Optimal sizing and siting of energy storage systems based on power grid vulnerability analysis: a trilevel optimization model

Zhenghui Zhao[a], Yingying Shang[a],Buyang Qi[b], Yang Wang[a,*],Qian Zhang[c]

[a]*School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China.*
[b]*China Electric Power Research Institute, Beijing 100192, China.*
[c]*School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China.*
**Corresponding author. E-mail addresses:*1000002156@ujs.edu.cn(Y. Wang).

Abstract

The integration of high proportions of renewable energy reduces the reliability and flexibility of power systems. Coordinating the sizing and siting of battery energy storage systems (BESS) is crucial for mitigating grid vulnerability. To determine the optimal capacity and location of BESS in high-penetration renewable energy systems, this paper proposes a trilevel optimization model for BESS sizing and siting. The upper-level is a pre-optimization layer that uses global vulnerability index to filter out potential BESS siting schemes that cannot mitigate system vulnerability, thereby enhancing the efficiency of capacity allocation and siting. The middle-level employs an improved particle swarm optimization algorithm to determine the optimal capacity and power configuration of BESS, aiming to maximize its equivalent annual revenue. The lower-level uses an improved butterfly algorithm to maximize the expected revenue of daily operation scheduling. The improved algorithms not only enhance computational efficiency but also significantly improve accuracy. The proposed model is validated through case studies based on the extended IEEE 39-bus system. The results demonstrate the model's effectiveness and superiority, showing that it can identify the optimal BESS sizing and siting scheme while mitigating grid vulnerability and enhancing economic benefits.

Keywords: Trilevel optimization model, A pre-optimization layer, Sizing and siting optimization, Vulnerability analysis, Economic benefits.

1. Introduction

The escalating energy and environmental crises have prompted the extensive utilization of diverse renewable resources such as wind turbines (WTs) and photovoltaics (PVs) to cater to consumer energy requirements[1]. The integration of a significant amount of renewable energy into the power system leads to a substantial increase in the uncertainty of power flow, thereby exerting a profound impact on voltage stability and frequency control within the power system[2].

Although energy storage systems (ESS) are indispensable flexible resources in power systems, their ability to transfer electrical energy spatially and temporally, along with their bidirectional power support capabilities, play a significant role in enhancing renewable energy absorption, smoothing renewable energy fluctuations, optimizing power flow, improving voltage stability, and addressing challenges like insufficient peak-shaving capacity and emergency response. ESS can effectively enhance the overall flexibility and reliability of the power grid[3]. Excessive implementation of ESS can result in the underutilization of resources and have a negative impact on the overall economic efficiency of the power grid[4]. Simultaneously, prolonged periods of inactivity for battery components lead to a reduction in operational efficacy. Moreover, it becomes challenging to leverage the flexible resource advantages of energy storage when the capacity configuration of ESS is insufficient. The inadequacy in the capacity of ESS not only hampers effective absorption of

excess generation of WTs and PVs but also restricts exploration of additional application scenarios for enhanced profitability[5]. Therefore, conducting a comprehensive analysis of investments and revenues across multiple locations to determine the optimal siting and capacity of ESS that can effectively meet system demands has become a crucial task[6].

