi(j)} (Q_{ij,s} - \tilde{I}_{ij,s} x_{ij}) + b_j \tilde{V}_{j,s} - Q_{j,s}^{PV} + Q_{j,s}^L \leq Mw_s \\ \quad \forall s \in S, \forall ij \in \Phi_b \end{array} \right. \quad (20)$$

$$\sum_{s \in S} \pi_s w_s \leq \delta \quad w_s \in \{0, 1\} \quad (21)$$

where M is a sufficiently large number; and π_s is the probability of the scenario s . If w_s is set to be 1, the associated constraint in the scenario s is ignored due to the large enough M in the Big- M algorithm. The total probability of scenario omission is constrained by (21). By restricting the number of w_s being 1, (21) works as an equivalence of chance constraint.

Remark: if $\delta=0$, we can obtain $w_s=0$ for all scenarios, which reduces the chance-constrained PV hosting capacity assessment model to the stochastic model.

As a result, the chance-constrained PV hosting capacity assessment problem is converted into a large-scale deterministic MISOCP problem. Although the MISOCP problem is easy to understand and program, numerous sampling scenarios are usually required to obtain an accurate approximation of the multivariate distribution, which leads to a large number of binary variables. Thus, the associated MISOCP problem cannot be efficiently solved by the off-the-shelf MISOCP solvers. It is necessary to develop an advanced algorithm to solve this problem.

IV. SOLUTION METHOD

Benders decomposition algorithm, which is a master-subproblem structured method, has been widely used in computing the scenario-based stochastic mixed-integer linear problem [32]-[34]. Herein, we extend the basic Benders decomposition algorithm to compute the stochastic MISOCP formulation and its chance-constrained variant. Furthermore, to improve the exactness of the SOCR, the cutting plane method [26] is introduced subsequently.

A. Benders Decomposition for Chance-constrained PV Hosting Capacity Assessment Problem

1) Subproblem

The role of subproblem in the Benders decomposition is to check the feasibility of planning results determined by the master problem in each scenario and feed the feasibility cut to the master problem at the next iteration.

$$\left\{ \begin{array}{l} \min \theta_s^k = \sum_{i \in \Psi_n} (\varepsilon_{i,s}^+ + \varepsilon_{i,s}^- + \tau_{i,s}^+ + \tau_{i,s}^-) \\ \text{s.t. (13)-(17)} \\ \sum_{k \in \partial(j)} P_{jk,s} - \sum_{i \in \pi(j)} (P_{ij,s} - \tilde{I}_{ij,s} r_{ij}) - P_{j,s}^{PV} + \\ P_{j,s}^L + \varepsilon_{i,s}^+ - \varepsilon_{i,s}^- = 0 \quad \forall ij \in \Phi_b, s \in S \\ \sum_{k \in \partial(j)} Q_{jk,s} - \sum_{i \in \pi(j)} (Q_{ij,s} - \tilde{I}_{ij,s} r_{ij}) + b_j \tilde{V}_{j,s} - Q_{j,s}^{PV} + \\ Q_{j,s}^L + \tau_{i,s}^+ - \tau_{i,s}^- = 0 \quad \forall ij \in \Phi_b, s \in S \\ \mu_{i,s}^k \cdot \hat{c}_i^{PV,k} = c_i^{PV,k} \quad \forall i \in \Psi_{PV}, s \in S \end{array} \right. \quad (22)$$

where θ_s^k is the objective of the subproblem in scenario s at iteration k ; $\varepsilon_{i,s}^+$, $\varepsilon_{i,s}^-$, $\tau_{i,s}^+$, $\tau_{i,s}^-$ are the non-negative slack variables of power balance constraints at bus i in scenario s ; $\hat{c}_i^{PV,k}$ is the fixed value of first-stage variables c_i^{PV} at iteration k ; and $\mu_{i,s}^k$ is the dual variables for (25) at iteration k .

By introducing non-negative slack variables $\varepsilon_{i,s}^+$, $\varepsilon_{i,s}^-$, $\tau_{i,s}^+$,

and $\tau_{i,s}^-$ to the power balance constraints (12) and (13), the feasibility of the subproblem is ensured. The objective of the subproblem is to minimize the sum of all slack variables. If the objective θ_s^k is larger than a preset threshold, e.g., 0 in this paper, then the feasibility cut (23) will be generated and added into the master problem.

$$-Mw_s \leq \hat{\theta}_s^k + \sum_{i \in \Psi_{PV}} \mu_{i,s}^k (c_i^{PV} - \hat{c}_i^{PV,k}) \leq Mw_s \quad (23)$$

where $\hat{\theta}_s^k$ is the optimal objective of the feasibility check subproblem in scenario s at iteration k . The binary variable w_s indicates the inclusion or omission of the feedback from the feasibility check subproblem in scenario s . If w_s is set to be 1, the Benders cut of scenario s will be ignored. Note that since the objective function (7) only contains the installed PV capacity, there is no need to return optimality cuts from subproblems to the master problem. We only check the feasibility of the subproblem in each scenario and feed the feasibility cut.

