Evaluating Connectable Capacity of Distributed Wind Generation in Distribution Networks Through a Bayesian Integrated Optimization Method

Evaluating Connectable Capacity of Distributed Wind Generation in Distribution Networks Through a Bayesian Integrated Optimization Method

Shijie Chen, Yue Xiang, *Senior Member, IEEE*, Wei Sun, *Member, IEEE*, Junyong Liu

***Abstract*—The effective evaluation of the connectable capacity of renewable energy plays a vital role in the development of a sustainable distribution network. Quantifying the capacity while considering network security and the local renewable accommodation policy is a challenge. This paper proposes a scenario-based bi-level mathematical model using a Bayesian integrated optimization method to evaluate and quantify the connectable capacity of distributed wind generation in distribution networks, which effectively integrates the characteristics of wind power and the local accommodation policy. The constraint on the generation curtailment ratio (CR) is innovatively designed to represent the renewable accommodation policy and integrated with network security constraints to coordinatively quantify the capacity. The model is solved by the Bayesian integrated optimization method. The regression-based algorithm greatly reduces the complexity of alternating iteration and improves the calculation efficiency. Practical cases are used to verify the effectiveness of the proposed method. Results indicate that the method is more efficient than traditional optimization algorithms, and CR integration ensures that the connectable capacity fits local renewable energy development policies well.**

***Index Terms*—Capacity evaluation, Bayesian optimization, distributed wind generation, renewable accommodation policy.**

I. Introduction

DISTRIBUTED wind generation (DWG) has been increasingly connected to distribution networks due to its environmental protection and greenness features. By December 2017, the total installed capacity of wind generation reached 539,123 MW worldwide [1]. The integration of DWG can effectively alleviate the operation pressure and increase the available capacity of the network [2]. However, wind power generation often needs curtailing due to the temporal uncertainty of wind speed and network constraints [3]. Therefore, a critical need exists for the policy of renewable accommodation to be considered when evaluating the connectable capacity of DWG in distribution networks.

DWG is a typical distributed generation (DG). The hosting capacity problem for DG units has been widely investigated [4]– [8]. "Hosting capacity" usually refers to the maximum DG capacity allowed to be connected to the distribution network under the premise of satisfying various network security constraints, such as node voltage, line capacity, and power quality [9]–[12]. Various methods of enhancing the hosting capacity have been proposed recently. An integrated heat and power dispatch model that considers the thermal inertia of the district heating network is proposed to improve the flexibility of the wind power accommodation in [13]. Other measures, such as energy storage coordination [4] and demand response [14], are also used to increase the hosting capacity.

The above studies contributed in evaluating and increasing the DG hosting capacity in a distribution network, but few studies have considered the impact of accommodation policies on hosting capacity. Thus, on the basis of "hosting capacity," "connectable capacity" is proposed to present the maximum capacity of the DG that can be connected to the distribution network with both policy and network security constraints. Given that the accommodation policy is mainly aimed at wind power, the evaluation problem is modeled for wind generations. Furthermore, a bi-level coordinative optimization model is proposed for evaluating the connectable capacity of DWG considering the renewable accommodation policy. The upper-level is formulated as a capacity evaluation model to evaluate the maximum capacity of DWG. The lower-level is formulated as a multi-period operation model to optimize the operation of DWG to improve renewable utilization. In addition, the constraint on the generation curtailment ratio (CR) is innovatively designed to present the renewable accommodation policies in different regions.

Intelligent optimization algorithms [15]–[17], such as genetic algorithms (GA) [15]–[16] and particle swarm optimization (PSO) [5][17], were used for solving the optimal capacity problem. However, existing intelligent optimization algorithms require many iterations to find the optimum and cannot deeply mine the complex coupling correlation between evaluation and operation to increase the convergence speed. In recent years, analytical methods were also used to evaluate the hosting capacity. Specifically, various methods such as polygonal inner-approximation [18], bender decomposition [19], and second-order cone relaxation [20], are used to convert non-convex and nonlinear power flow constraints into linear constraints, and the distribution network operation model is

This work was supported in part by the Young Scholar Support Program of the Chinese Society of Electrical Engineering under Grant CSEE-YESS-2018006, in part by the National Natural Science Foundation of China under Grant 51807127. (*Corresponding author: Yue Xiang*.).

S. Chen, Y. Xiang, J. Liu are with the College of Electrical Engineering, Sichuan University, Chengdu 610065, China (e-mail: ShijieCW@foxmail.com; xiang@scu.edu.cn; liujy@scu.edu.cn).

W. Sun is with School of Engineering, University of Edinburgh, Mayfield Road, Edinburgh EH9 3DW, U.K. (e-mail: W.Sun@ed.ac.uk).

converted into a mixed-integer linear programming (MILP) model, which is solved by commercial software such as CPLEX and GUROBI. However, when renewable accommodation policies are considered, the model will have bilinear terms and will be difficult to solve directly. Therefore, the above methods are inapplicable, and the efficiency of optimization must be urgently improved.

In this paper, a regression-based Bayesian integrated optimization method, which combines LP and Bayesian optimization, is proposed to solve the model. The lower-level is converted into LP through second-order cone relaxation [20] and polyhedral-based linearization [21] and then embedded in the upper-level as a soft constraint. Overall, an unconstrained model is established based on the penalty function method and solved by Bayesian optimization. The probabilistic surrogate model (PSM) is used to simulate the coupling correlation between evaluation and operation. The evaluation point (input of operation level) is actively selected based on the acquisition function.

Thus, the main contributions of this paper are as follows:

1) A novel bi-level coordinative optimization model, which effectively integrates the characteristics of wind power and the local accommodation policy to evaluate the connectable capacity of DWG, is proposed.
2) A regression-based Bayesian integrated optimization method, in which the mined correlation by Bayesian highly improves the computing efficiency, is proposed.
3) The CR is innovatively designed and integrated with traditional security constraints to quantify the connectable capacity of DWG. The integration of CR enables the evaluated capacity to satisfy local renewable energy development policies.

The rest of this paper is structured as follows. The proposed bi-level coordinative optimization model is introduced in Section II. Model solving based on Bayes' theorem is introduced in Section III. The proposed model is demonstrated in practical networks in Section IV. Finally, Sections V concludes this paper.

II. Bi-level Coordinative Optimization Model

A. Research Framework

As shown in Fig. 1, supposing that the output of each DWG can be controlled continuously in each time period, the operation level is optimized to minimize the average CR by controlling the DWG, and numerous physical constraints are included in it. The optimization variable is the power curtailment of each DWG in each time period, and the capacity is given by the evaluation level. CR presents the curtailment of renewable energy and defined as the ratio of the potential generation of the DWG to its power curtailment. After solving the operation level, the minimum average CR is sent to the evaluation level, which then updates the evaluation sample set, selects the next evaluation point, and sends it back to the operation level for another iteration until the stop criteria are met. Overall, the lower-level is converted into a soft constraint and embedded in the upper-level. Evaluation results that comply with the policy of renewable accommodation can be obtained after multiple iterations.