The task of optimally sizing and allocating battery energy storage systems (BESS) can vary based on different scenarios. However, at its core, it is always an optimization problem. Thus, significant research efforts have been dedicated to modeling and solving the problem of optimally sizing and placing BESS in power systems. These efforts can generally be summarized into three major categories: 1) two-stage framework [7-11], 2) bi-level optimization [12-16]and 3) data-driven based multi-stage algorithm [17, 18]. In particular, the two-stage framework has been extensively developed to address the optimal configuration of energy storage systems. Similarly, the bi-level optimization approach is well-suited for the optimal sizing and allocation of BESS due to the inherent interdependence between the planning and operation layers. For example, a bi-level multi-objective optimization model proposed in [13] optimally designs and operates renewable energy sources and storage systems. This model combines particle swarm optimization (PSO) with dynamic linear AC-optimal power flow (DLOPF). A bi-level optimization model proposed in [14] efficiently designs the distribution network by incorporating various functionalities of renewable energy and ESS to ensure both economic viability and reliable operation of the network. A bi-level integrated energy systems scheduling strategy in [19] considers the coordination of distributed battery energy storage systems (BESS). The upper layer optimizes the BESS capacity, while the lower layer focuses on the optimal operation of the power system. In contrast to the conventional two-stage framework and bi-level optimization approach, data-driven multi-stage algorithm can significantly alleviate the computational burden of the optimization model due to its ability to handle massive state variables [17]. In the day-ahead electricity market, [18] proposes a data-driven probabilistic characterization of the real-time balancing stage to inform the day-ahead scheduling problem faced by BESS owners. The aforementioned approaches and methods consider all nodes in the power system as potential locations for BESS when optimizing its placement. Although data-driven based multi-stage algorithms can reduce computational burdens when handling large power system networks, their high demand for computational resources and strong reliance on data can affect their accuracy and reliability in scenarios with insufficient or low-quality data[20]. A more prudent approach involves initially narrowing down the potential deployment locations for BESS. Specifically, if the sites for energy storage devices can be substantially reduced prior to implementing corresponding intelligent algorithms, redundant computation time can be minimized from the outset. Integrating a pre-optimization layer to screen and reduce potential installation sites of energy storage systems within conventional bi-level or two-stage optimization models represents a critical step in the optimization process[21].

Furthermore, once the optimization model is determined, selecting different algorithms can significantly impact the computational efficiency of the optimization process[22]. For instance, the bi-level optimization model proposed in [13] integrates PSO with DLOPF to optimize the design and operation of renewable energy sources and storage systems within an expanding electrical grid. This integration aims to enhance system efficiency and reliability amidst rising demand. The bi-level model proposed in [15] is efficiently solved using a nested genetic algorithm (GA) approach across multiple scenarios, providing optimal solutions for sizing and locating BESS within a reasonable time frame. In [23], multi-objective particle swarm optimization (MOPSO) is utilized to simulate and analyze bi-level optimization for sizing an

independent PV-wind-hydropower station with pumped-storage. Most swarm intelligence algorithms, including particle swarm optimization, genetic algorithm, and grey wolf optimizer, effectively tackle the optimization challenges related to the optimal sizing and placement of BESS[24]. The internal characteristics of these algorithms, however, may impose limitations on the optimization performance. For example, the PSO algorithm offers several advantages, including ease of implementation, minimal parameter requirements, and rapid convergence. The PSO algorithm can experience premature convergence, which leads to suboptimal solutions[25]. On the other hand, the grey wolf algorithm demonstrates exceptional adaptability and flexibility, making it suitable for addressing multi-dimensional optimization problems. However, its drawback lies in its high algorithmic complexity, which results in prolonged computational time[26]. Therefore, when dealing with multi-layer models, selecting appropriate algorithms to enhance efficiency and accuracy is crucial.

On the other hand, the increasing integration of renewable energies will significantly influence the optimization of sizing and siting BESS, due to their inherent characteristics such as stochasticity and intermittency. Therefore, much research is concentrated on managing the uncertainty linked to renewable energy sources concerning the sizing and siting of RESS. In [27], an optimal planning and operation architecture is proposed for multi-site renewable energy generators, effectively sharing an energy storage system on the generation side. In [28], a novel methodology based on deep reinforcement learning is proposed for real-time optimal planning of energy storage systems, considering the uncertainty of renewable energy sources. To effectively address the impacts of multiple uncertainties arising from renewable generation, a multi-objective optimization approach incorporating information gap decision theory (IGDT) was employed in [29] to handle uncertainties in both energy demand and renewable energy sources. In [30], an optimization model based on chance-constrained goal programming was proposed to address the uncertainty of renewable energy in the sizing and allocation of BESS. When employing mathematical programming, stochastic optimization (SO) and robust optimization (RO) techniques are widely used to optimize the objective function while accounting for uncertainties associated with renewable generation [31]. These methods effectively manage the uncertainty inherent in renewable energy generation. However, as power systems scale up and include more complex scenarios, addressing uncertainty in renewable energy generation through optimization solutions imposes substantial computational demands[32]. Additionally, employing the multi-scenario technique to characterize the uncertainty of renewable generation output in long-term planning has become the predominant approach in stochastic optimization for power system analysis[33]. However, including a large number of initial scenarios in the optimization solution significantly increases computational complexity. Addressing this computational burden necessitates finding feasible solutions, which is crucial.