2) Master Problem

The master problem aims to maximize the total PV hosting capacity, concerning constraints (8)-(10), (21), and the feasibility cut (23). The bilinear master problem of iteration k is described as:

$$\left\{ \begin{array}{l} \max \sum_{i \in \Psi_{PV}} c_i^{PV} \\ \text{s.t. (8)-(10), (21), (23)} \end{array} \right. \quad (24)$$

B. Iterative Correction of SOCR Based on Cutting Planes

The power flow constraint (14) is an SOCR formula, while the SOCR may not be bounded under some circumstances. To address this issue, we look back to the original power flow formulation before SOCR as in (25).

$$\tilde{I}_{ij,s} = \frac{P_{ij,s}^2 + Q_{ij,s}^2}{\tilde{V}_{i,s}} \quad \forall s \in S, \forall ij \in \Phi_b \quad (25)$$

References [35], [36] demonstrate that in a radial network, when the objective function is convex and monotonically increasing in each active power injection and the initial optimal power flow is feasible, (25) can be replaced by (26), known as SOCR, and further reformulated as constraint (14).

$$\tilde{I}_{ij,s} \geq \frac{P_{ij,s}^2 + Q_{ij,s}^2}{\tilde{V}_{i,s}} \quad \forall s \in S, \forall ij \in \Phi_b \quad (26)$$

However, due to the reversed power flow resulted from high PV penetration level, the objective function in the hosting capacity assessment problem is not monotonically increasing with the power injection [37], [38], but leads to over voltage, which obstructs the application of SOCR. Therefore, a cutting plane method is introduced to guarantee the exactness of SOCR [26].

The key idea of cutting plane method is to keep the branch currents at the minimum value when the objective function reaches the optimal. Thus, the branch current would not exceed its theoretical value, and the SOCR (26) will be tight and exact. To this end, linear inequalities are iteratively added to the relaxation and thus the gap between the optimal solution of the non-relaxed problem and the relaxed problem

is eliminated. The cutting plane at iteration $k+1$ is expressed as [26]:

$$\sum_{ij \in \Phi_B} r_{ij} \tilde{r}_{ij,s}^{k+1} \leq \sum_{ij \in \Phi_B} r_{ij} \frac{(P_{ij,s}^k)^2 + (Q_{ij,s}^k)^2}{\tilde{V}_{i,s}^k} \quad \forall s \in S \quad (27)$$

C. Overall Procedure for Solving Capacity Assessment Model

The whole procedure to solve the stochastic and chance-constrained PV hosting capacity assessment problem is summarized as follows.

1) Initialization

Set the iteration counter $k=0$, the lower bound $LB=0$, the upper bound $UB=+\infty$, and the optimality tolerance $\zeta=0.01$.

2) Iteration

Step 1: compute the master problem. Derive its optimal capacity solution $\hat{c}_i^{PV,k}$, and the scenario indicator w_s^k by an MISOCP solver. Update the upper bound $UB = \sum_{i \in \Psi_{PV}} \hat{c}_i^{PV,k}$.

Step 2: for all scenarios, solve the feasibility check sub-problem with cutting planes integrated by MISOCP solver, and derive the optimal value $\hat{\theta}_s^k$. If $\hat{\theta}_s^k$ is larger than zero, i. e., the original problem is infeasible, then generate the Benders feasibility cut (23) and feedback to the master problem.

Step 3: construct an MISOCP problem for deriving the LB . This problem is composed of the objective function (7) and constraints (8)-(10), (13)-(17), (19), (20) with fixed binary variable \hat{w}_s from the master problem and the cutting planes. Then, solve this problem and obtain its optimal objective. Finally, update LB .

3) Stop criteria

If $|UB-LB|/LB \leq \zeta$, terminate with the PV capacity solution associated with the latest lower bound LB . Otherwise, update the iteration $k=k+1$ and return to *Step 1*.

V. CASE STUDIES

In this section, the impact of spatial correlation of PV outputs on the MHC of distribution networks is demonstrated by the numerical tests based on the proposed PV hosting capacity assessment model and algorithm. First, the Gaussian Copula-based PV output sampling method is demonstrated. Then, a 59-bus 10 kV rural distribution network near Suzhou is used to show the difference before and after the correlation is considered in the capacity assessment scheme. The effectiveness of the proposed PV hosting capacity assessment model and algorithm is also demonstrated.

All the algorithms are executed on an HP Z840 workstation with Intel Xeon E5-2650v4 CPU running at 2.20 GHz with 16 GB RAM. The proposed model is programmed in MATLAB 2018b. The hosting capacity assessment model, along with the Benders decomposition algorithm, and the cutting plane method are programmed and solved using the General Algebraic Modeling System (GAMS) [39] software with commercial solver CPLEX 12.10.