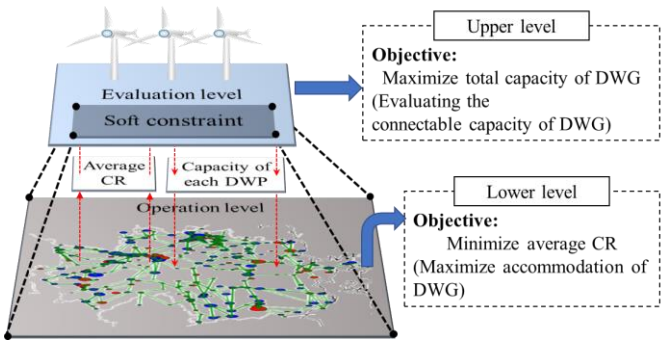

Fig. 1. Bi-level framework for evaluating the connectable capacity of DWG

B. Lower-lever: Multi-period operation of DWG

1) Objective Function

The objective of the lower-level is to minimize the average CR:

$$\min Z_1 = \lambda_{ave} = \frac{\sum_{\omega\in\Omega^S}\pi_\omega \sum_{t\in\Omega^T}\sum_{g\in\Omega^G}\lambda_{g,t,\omega}P_g^{Cap}}{N_t\sum_{g\in\Omega^G}P_g^{Cap}} \tag{1}$$

where λ_{ave} is the average CR, t is the index of the time period, and ω is the scenario index. λ_g is the CR of the g-th DWG, P_g^{Cap} is the capacity of the g-th DWG determined by the upper model, and N_t is the total number of time periods. A scenario-based method [22] is introduced to deal with the uncertainty of wind power generation. Let π_ω be the probability of the ω-th scenario and Ω^S the set of scenarios, while Ω^T and Ω^G are the respective sets of operating periods and DWGs.

2) Numerous Constraints

The model is subject to many constraints, and the constraints on the power flow are as follows:

$$\sum_{k\in\Omega_f(i)}\left[P_{ki,t,\omega}^{line}-R_{ki}\left(I_{ki,t,\omega}^{line}\right)^2\right]- \sum_{j\in\Omega_a(i)}P_{ij,t,\omega}^{line}+P_{i,t,\omega}^{WP}=P_{i,t,\omega}^{D}\quad \forall i\in\Omega^B,\forall t\in\Omega^T,\forall\omega\in\Omega^S \tag{2}$$

$$\sum_{k\in\Omega_f(i)}\left[Q_{ki,t,\omega}^{line}-X_{ki}\left(I_{ki,t,\omega}^{line}\right)^2\right]- \sum_{j\in\Omega_a(i)}Q_{ij,t,\omega}^{line}+Q_{i,t,\omega}^{WP}=Q_{i,t,\omega}^{D}\quad \forall i\in\Omega^B,\forall t\in\Omega^T,\forall\omega\in\Omega^S \tag{3}$$

$$\left(V_{i,t,\omega}\right)^2-\left(V_{j,t,\omega}\right)^2=2\left(R_{ij}P_{ij,t,\omega}^{line}+X_{ij}Q_{ij,t,\omega}^{line}\right)- \left(I_{ij,t,\omega}^{line}\right)^2\left[\left(R_{ij}\right)^2+\left(X_{ij}\right)^2\right]\quad \forall ij\in\Omega^{Line},\forall t\in\Omega^T,\forall\omega\in\Omega^S \tag{4}$$

$$\left(S_{ij,t,\omega}^{line}\right)^2=\left(P_{ij,t,\omega}^{line}\right)^2+\left(Q_{ij,t,\omega}^{line}\right)^2\quad \forall ij\in\Omega^{Line},\forall t\in\Omega^T,\forall\omega\in\Omega^S \tag{5}$$

$$\left(V_{i,t,\omega}\right)^2\left(I_{ij,t,\omega}^{line}\right)^2=\left(S_{ij,t,\omega}^{line}\right)^2\quad \forall ij\in\Omega^{Line},\forall t\in\Omega^T,\forall\omega\in\Omega^S \tag{6}$$

where Ω^B is the set of buses, Ω^{Line} is the set of lines, $\Omega_f(i)$ represents the end bus set of the branch with bus i as the head bus, and $\Omega_a(i)$ represents the end bus set of the branch with bus i as the head bus. R_{ki}, R_{ij}, X_{ki}, and X_{ij} are the resistance and reactance on branches ki and ij, respectively; I_{ki}^{line} and I_{ij}^{line} are the respective currents of lines ki and ij; P_{ij}^{line}, Q_{ij}^{line}, and S_{ij}^{line} are the respective active, reactive, and apparent powers of branch ij; P_{ki}^{line} and Q_{ki}^{line} are the active and reactive powers of branch ki, respectively; P_i^{WP} and Q_i^{WP} are the respective active and reactive powers injected by the DWG at bus i; and P_i^D and Q_i^D are the active and reactive powers of the load at bus i, respectively.

Constraints (2) and (3) represent the bus active and bus reactive power balances, respectively; Constraint (4) represents the voltage drop in the branch; Constraints (5) and (6) are branch power constraints. The main grid can supply power to the distribution network through the grid supply point (GSP). The active and reactive power constraints of the GSP are as follows:

$$\underline{P}^{GSP} \le P_{t,\omega}^{GSP} \le \bar{P}^{GSP} \quad \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{7}$$

$$\underline{Q}^{GSP} \le Q_{t,\omega}^{GSP} \le \bar{Q}^{GSP} \quad \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{8}$$

where $\underline{P}^{GSP}$ and $\bar{P}^{GSP}$ are the minimum and maximum output active powers of the GSP, respectively; and $\underline{Q}^{GSP}$ and $\bar{Q}^{GSP}$ are the minimum and maximum output reactive powers of the GSP, respectively.

Constraints (7) and (8) limit the GSP's active and reactive power output, respectively. When bus i is the GSP, Constraints (2) and (3) can be replaced with Constraints (9) and (10):

$$P_{t,\omega}^{GSP} - \sum_{j \in \Omega_a(i)} P_{ij,t,\omega}^{line} + P_{i,t,\omega}^{WP} = P_{i,t,\omega}^{D} \quad \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{9}$$

$$Q_{t,\omega}^{GSP} - \sum_{j \in \Omega_a(i)} Q_{ij,t,\omega}^{line} + Q_{i,t,\omega}^{WP} = Q_{i,t,\omega}^{D} \quad \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{10}$$

For network security, the bus voltages should be limited within a range, as shown in (11):

$$\underline{V}_i \le V_{i,t,\omega} \le \bar{V}_i \quad \forall i \in \Omega^N, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{11}$$

where $\underline{V}_i$ and $\bar{V}_i$ are the minimum v and maximum voltage levels of bus *i*, respectively.

3) Operation and Control of DWG

DWG adopts power factor and output control technology, which enables operation with leading or lagging power factor and the continuous control of the output. In practice, the power factor must be kept within a certain range. Given the difference between the local demand and output of DWG over time, if DWG injects its output into the distribution network, the bus voltage may exceed the limit and the network security will be violated. In this case, the DWG output must be curtailed. Power curtailment is represented by introducing a negative generation [8].

$$P_{g,t,\omega}^{WP} \bullet \tan\left(\cos^{-1}\left(PF_g^{\max}\right)\right) \le Q_{g,t,\omega}^{WP} \le P_{g,t,\omega}^{WP} \bullet \tan\left(\cos^{-1}\left(PF_g^{\min}\right)\right)$$
$$\forall g \in \Omega^G, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{12}$$

$$P_{g,t,\omega}^{WP} = P_g^{Cap}\beta_t - P_{g,t,\omega}^{Curt} \quad \forall g \in \Omega^G, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{13}$$

$$P_{g,t,\omega}^{Curt} = \lambda_{g,t,\omega} P_g^{Cap} \beta_{t,\omega} \quad \forall g \in \Omega^G, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{14}$$

where $PF_g^{\max}$ and $PF_g^{\min}$ respectively represent the maximum and minimum power factors that the *g*-th DWG can operate in; $\beta_{t,\omega}$ is the ratio of the maximum power that the DWG can generate to the DWG capacity in the *t*-th time period, which is determined by the wind speed in the *t*-th time period; and P_g^{Curt} is the power curtailment of the *g*-th DWG.