As large-scale renewable energy sources are integrated into power systems, their inherent randomness and intermittency contribute to the continual expansion of power system scale and the incorporation of diverse complex scenarios. When optimization schemes account for the stochastic nature of renewable energy output, significant computational demands arise, posing challenges in determining the sizing and siting of BESS. Existing studies typically employ traditional two-stage models or data-driven multi-stage algorithms, yet these approaches often fail to fundamentally reduce the computational burden of the optimization model. Moreover, traditional optimization models lack sufficient algorithmic improvements to effectively enhance efficiency and accuracy. Therefore, developing an

optimization model capable of determining the optimal sizing and siting for BESS is crucial to comprehensively enhance the flexibility and stability of the power system.

To tackle this challenge, this paper introduces a trilevel optimization model. Expanding upon the traditional two-level framework, a pre-optimization layer is incorporated to narrow down potential sites for BESS installation using vulnerability indicators. Subsequently, the model addresses scenarios arising from the stochastic nature of renewable energy output through multi-scenario techniques and a particle swarm optimization algorithm. Finally, an improved butterfly optimization algorithm (IBOA) is applied to solve the trilevel optimization model, significantly reducing computational burdens while improving algorithmic accuracy to determine optimal BESS sizing and siting.

The main contributions of this paper are as follows:

(1) An innovative pre optimization layer is proposed to screen potential installation sites for ESS. This layer integrates vulnerability analysis methods from electrical power systems, incorporating voltage stability assessment and connection loss metrics into a comprehensive average vulnerability index at nodes for the first time. The analytic hierarchy process (AHP) is used to determine the weights of these indicators, thereby eliminating non-optimal BESS capacity and location schemes, ultimately forming a novel trilevel optimization framework.

(2) The integration of cluster analysis technology into the scenario reduction method effectively addresses the challenges posed by the inherent randomness and intermittency of renewable energy, further reducing computational burdens and resolving the uncertainties faced during the optimization of BESS sizing and siting.

(3) An IBOA is introduced by incorporating an enhanced dynamic switching probability and dynamic variance Gaussian mutation strategy. This optimized algorithm is specifically designed to be embedded within the proposed trilevel framework, aiming to address high-dimensional nonlinear model problems in the optimal sizing and siting process of BESS. It reduces the risk of premature convergence to local optima and achieves significant improvements in convergence accuracy.

The remainder of this paper is organized as follows: Section 2 provides an overview of the trilevel optimization model and problem definition. Section 3 details the scenario reduction method using cluster analysis techniques and describes the application of the IBOA in solving the proposed trilevel optimization model. Section 4 validates the correctness and effectiveness of the proposed trilevel optimization model through case study. Finally, Section 5 presents the relevant conclusions.

2. Trilevel BESS sizing and siting optimization model

The proposed trilevel optimization model for BESS sizing and siting, as shown in Fig.1, consists of upper, middle, and lower levels. The upper-level is the pre-optimization layer, which establishes a global vulnerability index based on the distribution balance of power system vulnerabilities to reasonably evaluate the impact of integrated energy storage on the overall vulnerability of the power grid and analyze different energy storage configurations. The global vulnerability index includes the power grid voltage stability evaluation index and the connection loss index. In the upper-level optimization, the global vulnerability index is used as a penalty term. After determining the energy storage scheme, the changes in the global vulnerability index are calculated. If these changes do not mitigate the overall vulnerability of the power grid, the energy storage operation and capacity allocation schemes are directly eliminated. The goal of the upper-

level optimization is to reduce unsuitable BESS sizing and siting schemes, thereby reducing the computational burden of the model. The middle-level of the model primarily determines the capacity and power of the energy storage devices, aiming to maximize the annual profit of energy storage investments while assessing whether the proposed energy storage planning scheme can enhance the overall resilience of the power grid. By optimizing energy storage capacity and power allocation, the goal is to maximize the returns on energy storage investments and ensure that the deployment of the energy storage system can improve the reliability and resilience of the power grid. The lower-level of the model focuses on maximizing economic benefits through diversified applications of large-capacity standalone grid-side energy storage stations while clearly establishing the objective function. The goal for grid-side energy storage investors is to optimize the benefits brought by energy storage deployment, minimizing the construction and operational investment costs over the entire lifecycle of the energy storage system and reducing the impact on the distribution network. Additionally, the model supports operators in declaring additional configuration capacity for transactions in the ancillary services market. Through this trilevel optimization model, this paper proposes an innovative and practical solution for BESS sizing and siting optimization, effectively reducing computational burdens and improving optimization efficiency and accuracy.