A. Correlation Model Fitting and Performance

The Gaussian Copula-based joint probability distribution modeling method is tested based on the empirical correlation

of the 18 PV stations in Suzhou.

The quantile-quantile (Q-Q) plot and scatter plot for two arbitrary locations (PV1 and PV2) are given in Fig. 5 to visualize the fitting result. Since it is difficult to visualize more dimensions, only one location pair (PV1 and PV2) is presented in Fig. 5. Figure 5(a) and (b) demonstrates that the marginal distributions generated from the Copula-based method are the same as the measured data. The scatter plots in Fig. 5(c) and (d) show that the sampled result of the Copula-based method is consistent with the shape of measured result of the joint probability distributions. In summary, it is clear that the Copula-based method is able to tackle the correlation and fit the general shape of the joint probability distribution.

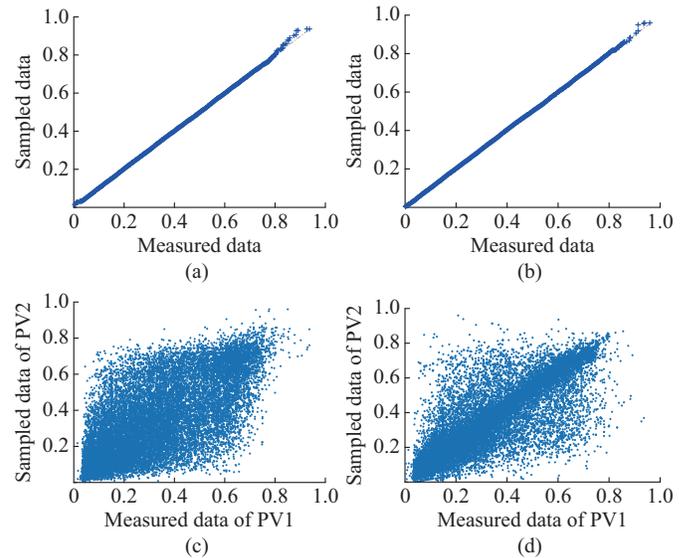


Fig. 5. Q-Q plot and scatter plot for PV1 and PV2. (a) Q-Q plot of measured data and sampled data of PV1. (b) Q-Q plot of measured data and sampled data of PV2. (c) Scatter plot of sampled data of PV1 and PV2. (d) Scatter plot of measured data of PV1 and PV2.

B. Hosting Capacity Assessment of a Real Distribution Network

The hosting capacity assessment test is twofold. First, the impact of the spatial correlation of PV outputs on the MHC based on the fixed and varied PV allocation cases is presented. Then, four plans with varied mean station separation distances are tested to demonstrate the impact of station separation distances on the MHC.

The tested distribution network is a typical 59-bus 10 kV rural distribution network in Suzhou. The single-line geographic diagram of the distribution network is depicted in Fig. 6. The voltage range of all nodes is set to be [0.93, 1.07] p.u. and the voltage of the substation node is set to be 1.0 p.u.. In this distribution network, there are 15 PV stations, located at buses 8, 14, 15, 19, 22, 24, 29, 35, 40, 41, 46, 50, 54, 55, and 59, respectively. The distance matrix and the corresponding correlation matrix of PV outputs of these candidate buses are shown in Fig. 7. The loading level has a significant impact on the hosting capacity, herein, the average load [40] at daytime (9 a.m. to 3 p.m.) is employed to ob-

tain the most common PV hosting capacity. The line apparent power is set to be 5.03 MW, and other line parameters can be found in [37].

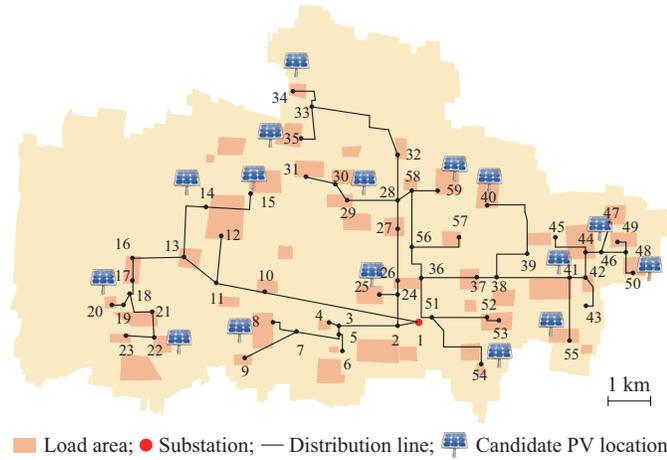


Fig. 6. Single-line geographic diagram of tested 59-bus 10 kV rural distribution network in Suzhou.