Constraint (12) limits the reactive power output of the DWG according to the range of the power factor; Constraint (13) represents the actual active output of the DWG; the CR of the DWG is defined in Constraint (14).

C. Upper-level: evaluating connectable DWG capacity

The upper-level is established to evaluate the connectable capacity of the DWG with the accommodation policy as the constraint:

$$\max Z_2 = \sum_{g \in \Omega^G} P_g^{Cap} \tag{15}$$

$$s.t. \quad Z_1 \le \bar{\lambda} \tag{16}$$

where $\bar{\lambda}$ is the limit of the average CR, presenting the policy of renewable energy accommodation.

The models represented by (1)–(14) are embedded in the upper-level as a soft constraint, and the upper-level represented by (15)–(16) is converted into an unconstrained optimization model by the penalty function. The details are introduced in Section III.

III. Bayesian Integrated Optimization Method

The lower-level is a strictly non-convex and non-linear model, which is difficult to solve directly. In this section, the nonlinear programming (NLP) is converted into LP through second-order cone relaxation [20] and polyhedral-based linearization [21]. By using the penalty function method, the linearized lower-level is converted into a soft constraint and embedded in the upper-level. Then, an unconstrained model is established and solved by the proposed Bayesian integrated method.

A. Linearization of power flow constraints

Given the strict nonlinearity of Constraints (2)–(6), the proposed model is difficult to solve directly in its original NLP formulation. Thus, the lower-level must be transformed into a LP model.

By introducing auxiliary variables, namely, $\left(I_{ki,t,\omega}^{line}\right)^2 = \tilde{I}_{ki,t,\omega}^{line}$, $\left(I_{ij,t,\omega}^{line}\right)^2 = \tilde{I}_{ij,t,\omega}^{line}$, $\left(V_{i,t,\omega}\right)^2 = \tilde{V}_{i,t,\omega}$, and $\left(V_{j,t,\omega}\right)^2 = \tilde{V}_{j,t,\omega}$, Constraints (2)–(4), (6), and (11) can be converted into the following constraints:

$$\sum_{k \in \Omega_f(i)} \left[P_{ki,t,\omega}^{line} - R_{ki}\tilde{I}_{ki,t,\omega}^{line}\right] - \sum_{j \in \Omega_a(i)} P_{ij,t,\omega}^{line} + P_{i,t,\omega}^{WP} = P_{i,t,\omega}^{D}$$
$$\forall i \in \Omega^N, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{17}$$

$$\sum_{k \in \Omega_f(i)} \left[Q_{ki,t,\omega}^{line} - X_{ki}\tilde{I}_{ki,t,\omega}^{line}\right] - \sum_{j \in \Omega_a(i)} Q_{ij,t,\omega}^{line} + Q_{i,t,\omega}^{WP} = Q_{i,t,\omega}^{D}$$
$$\forall i \in \Omega^N, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{18}$$

$$\hat{V}_{i,t,\omega} - \hat{V}_{j,t,\omega} = 2\left(R_{ij}P_{ij,t,\omega}^{line} + X_{ij}Q_{ij,t,\omega}^{line}\right) - \hat{I}_{ij,t,\omega}^{line}\left[\left(R_{ij}\right)^2 + \left(X_{ij}\right)^2\right]$$
$$\forall ij \in \Omega^{Line}, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{19}$$

$$\hat{V}_{i,t,\omega}\hat{I}_{ij,t,\omega}^{line} = \left(S_{ij,t,\omega}^{line}\right)^2 \quad \forall ij \in \Omega^{Line}, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{20}$$

$$\left(\underline{V}_i\right)^2 \le \hat{V}_{i,t,\omega} \le \left(\bar{V}_i\right)^2 \quad \forall i \in \Omega^N, \forall t \in \Omega^T, \forall \omega \in \Omega^S \tag{21}$$

Constraints (5) and (20) are still non-convex constraints. The left side of (20) can be rewritten as:

$$\hat{V}_{i,t,\omega}\hat{I}_{ij,t,\omega}^{line} = \left[\left(\hat{V}_{i,t,\omega} + \hat{I}_{ij,t,\omega}^{line}\right)/2\right]^2 - \left[\left(\hat{V}_{i,t,\omega} - \hat{I}_{ij,t,\omega}^{line}\right)/2\right]^2 \tag{22}$$

Then, based on the second-order cone relaxation, Constraints (5) and (22) are transformed into convex constraints, as shown in (23) and (24):

$$S_{ij,t,\omega}^{line} \ge \sqrt{\left(P_{ij,t,\omega}^{line}\right)^2 + \left(Q_{ij,t,\omega}^{line}\right)^2} \tag{23}$$

$$\left[\left(\hat{V}_{i,t,\omega} + \hat{I}_{ij,t,\omega}^{line}\right)/2\right]^2 \ge \sqrt{\left(S_{ij,t,\omega}^{line}\right)^2 + \left[\left(\hat{V}_{i,t,\omega} - \hat{I}_{ij,t,\omega}^{line}\right)/2\right]^2} \tag{24}$$

Considering the errors caused by the second-order cone relaxation [20], the objective function of the lower-level model

is modified as:

$$\min Z_1^{'} = \min Z_1 + \min \gamma \sum_{\omega \in \Omega^S} \sum_{ij \in \Omega^{Line}} \sum_{t \in \Omega^T} S_{ij,t,\omega}^{line} \tag{25}$$

where γ is a sufficiently small number compared to min Z_1 which is 10^{-6} here.

Constraints (23) and (24) have the following form:

$$d_3 = \sqrt{(d_1)^2 + (d_2)^2} \tag{26}$$

A highly accurate method based on polyhedral approximation is used to linearize Constraints (23) and (24) [21]:

$$\tau_l \geq |d_1|, \ \alpha_l \geq |d_2|, \ l = 0 \tag{27}$$

$$\begin{cases} \tau_l = \tau_{l-1} \cos\left(\dfrac{\pi}{2^{l+1}}\right) + \alpha_{l-1} \sin\left(\dfrac{\pi}{2^{l+1}}\right) \\ \alpha_l \geq \left| \alpha_{l-1} \cos\left(\dfrac{\pi}{2^{l+1}}\right) - \tau_{l-1} \sin\left(\dfrac{\pi}{2^{l+1}}\right) \right| \end{cases}, \forall l = 1, \cdots, L \tag{28}$$

$$\tau_l \leq z_3, \quad \alpha_l \leq \zeta_l \tan\left(\frac{\pi}{2^{l+1}}\right), l = L \tag{29}$$

where τ_l and α_l are auxiliary variables, and L is a constant.

Linear Constraints (27)–(29) approximate (26) in such a way that [23]:

$$d_3 \cos\left(\pi / 2^{L+1}\right) \geq \sqrt{(d_1)^2 + (d_2)^2} \tag{30}$$

B. Unconstrained processing of the upper model

A bi-level mathematical model is raised to evaluate the connectable capacity of the DWG. However, the operation model composed of Formulas (1)–(14) and the evaluation model composed of Formulas (15)–(16) are two independent optimization models, which make the proposed bi-level model difficult to solve directly. Therefore, an optimization model coupled with operation and evaluation models must be built. Thus, the penalty function method is used to convert Constraint (16) into a soft constraint, and the upper-level is converted into an unconstrained model. This reformulation makes possible the gradually finding of an optimal solution that meets the soft constraint using the Bayesian integrated method. The upper-level is then converted into the following form:

$$\max Z_2^{'} = Z_2 - M \bullet \delta \tag{31}$$

$$\delta = \begin{cases} 0 & Z_1 \leq \bar{\lambda} \\ \left(Z_1 - \bar{\lambda}\right)^2 & Z_1 > \bar{\lambda} \end{cases} \tag{32}$$

where M is a large positive number and δ is a penalty factor.