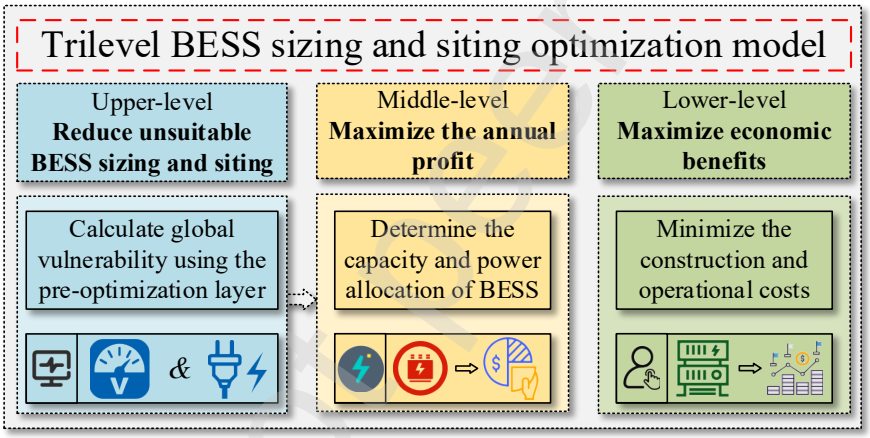


Fig.1 Trilevel BESS optimization model

2.1. Upper-level: Initial BESS location selection based on vulnerability assessment method

The vulnerability assessment indicators primarily capture the inherent structural characteristics of the power grid itself. In this approach, the power network is simplified into a connected network comprising n edges and m nodes, where nodes represent generators, loads, substations, and bus bars. Edges correspond to transmission and distribution lines. All system nodes are categorized into four distinct sets: generator node set, substation node set, load node set, bus node set. The power grid network is represented as an undirected weighted graph comprising of N nodes and K edges. $G=(V,L)$ and $V=\{v_1,\cdots v_N\}$ represents the notes set. $L=\{L_1,\cdots,L_K\}$ denotes the edges set. The degree of a node refers to the number of directly connected nodes. A higher degree implies more connections to other nodes and edges, indicating greater importance within the network[34].The formula for calculating node degree is as follows:

$$D(i)=\sum_{j=1}^{L} y_{ij} \quad y_{ij}\in\{0,1\} \tag{1}$$

Where, $D(i)$ denotes the degree of node i, j denotes the branch, y_{ij} indicates whether branch j is associated

with node i, L denotes the total number of branches.

The number of nodes is utilized to evaluate the extent to which a node functions as a pivotal connection in a distribution network system, i.e., its significance in facilitating information transfer among diverse nodes[35]. A higher median centrality of a node corresponds to an amplified influence within the entire system[36]. Assuming energy transmission occurs via the shortest path, the numerical representation of each node is expressed as follows:

$$B(i) = \sum_{j \neq k \in N} \frac{n_{jk}(i)}{n_{jk}} \tag{2}$$

Where, n_{jk} denotes the number of shortest paths between node j and node k, N denotes the total number of nodes.

The voltage stability evaluation index primarily encompass state and margin indices. The state index assesses system stability based on the current operating conditions, while the margin index comprehensively considers various factors within the system. Specifically, the voltage stability index is predominantly expressed as follows:

$$VI_i^a = \left| U_i(t) - U_{cr} \right| \tag{3}$$

$$VI_i^b = \left| U_i(t) - U_{cr} \right| / \left| U_{i0} - U_{cr} \right| \tag{4}$$

$$VI_i^c = \frac{\partial U_i}{\partial p} \tag{5}$$

Where, $U_i(t)$ indicates the voltage of node i at time t, U_{cr} denotes the threshold voltage of node i, U_{i0} denotes the voltage at node i at the initial moment, p indicates the status parameter of the system.