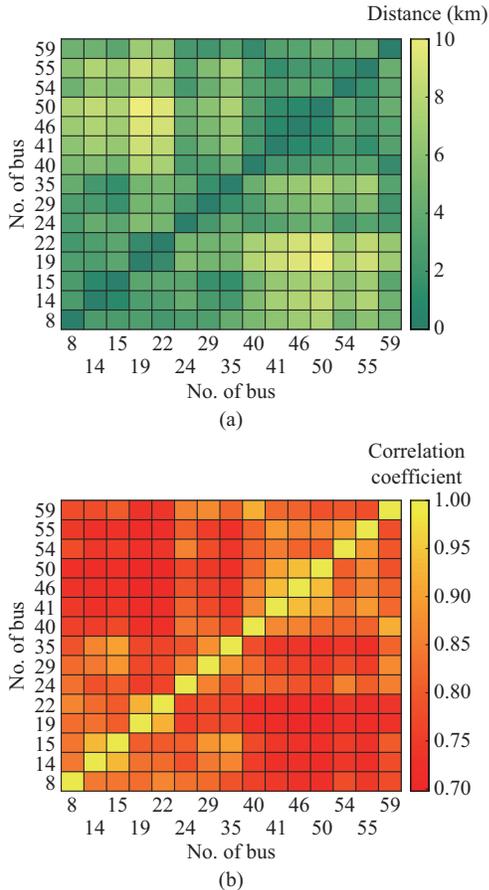


Fig. 7. Distance and correlation coefficient matrices of 15 candidate buses. (a) Distance matrix. (b) Correlation coefficient matrix.

1) Impact of Spatial Correlation of PV Outputs on MHC

Two cases, i.e., fixed and varied, are tested in this part to show the impact of spatial correlation on the MHC.

1) Fixed case: one arbitrarily chosen PV output is used as a representative for the whole network.

2) Varied case: the Gaussian Copula with the correlation value from the spatial correlation model in (3) is applied to model the PV output at each PV station.

In the fixed case, the PV outputs are sampled directly from an arbitrary chosen marginal distribution and applied to all sites. In the varied case, correlated values are sampled from the joint N -dimensional distribution, using the Gaussian Copula. The fixed case is a benchmark, which shows the assessment result of the PV hosting capacity without considering the spatial correlation, i.e., the correlation coefficient is regarded to be 1, while the varied case shows the PV hosting capacity assessment result considering the spatial correlation (correlation declines with the distance). The stochastic PV hosting capacity assessment model is tested first to show the impact of PV correlation on stochastic capacity assessment result. Then, the chance-constrained PV hosting capacity assessment model is tested to show the impact of curtailment risk level on the chance-constrained capacity assessment result. For each case, one thousand scenarios are generated to achieve the final MHC result.

The hosting capacity assessment results of the two cases are given in Table I. It can be observed from Table I that the total hosting capacities in the fixed case are highly underestimated than that in the varied case. It is because the high spatial correlation will aggravate the fluctuation of the total PV outputs and leads to more extreme scenarios. The results given by full correlation assumption will waste branch capacity of the existing distribution network and result in higher investment cost than actual cases. Therefore, it is important to take spatial correlation into consideration when assessing the PV hosting capacity of the distribution network.

TABLE I
HOSTING CAPACITY ASSESSMENT RESULTS OF 59-BUS DISTRIBUTION NETWORK WITH AND WITHOUT SPATIAL CORRELATION

Bus No.	Hosting capacity assessment result (MW)	
	Fixed case	Varied case
8	2.857	0.259
14	1.207	0
15	0	2.438
19	0.740	4.037
22	0.143	0
24	2.636	2.257
29	0	0
35	0	0
40	0	0
41	0	0
46	0	0
50	0.672	0
54	0	0
55	0	4.971
59	0	0
Total	8.255	13.962

Then, the impact of chance constraints is explored by solving the chance-constrained PV hosting capacity assessment model with Benders decomposition. The MHCs at each can-

didate bus with curtailment risk levels of 95%, 90%, 85%, and 80% are presented in Fig. 8(a). Note that the chance-constrained PV hosting capacity assessment model reduces to the stochastic model when the risk level δ is set to be 0. Figure 8(b) shows the total hosting capacity increases monotonically with the curtailment risk level, and the impact of correlation is even more pronounced in cases where the chance constraints are considered. Hence, by adjusting the curtailment risk level, the DNO can have a balance between the cost of the expansion plan which results in a high-capacity system and the desired level of security against risks.