C. Model solving based on Bayesian optimization

The correlation between the evaluation and operation levels presents a black box characteristic, and the computation efficiency can be improved by effectively mining the above correlation. Therefore, regression-based Bayesian optimization is designed to solve the model. The PSM and the acquisition function are two parts of the algorithm.

1) Probabilistic Surrogate Model

The PSM is updated based on Bayes' theorem:

$$p(f \mid C_{1:r}) = \frac{p(C_{1:r} \mid f)\, p(f)}{p(C_{1:r})} \tag{33}$$

where f is the unknown objective function, in this study, $C_{1:r}=\{(x_1,f_1), (x_2,f_2), \cdots, (x_r,f_r)\}$ is the set of evaluated sample points, f_r is the function value obtained by the last evaluation, $p(C_{1:r}| f)$ is the likelihood distribution of f, $p(f)$ is the prior probability distribution model of f, $p(C_{1:r})$ is the marginal likelihood distribution after f is marginalized, and $p(f |C_{1:r})$ is the posterior probability distribution of f.

Here, $x = \left(P_1^{Cap}, P_2^{Cap}, \cdots, P_{N_g}^{Cap}\right)$ is an evaluation point on a N_g-dimensional space as the input of the operation level which represents the capacity of each DWG, where N_g is the number of DWG. f denotes the objective function Z_2 of the upper-level.

PSM includes parametric and non-parametric models [24]. Among them, Gaussian Process (GP) is widely used and has a great fitting performance. GP is expressed as follows:

$$\begin{cases} f(x) \sim gp\left(m(x), k\left(x, x^{'}\right)\right) \\ m(x) = E\left[f(x)\right] \\ k\left(x, x^{'}\right) = E\left[\left(f(x) - m(x)\right)\left(f\left(x^{'}\right) - m\left(x^{'}\right)\right)\right] \end{cases} \tag{34}$$

where $k\left(x,x^{'}\right)$ is the covariance function,and $m(x)$ is the mean function. To simplify the calculation, suppose $m(x)$=0, then $f(x) \sim gp\left(0, k\left(x, x^{'}\right)\right)$. The prior distribution of the objective function can be expressed as:

$$p\left(f_{1:r} \mid \mathrm{C}_{1:r}\right) \sim \mathcal{N}\left(0, K_r\right) \tag{35}$$

where $\mathcal{N}(\bullet)$ is the Gaussian distribution probability density function, $f_{1:r} = \{f_1, f_2, \cdots, f_r\}$ is the set of evaluated sample points, and K_r is the covariance matrix from the covariance function as shown in (36).

$$K_r = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_r) \\ \vdots & \ddots & \vdots \\ k(x_r, x_1) & \cdots & k(x_r, x_r) \end{bmatrix} \tag{36}$$

When a new evaluation sample $\left(x_{r+1}, f_{r+1}\right)$ is added to the sample set, (37) is used to update the covariance matrix, and (38) is used to estimate the posterior probability of f_{r+1}:

$$\begin{cases} K_{r+1} = \begin{bmatrix} K_r & k_r^T \\ k_r & k\left(x_{r+1}, x_{r+1}\right) \end{bmatrix} \\ k_r = \left[k\left(x_{r+1}, x_1\right), k\left(x_{r+1}, x_2\right), \cdots, k\left(x_{r+1}, x_r\right)\right] \end{cases} \tag{37}$$

$$\begin{cases} p\left(f_{r+1} \mid C_{1:r+1}, x_{r+1}\right) \sim \mathcal{N}\left(\mu_{r+1}, \sigma_{r+1}^2\right) \\ \mu_{r+1} = k_{r+1}^T K_{r+1}^{-1} f_{1:r+1} \\ \sigma_{r+1}^2 = k_{r+1}\left(x_{r+1}, x_{r+1}\right) - k_{r+1}^T K_{r+1}^{-1} k_{r+1} \\ k_{r+1} = \left[k\left(x_{r+1}, x_1\right), k\left(x_{r+1}, x_2\right), \cdots, k\left(x_{r+1}, x_{r+1}\right)\right] \end{cases} \tag{38}$$

where μ_{r+1} represents the predicted mean, and σ_{r+1}^2 represents the predicted covariance。

2) Acquisition function

The acquisition function is the basis for searching the next input point of the operation level. Here, expected improvement (EI) is used as the acquisition function. EI is defined as the expectation of the improved function $u(x) = \max\left(0, f(x) - f^*\right)$, where f^* is the optimal value in the evaluated sample set. The acquisition function is defined as follows:

$$\begin{aligned} AC &= E\left[u_r(x) \mid x_{1:r}, C_{1:r}\right] \\ &= (\mu_r - f_r^*) \bullet \Phi\left(\frac{\mu_r - f_r^*}{\sigma_r}\right) + \sigma_r \bullet \phi\left(\frac{\mu_r - f_r^*}{\sigma_r}\right) \end{aligned} \tag{39}$$

where $\Phi(\bullet)$ is the cumulative density function of the standard

normal distribution, and $\phi(\cdot)$is the probability density function of the standard normal distribution. The next evaluation point is collected according to (40):

$$x_{r+1} = \arg\max_x AC(x) \tag{40}$$

D. Model solving process

The proposed bi-level optimization model is solved by the Bayesian integrated method. Here, two stopping criteria are used: "Max iterations number" and "Max stall iterations number." The flowchart and the specific implementation are as follows:

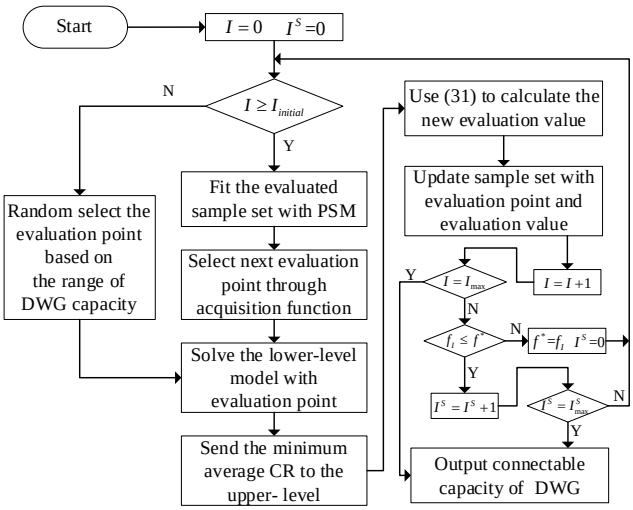

Fig. 2. Model solving process.

1) Set the iteration counter and the stall iteration counter to $I=0$ and $I^S=0$, respectively.
2) If the iteration counter is greater than or equal to the initial number of iterations, go to Step 3); otherwise, randomly select the initial evaluation point based on the range of the DWG capacity, and then go to Step 4).
3) Fit the set of evaluated samples using the PSM and select the next evaluation point through the acquisition function.
4) Solve the multi-period DWG operation model is solved according to the selected evaluation point. Obtain and send the minimum average CR to the upper-level.
5) Calculate the objective function of the upper-level using Formula (31). Update the sample set with the evaluation and evaluation values.
6) Update iteration counter: $I=I+1$.
7) If the iteration counter equals the maximum number of iterations, output the connectable capacity of the DWG; otherwise, go to Step 8).
8) If the evaluation value is less than or equal to the optimal value in the evaluated sample set, update the stall iteration counter $I^S=I^S+1$, and go to Step 9); otherwise, update the optimal value in the evaluated sample set $f^*=f_l$, reset the stall iteration counter to $I^S=0$, and go to Step 2).
9) If the stall iteration counter equals the maximum number of stall iterations, output the connectable capacity of the DWG; otherwise, go to Step 2).