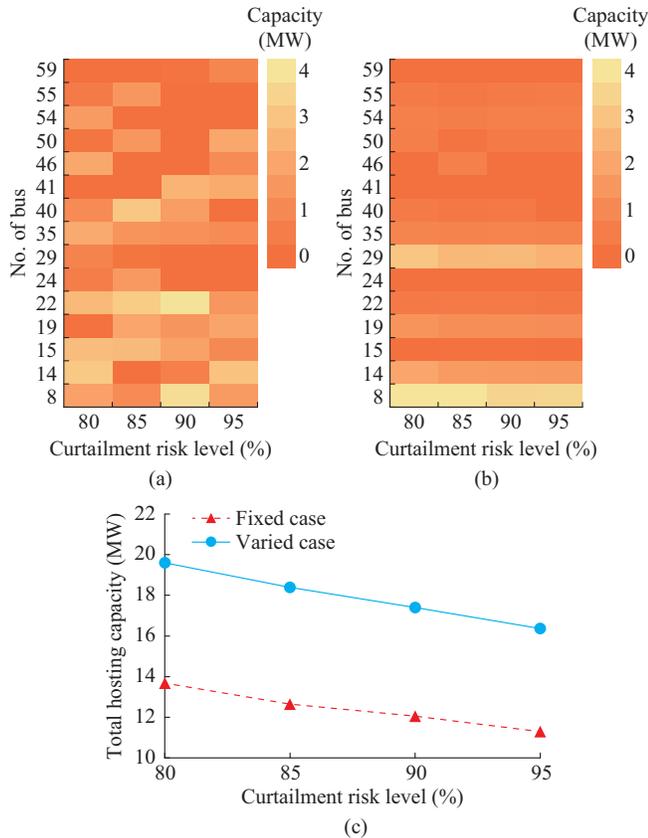


Fig. 8. Hosting capacity of 59-bus distribution network at different curtailment risk levels. (a) Hosting capacity at each candidate bus in fixed case. (b) Hosting capacity at each candidate bus in varied case. (c) Total hosting capacities of two cases with different curtailment risk levels.

2) Impact of Station Separation Distance on MHC

In order to test the impact of increasing station separation distances on the MHC of PV outputs, four different allocation plans for eight PV stations, shown as Fig. 9, are tested in this case study. The mean station separation distance is used to represent the degree of separation, and a larger mean station separation distance means a larger separation. In this case study, the mean station separation distances for allocation plans A, B, C, and D are set to be 0.87, 2.16, 3.05, and 4.26 km, respectively. The exponential correlation function in (3) is also employed to calculate the spatial correlation among stations in the network.

Similar to the fixed and varied cases in the last part, we also consider the fixed and varied cases in each plan.

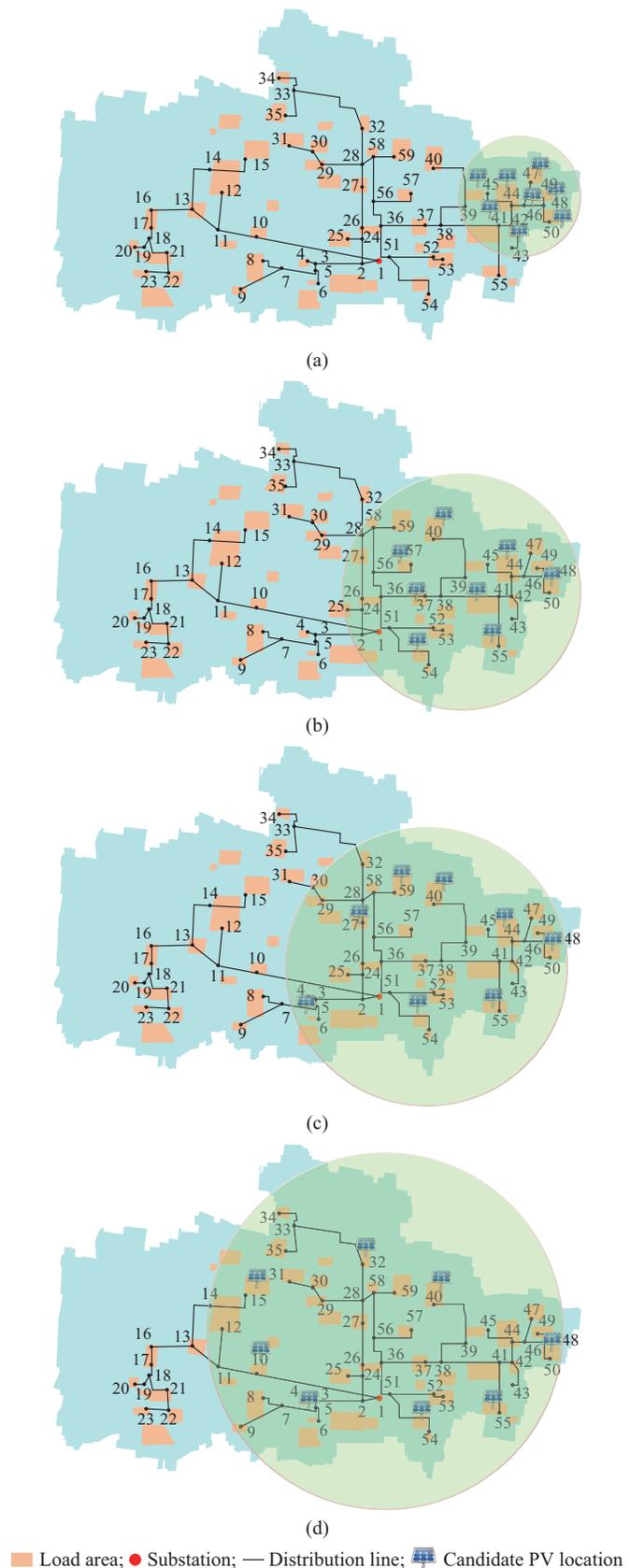


Fig. 9. Four PV allocation plans with different mean station separation distances. (a) Plan A. (b) Plan B. (c) Plan C. (d) Plan D.

Hosting capacity assessment results of the two cases with different mean station separation distances are given in Table

II. Similar to the hosting capacity assessment result in the last part, it can be seen from Table II that the total hosting capacity will be underestimated if the varied spatial correlation is not considered. Furthermore, by comparing the hosting capacity assessment results of plans A to D, we find that the results have a monotonically increasing relation with the mean station separation distance. That is because PV can be installed in different feeders when the geography room is not limited.

TABLE II
HOSTING CAPACITY ASSESSMENT RESULTS WITH DIFFERENT STATION SEPARATION DISTANCES

Plan	Mean distance (km)	Hosting capacity in fixed case (MW)	Hosting capacity in varied case (MW)	Difference (%)
A	0.87	2.14	2.36	10.28
B	2.16	3.29	3.78	14.89
C	3.05	4.54	5.37	18.28
D	4.26	6.89	8.76	27.14

Such separation significantly enhances the hosting capacity of the distribution network. Comparing the fixed and varied cases with different separation distances, we can also find that the increase of hosting capacity is even more pronounced when considering the distance-induced correlation, specifically, the hosting capacity increases about 10%-27% if the correlation is related to the geographical distance. And a higher separation distance leads to a higher increase.

3) Computational Performance

Table III provides the computational time of the Big- M and Benders decomposition algorithms with different curtailment risk levels.

TABLE III
COMPUTATIONAL TIME OF BIG- M AND BENDERS DECOMPOSITION ALGORITHMS WITH DIFFERENT CURTAILMENT RISK LEVELS

Algorithm	Case	Computational time (s)				
		100%	95%	90%	85%	80%
Big- M	Fixed	524.10	10022.26	9960.42	10054.85	10035.87
	Varied	581.54	8822.91	10042.88	10110.41	10035.32
Benders	Fixed	276.73	3894.39	5889.25	7094.14	7256.33
	Varied	221.89	3740.02	5637.69	6716.69	7198.43

From Table III, it can be observed that the Benders decomposition algorithm significantly accelerates the overall computation progress. Specifically, the Benders decomposition algorithm is about twice as fast as the traditional Big- M algorithm in both fixed and varied cases. Compared with the full accommodation scenario ($1-\delta=100\%$), the introduction of chance constraints significantly increases the computational time. However, the variation of the curtailment risk level has little impact on the computational time of the traditional Big- M algorithm. On the contrary, the Benders decomposition algorithm is more sensitive to the curtailment risk level. A larger risk level leads to a heavier computational burden. It is understandable because a larger risk level brings more integer variables and thus increases computational costs. Al-

though the computational time is longer than an hour in most cases, the proposed model and algorithm are still usable for a distribution network planning problem.

The Gaussian Copula fitting and sampling are employed on the same platform. According to our experience, the running time of the fitting and sampling program is less than 1 min, which is negligible in the whole hosting capacity assessment process.

VI. CONCLUSION

This paper investigates the potential of spatial correlation to improve the MHC of PV in the distribution network. At first, an exponential model is fitted to the empirical correlation of PV output samples in Suzhou, China. Then, the Gaussian Copula is employed to build the joint probability distribution of PV outputs. After that, a stochastic and chance-constrained PV hosting capacity assessment model is proposed to evaluate the impact of the spatial correlation on the PV hosting capacity, and the Benders decomposition algorithm is applied to reduce the computational cost.

Numerical results on a 59-bus rural distribution network in Suzhou, China, verify the effectiveness of the proposed model. The testing results give us four useful insights: ① it demonstrates that the correlation of PV outputs in a medium geographical size (10 km) has an exponential decay relation with the distance between two PV sites; ② it demonstrates that the correlation of PV outputs has a strong impact on the PV hosting capacity for the distribution network, specifically, the hosting capacity would be much higher when considering the relation of spatial correlation and geographical distance; ③ the curtailment risk level of PV outputs also has a significant impact on the hosting capacity; ④ the proposed hosting capacity assessment algorithm can consider the correlation among distributed PVs and figure out the hosting capacity effectively and efficiently.

Future studies include considering the temporal correlation of PV outputs, taking active distribution management and integrated energy system into account, discussing the impact of energy storage on the hosting capacity, etc.