IV. Case Study

A. Case description

1) Scenario analysis

To choose a reasonable operation period, scenario analysis is performed firstly, based on one-year wind speed data in Northern Scotland [25] as shown in Fig.3.

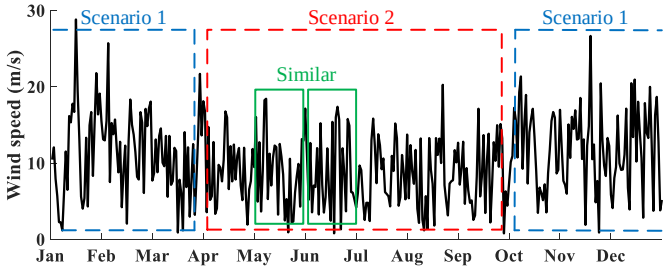

Fig. 3. One-year wind speed data

The change in wind speed throughout the year presents seasonal characteristics. Obviously, the average wind speed and fluctuations in summer are smaller than those in winter. Therefore, all wind speed scenarios can be divided into two typical scenarios. Meanwhile, the wind speed scenarios of different months in the same typical scenarios are similar. For example, the wind speed scenario in May is like that in June. In addition, one-month wind speed scenarios are more abundant than one-day scenarios, and the timing characteristics are more significant. Thus, the operation period is set to one month. Furthermore, given that January is the strongest wind month and stresses the network to the most extent, January is selected as the reference scenario. Hourly wind power production (relative to capacity) and demand (relative to peak demand) are based on those provided in [8], as shown in Fig. 4. Five typical scenarios are generated to describe the uncertainty of DWG by *K*-means, as shown in Appendix A.

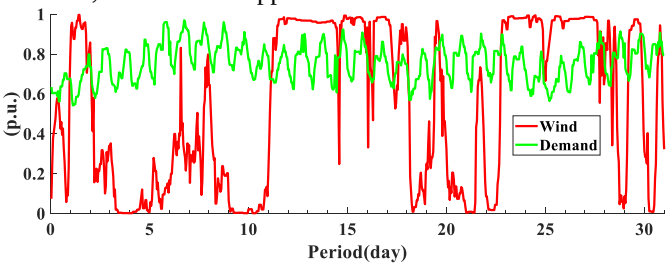

Fig. 4. Hourly wind power generation (reference scenario, relative to capacity) and demand (relative to peak demand) [8].

2) Test system and the data

The proposed method is applied to a sample radial section of the Irish 38-kV distribution network [26], as shown in Fig. 5. The line impedance is given in Table I. Each line between the DWG and the bus is assumed to be short enough for its impedance to be ignored. The voltage limits of each bus are given in Table II. The peak demand data are provided in Table III. The minimum power factor of each DWG is set to 0.9. The model is simulated on Matlab2018a, and the lower-level is solved with CPLEX.

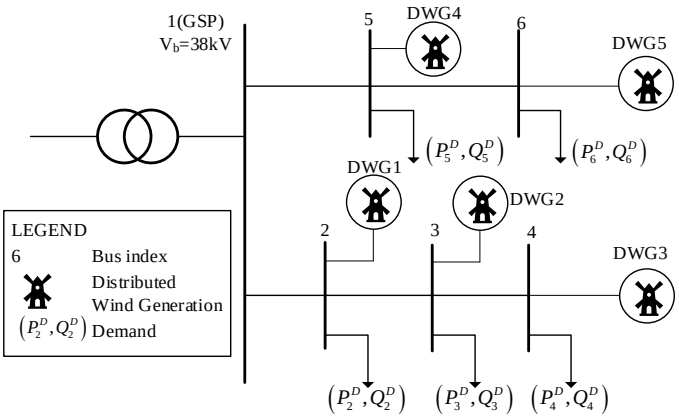

Fig. 5. Test system [26].

TABLE I
TEST SYSTEM IMPEDANCE

Line		Resistance(Ω)	Reactance (Ω)
From	To		
1	2	1.19	1.176
2	3	2.98	3.14
3	4	9.32	9.80
1	5	3.36	3.53
5	6	10.44	10.98

TABLE II
VOLTAGE LIMITS

Bus	1	2	3	4	5	6
Max	1.0789	1.0718	1.0526	1.0487	1.0508	1.0474
Min	1.0789	1.0508	0.9561	0.9237	1.0097	1.0039

TABLE III
PEAK DEMAND OF TEST SYSTEM

Bus	2	3	4	5	6
Active power (MW)	0.68	4.12	4.95	0.68	4.67
Reactive power (MVar)	0.17	1.35	1.44	4.87	1.53
Index of DWG	1	2	3	4	5

B. *Bayesian integrated capacity evaluation trace*

The evaluation trace of the Bayesian integrated method with the average CR not exceeding 10% is obtained, as shown in Fig. 6. To observe the iteration trace of the method, 50 iterations were performed.

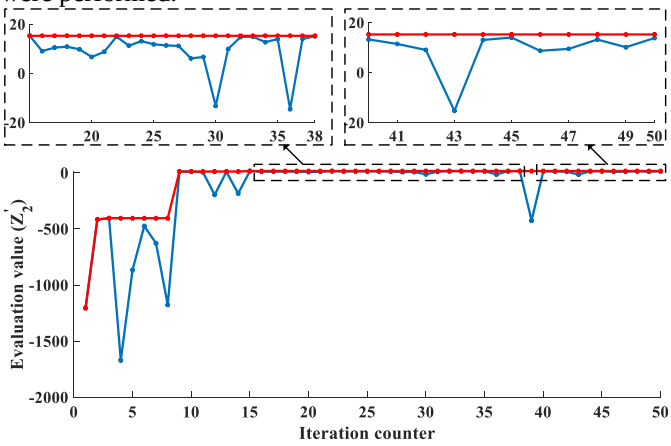

Fig. 6. Evaluation trace with the average CR not exceeding 10% based on Bayesian integrated method.

The overall trend shows that when the iterations counter is less than 14, the evaluation value ($Z_2^{'}$ in objective function (31)) changes significantly. Although the correlation between evaluation and operation can be described by the PSM based on the evaluated sample set, after the steady improvement of the evaluation values in the 8th to 11th iterations, the values in the 12th to 15th iterations show significant volatility due to the very small number of iterations (8 iterations compared to 5-dimensions of the evaluation point), which results in the insufficient accuracy of the PSM-based fitting.

Compared with the first 14 iterations, the evaluation values in the following iterations are similar, and the average CR obtained from the operation-level can meet the CR limit, indicating that after 14 iterations, the coupling correlation of operation and evaluation can be described by the PSM. The entire trace shows that when the obtained average CR cannot meet the CR limit, the value will become extremely small, demonstrating that the penalty term ($M \bullet \delta$) in Formula (31) can make an evaluation point gradually meet the CR limit and an optimal point fall to the feasible region.

C. *Diverse evaluation results*

Several reasonable connectable capacity evaluation results with an average CR limit of 10% are shown in Table IV. The average CR in each result is approximately 10%, and the total capacity in each result is not very different. The CR is a relative value, as shown in Formula (14). Thus, for the first result, although the CR of the 2nd DWG seems significantly different from that of the 4th DWG, the absolute average generation curtailment between the two DWGs is not very different (the absolute curtailment of the 2nd DWG is 0.808 MW, while that of the 4th DWG is 0.717MW). In other results, a large CR does not mean a large absolute curtailment.