REFERENCES

- [1] A. Calcabrini, H. Ziar, O. Isabella *et al.*, "A simplified skyline-based method for estimating the annual solar energy potential in urban environments," *Nature Energy*, vol. 4, no. 3, pp. 206-215, Mar. 2019.
- [2] Y. Riffonneau, S. Bacha, F. Barruel *et al.*, "Optimal power flow management for grid connected PV systems with batteries," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 3, pp. 309-320, Jul. 2011.
- [3] M. Kolenc, I. Papic, and B. Blazic, "Assessment of maximum distributed generation penetration levels in low voltage networks using a probabilistic approach," *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 505-515, Jan. 2015.
- [4] M. M. Hague and P. Wolfs, "A review of high PV penetrations in LV distribution networks: present status, impacts and mitigation measures," *Renewable & Sustainable Energy Reviews*, vol. 62, pp. 1195-1208, Sept. 2016.
- [5] R. Torquato, D. Salles, C. O. Pereira *et al.*, "A comprehensive assessment of PV hosting capacity on low-voltage distribution systems," *IEEE Transactions on Power Delivery*, vol. 33, no. 2, pp. 1002-1012, Apr. 2018.
- [6] M. S. S. Abad, J. Ma, D. Zhang *et al.*, "Probabilistic assessment of hosting capacity in radial distribution systems," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1935-1947, Mar. 2018.
- [7] A. Dubey and S. Santoso, "On estimation and sensitivity analysis of