TABLE IV
DIVERSE RESULTS

DWG index	1	2	3	4	5	Overall
Capacity (MW)	1.07	6.12	1.41	1.80	4.72	15.11
CR (%)	0.27	13.21	0	39.83	0	10.02
Capacity (MW)	0.83	4.46	0.72	1.68	7.94	15.63
CR (%)	11.95	10.26	0	43.79	4.91	10.77
Capacity (MW)	1.77	2.16	2.30	1.75	7.28	15.26
CR (%)	46.91	11.11	0.60	73.27	3.80	9.91
Capacity (MW)	1.43	4.16	1.21	6.54	2.01	15.34
CR (%)	0	0	0.03	25.68	0	10.95
Capacity (MW)	0.96	2.17	1.09	8.61	1.96	14.80
CR (%)	0	0	0	17.36	0	10.10

Theoretically, when the output of DWG great than load demand, the larger the capacity, the larger the power that needs to be curtailed, and the CR will increase. However, the total capacity of the 1st result is 15.11 MW greater than the 14.80 MW of the 5th result, but the average CR of the 1st result is 10.02% less than the 10.10% of the 5th result. The 3rd and 5th results have a similar relation, indicating that excessive generation is curtailed not only because it exceeds the demand but also the distribution network capacity.

D. *Multi-period DWG operation*

On the basis of the evaluated DWG capacity (1.77, 2.16, 2.30, 1.75, and 7.28 MW), the multi-period DWG operation model is solved with Scenario 1, and the CR of each DWG in each period is obtained, as shown in Fig. 7. In each period with a large CR, the CRs of the 1st and 4th DWGs are higher than those of the 2nd, 3rd, and 5th DWGs. This situation is caused by the different installation locations of various DWGs, such as, the demand for the 1st and 4th DWG installation points (buses 2 and 5) is much lower than that of the others. During periods with fast wind speed, most of the DWG output on the bus with a large demand can be consumed by the local demand. If the

DWG on the bus with small demand does not curtail power or its CR is extremely small, the large volumes of generation will be exported to the distribution network, resulting in difficult to meet power flow and network security constraints. Therefore, the power curtailment of the DWG with a small demand is greater than that of the DWG with a large demand, verifying that the proposed "connectable capacity" depends not only on the policy constraint but also on security constraints.

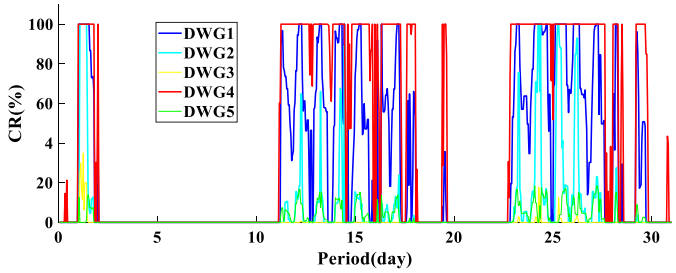

Fig. 7. CRs of each DWG in each period.

The total DWG output, network losses, and demand in each period are shown in Fig. 8. To facilitate analysis, all data is converted into per-unit values. Generally, the integration of DWG can reduce network losses, but the general trend indicates that when the DWG output is greater, the network loss is also relatively largeer. This phenomenon may be attributed to the goal of the operation, which is to accommodate the output of DWG rather than minimize network losses. When the wind speed is fast, the power flow direction in the distribution network may change from the end bus to the head bus. According to the system parameters, the resistance of line 56 is 10.44 ohms, which is the largest resistance among all lines, and the capacity of the DWG installed at bus 6 is 7.28 MW, which far exceeds the peak demand in bus 6, indicating that some periods will have a large amount of power from bus 6 to bus 5, causing huge network losses.

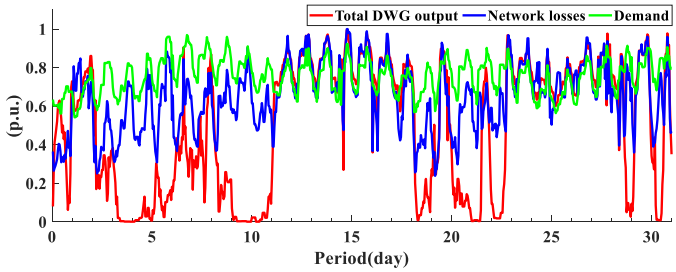

Fig. 8. Total DWG output (relative to 14MW), network loss (relative to 0.45MW) and demand (relative to peak demand) in each period.

E. Parameter analysis on CR

1) Performance of the method under different CR limits

To analyze the performance of the Bayesian integrated algorithm under different CR limits, the model is solved with a CR limit = 15%, 20%, 30%, as shown in Fig. 9. Although the Bayesian integrated algorithm converges to the optimal value after 9 iterations when the CR limit = 10%, 35 iterations are still performed here for stability. Furthermore, the selection of the initial evaluation points (the first four iterations) is the same to compare the difference in objective function values under different CR limits. For the first four iterations, the objective function value increases with the CR limit because as the CR limit increases, the penalty term in Equation (31) decreases. When the CR limit is 10%, 15%, 20%, and 30%, the curves start to flatten from the 15th, 13th, 12th, and 9th iterations, respectively, indicating that the larger the CR limit is, the faster the algorithm converges. This phenomenon occurs because models with different CR limits are solved within the same DWG capacity search range (each DWG capacity search range is [0MW, 20MW]); the larger CR limit is, the larger the feasible range is, and the correlation between operation and evaluation can be better described by the Bayesian integrated method.

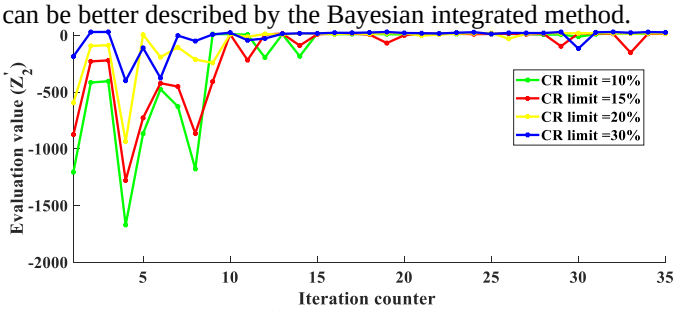

Fig. 9. Evaluation trace with different CR limits.

2) Constraints without renewable accommodation policy

Constraints without a renewable accommodation policy are implemented (CR limit = 0), and the model is solved with the first scenario. In this case, the entire output of the DWG is injected into the network, and the connectable capacity of the DWG is only constrained by the distribution network security, which is the defined basic hosting capacity. Thus, the capacity (8.916 MW) is significantly less than that of CR = 10%. The total DWG outputs and network losses under different CR limits are compared, as shown in Fig. 10. After power curtailment is implemented, renewable accommodation is greatly promoted, and network losses are significantly reduced.

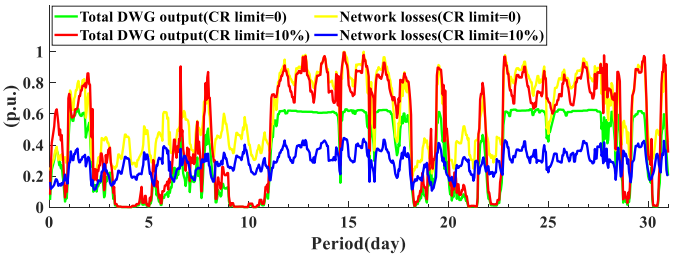

Fig. 10. Total DWG output (relative to 14MW) and network loss (relative to 0.7MW) in each period with different CR limits.

3) Full sensitivity analysis

The sensitivity curve that represents the connectable capacity of the DWG under a wide range of CR limits is shown in Fig. 11. On the basis of the observed points, inverse regression, linear regression, and quadratic regression were used to fit the curve.

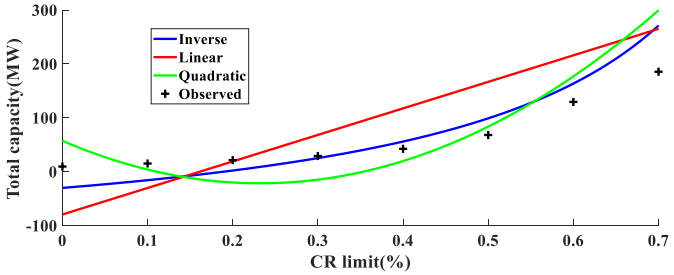

Fig. 11. Sensitivity curve of connectable capacity fitted with different regression models.

The fitting function and effect are shown in Table V. According to the definition of CR in (14), the average CR is inverse the total capacity. Therefore, the fitting effect of inverse proportional regression is better than those of linear regression

and quadratic regression. The sensitivity curve indicates that the connectable capacity of DWG increase as the limit on allowed curtailments is relaxed, and they show an inverse correlation.

TABLE V
FITTING FUNCTION AND FITTING PERFORMANCE WITH DIFFERENT MODELS

Regression model	Function	R-squared
Inverse proportional	$y=-\frac{129.069}{x-1}-159.068$	0.929
Linear	$y=492.177x-79.305$	0.594
Quadratic	$y=1465.276x^2-680.044x+57.454$	0.865

Note: x present CR Limit. y present evaluated connectable capacity. R-squared is the goodness of fit, indicates how well the regression curve fits the observations, the larger the R-squared, the better the fitting effect.

F. Accuracy and efficiency of Bayesian integrated method

To verify the effectiveness and accuracy of the proposed method. The initial NLP model is compared with the LP model. Moreover, the Bayesian integration method is compared with the GA, PSO, and the simulated annealing algorithm (SAA). All methods are implemented on the reference scenario, and the results are shown in Table VI.

TABLE VI
EFFECTIVENESS AND ACCURACY OF DIFFERENT MODELS/ ALGORITHMS

Model/ Algorithm	Total Capacity (MW)	Solving time
NLP (lower-level)	Infeasible	Infeasible
LP (lower-level)	Infeasible	7.9min
GA	14.95	14.03h
PSO	14.95	12.62h
SAA	14.50	9.22h
Bayesian integrated method	15.26	71min

The initial model is not solvable using CPLEX. After gradually linearizing, the lower-level is converted into a LP model, which can be solved by CPLEX in 7.9 min. Fig. 6 shows that the proposed regression-based method converged to the optimal value after nine iterations, which takes 4278 s (i.e., approximately 71 min) (on a PC, Intel Core5 2.8GHz, 8GB of RAM). If GA, SAA, or PSO is used with the same lower-level LP model (the population size for GA and PSO is set to 20), the time consumed is at least 9.22 h. This result shows that Bayesian optimization with regression as the essence can effectively simulate the correlation between evaluation and operation to increase the convergence speed. Meanwhile, the evaluation results of the GA, PSO, the SAA, and the Bayesian integrated method are very close, verifying the accuracy of the proposed method. Overall, the computational cost can be significantly reduce using proposed method, and this advantage will become more apparent as the system scale or operation period increases.

G. Larger system evaluation

To further verify the efficiency and benefit of the method, 33 [27], 69 [28], 85 [29], and 141-bus systems [30] are tested on the reference scenario via the proposed method. Since that the purpose of this part is to verify the efficiency, 20 operation periods are sampled from 744 periods for evaluation based on Monte Carlo simulation. The number of iterations and the convergence time are summarized in Table VII. The results indicate that the maximum number of lower-level evaluations (i.e., iterations) in all systems is only 12. This result again verifies that the regression-based Bayesian integrated optimization method would reduce the number of evaluations by effectively simulating the coupling correlation between evaluation and operation. Furthermore, given the computational time of each evaluation depends on the system scale, the overall convergence time of the 141-bus system is much longer than that of the 6-bus system. In addition, as the scale of the system increases, the difference between the solving time of the Bayesian integrated optimization method and that of the traditional method continuously increases, again showing that the efficiency advantage of the proposed method will become more apparent as the system scale increases.

The "number of iterations to converge" does not depend on the scale of the system. The proposed method can be regarded as a regression-based search algorithm. The iteration numbers mainly depend on the correlation between the search range and the feasible region. For all systems, except the 6-bus system, the DWG capacity search range is set to [0 MW, 15 MW]. Obviously, under the same search range, the smaller the feasible region is, the more iterations are needed. According to the processing of soft constraints by Formulas (31) and (32), when the average CR exceeds the CR limit, the objective function will be penalized, and its value will decrease. Describing the correlation between operation and evaluation is unfavorable for the PSM. In this case, more evaluation points must be obtained through iteration. Thus, under the same search range, the greater the optimal value (connectable capacity) is, the larger the range of DWG capacity that meets Constraint (16) and the fewer the iterations needed are.

TABLE VII
EFFECTIVENESS OF DIFFERENT SYSTEM

Test system	Connectable Capacity (MW)	Proposed method		Convergence time for PSO (s)
		Number of iterations to converge	Convergence time (s)	
6-bus	15.26	9	81.48	802.8
33-bus	3.79	9	493.33	5079.56
69-bus	3.25	12	1600.02	12887.11
85-bus	2.55	12	1919.02	15262.08
141-bus	12.05	8	2021.95	24389.75

V. CONCLUSION

A Bayesian integrated optimization method is designed to solve the proposed bi-level coordinative model for evaluating the connectable capacity of DWG in distribution networks when the policy of renewable accommodation is in place. The following conclusions are drawn.

1) The coordinative optimization of evaluation and operation can provide advice for the variable capacity expansion of the distribution network through maximizing the capacity of DWG and improving energy utilization under a specific renewable policy.
2) Efficiency analysis shows Bayesian optimization with regression as the essence that enables the proposed method to find the optimum after a few iterations. The efficiency of the method is at least 7–11 times better than that of traditional intelligent optimization algorithms. This advantage will become more apparent as the system scale increases.
3) The integration of CR and DWG control can significantly promote renewable accommodation and reduce network losses. Parameter analysis shows that the proposed method can efficiently evaluate connectable capacity under different renewable accommodation policies, and the CR

limit is inverse the connectable capacity.

In the next study, accommodation methods such as energy storage and demand response could be incorporated into the model to further increase the DWG connectable capacity. Considering the high efficiency requirements of rolling planning, the proposed method is worthy of further research.

APPENDIX A

The algorithm flow of *K*-means is as follows:

Taking N_m samples $\{S^1, S^2, \cdots, S^M\}$ in the sampling period as the initial sample set, *K*-means clustering method is used to cluster the initial samples into *K* scenarios based on similarity. The calculation steps are as follows:

1) Randomly select *K* samples from N_m samples as cluster centers denoted as $\mu_1, \mu_2, \cdots, \mu_K$, and denote the *k*-th cluster as C_k.

2) For sample S^m, calculate the distance between it and each cluster center based on the Euclidean distance shown in the (A1), select the cluster center with the smallest distance, and classify it into the class where the cluster center is located.

$$d(S^m, \mu) = \sqrt{\left(s_1^m - \mu_1\right)^2 + \cdots + \left(s_T^m - \mu_T\right)^2} = \sqrt{\sum_{t=1}^{T}\left(s_t^m - \mu_t\right)^2} \tag{A1}$$

where *T* represents the dimension of the sample, which in this paper represents the number of operation periods.