- distribution circuit's photovoltaic hosting capacity," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2779-2789, Jul. 2017.
- [8] E. Zio, M. Delfanti, L. Giorgi *et al.*, "Monte Carlo simulation-based probabilistic assessment of DG penetration in medium voltage distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 852-860, Jan. 2015.
- [9] F. Ding and B. Mather, "On distributed PV hosting capacity estimation, sensitivity study, and improvement," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1010-1020, Jul. 2017.
- [10] H. M. Ayres, W. Freitas, M. C. D. Almeida *et al.*, "Method for determining the maximum allowable penetration level of distributed generation without steady-state voltage violations," *IET Generation, Transmission & Distribution*, vol. 4, no. 4, pp. 495-508, Apr. 2010.
- [11] H. Al-Saadi, R. Zivanovic, and S. F. Al-Sarawi, "Probabilistic hosting capacity for active distribution networks," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2519-2532, Apr. 2017.
- [12] L. F. Ochoa, C. J. Dent, and G. P. Harrison, "Distribution network capacity assessment: variable DG and active networks," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 87-95, Nov. 2010.
- [13] E. Samani and F. Aminifar, "Tri-level robust investment planning of DERs in distribution networks with AC constraints," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3749-3757, Apr. 2019.
- [14] F. Capitanescu, L. F. Ochoa, H. Margossian *et al.*, "Assessing the potential of network reconfiguration to improve distributed generation hosting capacity in active distribution systems," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 346-356, May 2015.
- [15] S. Wang, S. Chen, L. Ge *et al.*, "Distributed generation hosting capacity evaluation for distribution systems considering the robust optimal operation of OLTC and SVC," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 1111-1123, Mar. 2016.
- [16] X. Chen, W. Wu, and B. Zhang, "Robust capacity assessment of distributed generation in unbalanced distribution networks incorporating ANM techniques," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 651-663, Sept. 2018.
- [17] D. Xiao, J. C. do Prado, and W. Qiao, "Optimal joint demand and virtual bidding for a strategic retailer in the short-term electricity market," *Electric Power Systems Research*, vol. 190, p. 106855, Jan. 2021.
- [18] M. K. AlAshery, D. Xiao, and W. Qiao, "Second-order stochastic dominance constraints for risk management of a wind power producer's optimal bidding strategy," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1404-1413, Jul. 2020.
- [19] U. Focken, M. Lange, K. Mönnich *et al.*, "Short-term prediction of the aggregated power output of wind farms - a statistical analysis of the reduction of the prediction error by spatial smoothing effects," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 90, no. 3, pp. 231-246, Mar. 2002.
- [20] W. Wu, K. Wang, B. Han *et al.*, "A versatile probability model of photovoltaic generation using pair copula construction," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1337-1345, Jun. 2015.
- [21] E. Nuño, M. Koivisto, N. A. Cutululis *et al.*, "On the simulation of aggregated solar PV forecast errors," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1889-1898, Mar. 2018.
- [22] J. C. Hernández, F. J. Ruiz-Rodríguez, and F. Jurado, "Technical impact of photovoltaic-distributed generation on radial distribution systems: stochastic simulations for a feeder in Spain," *International Journal of Electrical Power & Energy Systems*, vol. 50, pp. 25-32, Sept. 2013.
- [23] J. Widén, M. Shepero, and J. Munkhammar, "Probabilistic load flow for power grids with high PV penetrations using Copula-based modeling of spatially correlated solar irradiance," *IEEE Journal of Photovoltaics*, vol. 7, no. 6, pp. 1740-1745, Sept. 2017.
- [24] H. Wu, L. Wu, Y. Yuan *et al.*, "A PV hosting capacity assessment method with spatial correlation," in *Proceedings of 2019 IEEE Sustainable Power and Energy Conference (iSPEC)*, Beijing, China, Nov. 2019, pp. 1102-1007.
- [25] J. M. Morales, R. Mínguez, and A. J. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, vol. 87, no. 3, pp. 843-855, Mar. 2010.
- [26] S. Y. Abdelouadoud, R. Girard, F. P. Neirac *et al.*, "Optimal power flow of a distribution system based on increasingly tight cutting planes added to a second order cone relaxation," *International Journal of Electrical Power & Energy Systems*, vol. 69, pp. 9-17, Jul. 2015.
- [27] H. Gao, J. Liu, and L. Wang, "Robust coordinated optimization of active and reactive power in active distribution systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4436-4447, Jan. 2018.
- [28] J. Munkhammar and J. Widén, "Correlation modeling of instantaneous solar irradiance with applications to solar engineering," *Solar Energy*, vol. 133, pp. 14-23, Aug. 2016.
- [29] J. Munkhammar, J. Widén, and L. M. Hinkelman, "A Copula method for simulating correlated instantaneous solar irradiance in spatial networks," *Solar Energy*, vol. 143, pp. 10-21, Feb. 2017.
- [30] J. Widén, M. Shepero, and J. Munkhammar, "On the properties of aggregate clear-sky index distributions and an improved model for spatially correlated instantaneous solar irradiance," *Solar Energy*, vol. 157, pp. 566-580, Nov. 2017.
- [31] Z. Liu, Q. Wu, S. S. Oren *et al.*, "Distribution locational marginal pricing for optimal electric vehicle charging through chance constrained mixed-integer programming," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 644-654, May 2018.
- [32] M. Shahidehpour and F. Yong, "Benders decomposition: applying Benders decomposition to power systems," *IEEE Power and Energy Magazine*, vol. 3, no. 2, pp. 20-21, Mar. 2005.
- [33] B. Zeng, Y. An, and L. Kuznia. (2014, Mar.). Chance constrained mixed integer program: bilinear and linear formulations, and benders decomposition. [Online]. Available: <http://arxiv.org/abs/1403.7875>
- [34] S. M. Frank and S. Rebennack, "Optimal design of mixed AC-DC distribution systems for commercial buildings: a nonconvex generalized benders decomposition approach," *European Journal of Operational Research*, vol. 242, no. 3, pp. 710-729, May 2015.
- [35] L. Gan, N. Li, U. Topcu *et al.*, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Transactions on Automatic Control*, vol. 60, no. 1, pp. 72-87, Jun. 2015.
- [36] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 92-107, Aug. 2012.
- [37] S. Huang, Q. Wu, J. Wang *et al.*, "A sufficient condition on convex relaxation of AC optimal power flow in distribution networks," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1359-1368, Jun. 2017.
- [38] M. Farivar and S. H. Low, "Branch flow model: relaxations and convexification - part I," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554-2564, Aug. 2013.
- [39] GAMS. (2019, Jul.). GAMS - cutting edge modeling. [Online]. Available: <http://www.gams.com>
- [40] IEEE Dataport. (2019, Aug.). 59-bus distribution system in China. [Online]. Available: <http://dx.doi.org/10.21227/b5pg-rb19>
- Han Wu** received the M.E. degree and Ph.D. degree in electrical engineering from Hohai University, Nanjing, China, in 2015 and 2020, respectively. He is currently a Lecturer with Nanjing Institute of Technology, Nanjing, China. His research interests include distribution system planning, probabilistic power system analysis, and power system optimization.
- Yue Yuan** received the B.E. and M.Sc. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1987 and 1990, respectively, and the Ph.D. degree from Hiroshima University, Hiroshima, Japan, in 2002. He joined the Faculty of Xi'an Jiaotong University, in 1990. He has been with Hohai University, Nanjing, China, as a Professor, since 2003. His research interests include power system operation and optimization, renewable energy, and distributed generation.
- Junpeng Zhu** received the B.S. degree in applied mathematics and the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 2012 and 2017, respectively. He is currently a Lecturer with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. His research interests include distributed generations integration, optimization, and control of active distribution networks.
- Kejun Qian** received the Ph. D. degree in electrical engineering from Glasgow Caledonian University (GCU), Glasgow, UK, in 2010. He joined the Department of Electrical and Electronic Engineering, Xi'an Jiaotong Liverpool University, Suzhou, China, as an Academic Staff. He is currently a Senior Engineer with the State Grid Jiangsu Electric Power Company, Suzhou, China. His research interests include modeling of charging load of electric vehicles, V2G, smart grids, and renewable energy systems.
- Yundai Xu** received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2012. He is currently working toward the Ph.D. degree in electrical engineering from Hohai University, Nanjing, China. His research interests include optimal power flow and renewable energy forecasting.