3) After all samples are classified, calculate the number of samples in each cluster and take the average of all the samples classified into the cluster to obtain new cluster centers:

$$\mu_k = \frac{1}{N_k}\sum_{S^m \in C_k} S^m \tag{A2}$$

4) Calculate the convergence criterion, where τ is the number of iterations:

$$E(\tau) = \sum_{k=1}^{K}\sum_{S^m \in C_k} d\left(S^m, \mu_i\right) \tag{A3}$$

5) Judge whether the iteration has converged. If $\left|E(\tau) - E(\tau - 1)\right| \le \xi$, stop the iteration and output the iteration result. Otherwise, go to 2).

6) After convergence, calculate the probability of each scenario:

$$\pi_k = \frac{N_k}{N_m} \tag{A4}$$

The generated scenarios, which are expressed by the difference with the reference scenario, are shown as Fig. A1.

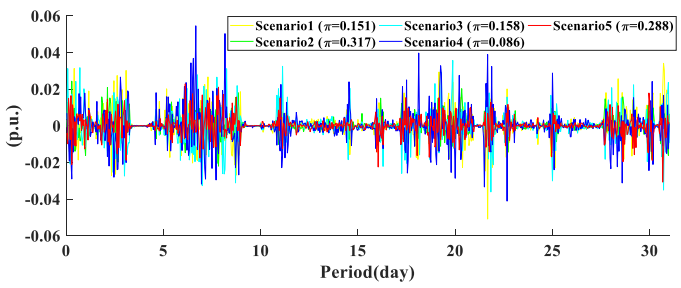

Fig. A1 Typical scenarios of wind power generation (compared with the reference scenario).

REFERENCES

[1] Global Wind Energy Council, "Global wind report", Global Wind Energy Council, Brussels, Belgium, April, 2019.

[2] R.A. Walling, R. Saint, R.C. Dugan, J. Burke, and L.A. Kojovic, "Summary of distributed resources impact on power delivery systems," *IEEE Trans. Power Del*., vol.23, no.3, pp.1636-1644, Jul.2008.

[3] R. A. F. Currie, G. W. Ault, R. W. Fordyce, D. F. MacLeman, M. Smith and J. R. McDonald, "Actively managing wind farm power output," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1523-1524, Aug. 2008.

[4] Y. Xiang, W. Han, J. Zhang, J. Liu and Y. Liu, " Optimal sizing of energy storage system in active distribution networks using Fourier–Legendre series based state of energy function." *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2313-2315, March 2018.

[5] Y. Xiang, Lili Zhou, *et al.*, " Coordinated DG-tie planning in distribution networks based on temporal scenarios," *Energy*, vol. 159, no. 15, pp. 774-785, September 2018.

[6] Y. Xiang, *et al.*, " An economic criterion for distributed renewable generation planning," *Electric Power Components and Systems*, vol. 45, no. 12, pp. 1298-1304, Nov. 2017.

[7] Y. Xiang, H. Cai, et al., " Cost-benefit analysis of integrated energy system planning considering demand response," *Energy*, vol. 192, no. 1, pp. 116632, Feb. 2020.

[8] 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, Feb. 2010.

[9] E. Mulenga, et al. " A review of hosting capacity quantification methods for photovoltaics in low-voltage distribution grids," *International Journal of Electrical Power & Energy Systems*, vol. 115, pp. 105445.1-105445.13, -Feb. 2020.

[10]X. Chen, W. Wu, B. Zhang and C. Lin, "Data-Driven DG Capacity Assessment Method for Active Distribution Networks," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3946-3957, Sept. 2017.

[11]S. Wang, Y. Dong, L. Wu and B. Yan, "Interval Overvoltage Risk Based PV Hosting Capacity Evaluation Considering PV and Load Uncertainties," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2709-2721, May 2020

[12]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, July 2017

[13]J. Zheng et al., " Integrated heat and power dispatch truly utilizing thermal inertia of district heating network for wind power integration," Applied Energy, vol. 211, no. 1, pp. 865-874, Feb. 2018.

[14] Z. Zhao, L. Wu, S. Zhang and X. Li, "An enhanced network-constrained UC model for leveraging system operation cost and financial profitability of incentive-based DR loads," IEEE Transactions on Smart Grid, vol. 9, no. 1, pp. 3-13, Jan. 2018.

[15]S. Ganguly and D. Samajpati, "Distributed generation allocation on radial distribution networks under uncertainties of load and generation using genetic algorithm," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 688-697, July 2015.

[16]B. R. Pereira, G. R. M. da Costa, J. Contreras and J. R. S. Mantovani, "Optimal distributed generation and reactive power allocation in electrical distribution systems," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 3, pp. 975-984, July 2016.

[17]A. Ameli, S. Bahrami, F. Khazaeli and M. Haghifam, "A multiobjective particle swarm optimization for sizing and placement of DGs from DG owner's and distribution company's viewpoints," *IEEE Transactions on Power Delivery*, vol. 29, no. 4, pp. 1831-1840, Aug. 2014.

[18]S. Wang, S. Chen, L. Ge and L. Wu, "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, July 2016.

[19]X. Xu, J. Li, et al., "Enhancing photovoltaic hosting capacity—A stochastic approach to optimal planning of static var compensator devices in distribution networks," Applied Energy, vol. 238, no. 15, pp.952-962, 2019.

[20]Q. Guo *et al*. "A model for multi-objective coordination optimization of voltage and reactive power in distribution networks based on mixed integer second-order cone programming." *Proceedings of The Chinese Society for Electrical Engineering*, vol. 38, no. 5, pp. 1385-1396, 2010.

[21]A. Zare, C. Y. Chung, J. Zhan and S. O. Faried, "A distributionally robust chance-constrained milp model for multistage distribution system planning with uncertain renewables and loads," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5248-5262, Sept. 2018.

[22]J. Xu, X. Yi, Y. Sun, T. Lan and H. Sun, "Stochastic Optimal Scheduling Based on Scenario Analysis for Wind Farms," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1548-1559, Oct. 2017.

[23]A. Ben-Tal and A. Nemirovski, "On polyhedral approximations of the second-order cone," *Math. Ope: Res.*, vol.26, pp.193-205, 2001.

[24]W. Lyu *et al*., "An efficient Bayesian optimization approach for automated optimization of analog circuits," *IEEE Transactions on Circuits and Systems I*, vol. 65, no. 6, pp. 1954-1967, June 2018.

[25]T. Boehme,J. Taylor,A.R. Wallace, and J.W. Bialek, *Matching Renewable Electricity Generation With Demand*. Edinburgh,U.K.: Scottish Executive, Feb.2006.

[26]A. Keane, L. F. Ochoa, E. Vittal, C. J. Dent and G. P. Harrison, "Enhanced utilization of voltage control resources with distributed generation," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 252-260, Feb. 2011.

[27]M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401-1407, April 1989. A. Keane, L. F.

[28]D. Das, "Optimal placement of capacitors in radial distribution system using a Fuzzy-GA method," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6-7, pp. 361-367, Sep. 2008.

[29]D. Das, D. P. Kothari, A. Kalam, " Simple and efficient method for load flow solution of radial distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 17, no. 5, pp. 335-346, Oct. 1995.

[30]H. M. Khodr, F. G. Olsina, P. M. De Oliveira-De Jesus, and J. M. Yusta, "Maximum savings approach for location and sizing of capacitors in distribution systems," *Electric Power Systems Research*, vol. 78, no. 7, pp. 1192-1203, July 2008.