Considering that both state and structural factors influence nodes, neither type of indicator alone can comprehensively evaluate the distribution network system. Therefore, the AHP is used to determine the weights of various evaluation indicators from different aspects to identify the importance of each index.

The vulnerability of nodes is comprehensively assessed by constructing a comprehensive index, enabling the preliminary screening of energy storage access locations. The number of nodes and connection loss are chosen as evaluation metrics for quantifying the vulnerability of the grid structure, while voltage stability is also selected as a state vulnerability index, taking into account the scenario of distributed power integration. However, conventional approaches for assessing node vulnerability often fail to consider the potential impact resulting from the widespread integration of renewable energy sources. The conventional node interface index, commonly employed in node vulnerability assessment, is widely recognized as the shortest path propagation during power flow calculation, which evidently deviates from the actual operation of the power system. In order to accurately represent the power flow distribution in the power system, while considering the impact of DG output fluctuations on the power grid. To accurately depict the power flow distribution in the system, while considering the impact of fluctuations in distributed generation (DG) output on the grid, we define the improved intermediate index of node i as formulation (6).

$$J(i,t) = \sum_{i,e \in V_1; w \in V_2} \frac{\sqrt{S_{\mathrm{pe}}(t) S_{\mathrm{lw}}(t)}}{S_{\mathrm{b}}} R_{ew}(i) \tag{6}$$

Where, V_1 represents a collection of equivalent power supply nodes, while V_2 denotes a group of equivalent load

nodes. $S_{\mathrm{pe}}(t)$ and $S_{\mathrm{lw}}(t)$ represent the apparent power output by power node e at time t and the apparent power consumed by load node w, respectively. S_b denotes the base capacity of the system. The function $R_{ew}(i)$ determines whether the shortest path between node e and node w includes node i. If it does, the value of $R_{ew}(i)$ is 1; otherwise, it is 0.

The conventional metric for quantifying connection loss is defined as the ratio of load nodes unable to establish a connection with the power node upon disconnection from the network. Clearly, within the framework of closed-loop design and open-loop operation of the distribution network, a higher level of connection loss implies that a greater number of nodes will encounter power outages subsequent to disconnection, thereby amplifying its impact on the power network[37]. The current research lacks a comprehensive consideration of the influence of DG output.

Therefore, this paper introduces the weight of power for the power node, which is a novel addition. Consequently, an improved connection loss index is defined as follows:

$$S(i,t) = 1 - \sum_{i,e\in V_1} \frac{S_{\mathrm{pe}}(t)}{S_{\mathrm{sum}}} E_e(i) / N \tag{7}$$

Where, S_{sum} represents the cumulative apparent power output of all equivalent power supplies, and $E_e(i)$ denotes the judgment function. Specifically, if node i is connected to power node e, its value is 1; otherwise, it is 0. It should be noted that a connection loss of 0 occurs only when the node is connected to all power nodes, while a connection loss of 1 indicates complete disconnection between nodes and power sources, thus still adhering to the defined concept of connection loss.

The influence of DG power fluctuation on the power grid is quantitatively represented in order to assess the anti-risk capability of the grid by analyzing the degree of voltage offset standard value at each node. The higher the vulnerability of a node in the power system, the more susceptible it becomes to voltage disturbances, thereby compromising its robustness. The vulnerability of a node i in the network at time t can be expressed as follows:

$$v(i,t) = \frac{\left|U_{i,t} - U_{i,cr}\right|}{\Delta U_{\mathrm{lim}}} \tag{8}$$

Where, $U_{i,t}$ represents the voltage of node i at time t, while $U_{i,cr}$ denotes the reference rated voltage of node i, ΔU_{lim} signifies the maximum allowable voltage deviation for a given node.

Using appropriate weights and weighted sums of indicators can overcome the limitations of single-indicator evaluations, allowing for a more accurate identification of vulnerable nodes in the power system. The comprehensive vulnerability indicator of node i at time t is defined as follows:

$$CV(i,t) = \omega_1 J(i,t) + \omega_2 S(i,t) + \omega_3 v(i,t) \tag{9}$$

Given the subjective nature of assigning weights to indicators by decision-makers, this study employs the analytic hierarchy process for determining indicator weights. Hence, to enhance the decision-making process, we define the average vulnerability index of node i at time t during a typical day in a natural scenario as follows:

$$CV_{av}(i) = \sum_{t=1}^{T} CV(i,t) / T \tag{10}$$

Considering the vulnerability of a single node only reflects local changes in the grid, solely relying on the vulnerability index of individual nodes cannot comprehensively evaluate the impact of energy storage integration on the overall resilience of the grid. To effectively use this vulnerability indicator for optimal configuration decisions, it is necessary to assess the system's overall vulnerability from a holistic perspective. According to equation (10), the comprehensive vulnerability index of the power grid is defined as follows:

$$CV_g = \sum_{i=1}^{N} CV_{av}(i) / N \tag{11}$$

On the other hand, the uniform distribution of node vulnerability in the power grid also impacts the overall vulnerability, and voltage collapse triggered by any weak node can initiate a cascading failure of other nodes. Therefore, it is imperative to consider the equilibrium level of node vulnerability distribution. Only when the voltage deviation of each node in the network falls within a reasonable range can an accurate assessment be made regarding overall voltage stability.

The parameter of Gini coefficient serves as a classical indicator for quantifying uneven distribution, effectively capturing the extent of unevenness[38]. Therefore, based on the concept of quantifying parameter imbalance using the Gini coefficient, this paper introduces the novel notion of equilibrium degree in vulnerability distribution for power systems. The mathematical expression is presented as follows：

$$E = 1 - \frac{1}{N}\left[2\sum_{i=1}^{N-1} v_{av}(i) + 1\right] \tag{12}$$

Where $v_{av}(i)$ represents the average voltage stability vulnerability index of node i, a smaller value of E indicates reduced voltage fluctuation caused by energy storage access and a more balanced distribution of voltage, thereby enhancing power supply security in the grid.

By establishing a power system global vulnerability index based on the vulnerability distribution equilibrium E, the overall vulnerability of the grid can be reasonably evaluated. The global vulnerability index can be expressed as follows:

$$H = \zeta_g CV_g + \zeta_e E \tag{13}$$

The weights ζ_g and ζ_e represent the equilibrium degree of vulnerability distribution in the power system and the overall comprehensive vulnerability index, respectively, when calculating the global vulnerability index. Their values range from 0 to 1, with a sum equal to 1.

The network model is established by inputting the topology information, enabling the selection of a representative daily scenario for power flow calculation in each time section based on DG output.

Based on the results of power flow calculations, an average comprehensive vulnerability index was computed for each point and nodes were ranked in descending order according to their vulnerability indices. The node with the highest vulnerability index was selected as the primary location for energy storage placement, thus completing preliminary site selection.

2.2. Middle-level: sizing and sitting

The objective function of the second-level optimization model is expressed as:

$$F_{up} = \max\left[D_y^{ess} \bullet F_{lw}(E_{ess}, P_{ess}) - C_{inv} - C_{op} - pun_{CV} \right] \tag{14}$$

The fitness optimal result of the lower layer model, denoted as $F_{lw}(E_{ess}, P_{ess})$ / $F_{lw}(\bullet)$, represents the maximum daily operating revenue of storage. D_y^{ess} refers to the annual operating days of storage. C_{inv} represents various costs associated with energy storage, such as building cost and ontology-related expenses.

$$C_{inv} = \eta \bullet (C_E E_{ess} + C_P P_{ess}) \tag{15}$$

$$\eta = (1 + \lambda_r / 1 + d_r)^{T_{life}} \tag{16}$$

The variable η, which can be calculated via formulation (16), represents the discount factor that annualizes the investment cost of the energy storage device throughout its entire life cycle. The variable λ_r represents the inflation rate, d_r represents the discount rate, T_{life} denotes the service life of the energy storage battery, C_E stands for the unit energy cost of the energy storage, and C_P signifies the unit power cost of the energy storage.

The variable C_{op} represents the yearly expenditure for operating and maintaining energy storage systems, which can be calculated by formulation (17). And the variable C_w represents the yearly expenditure for operating and maintaining energy storage systems.

$$C_{op} = C_w P_{ess} \tag{17}$$

The vulnerability penalty term pun_{cv} represents the comprehensive measure of power grid vulnerability and is employed to exclude schemes that fail to meet the overall vulnerability requirements of the distribution network.

$$pun_{CV} = \chi_1 \Delta H \tag{18}$$

The penalty constraint coefficient, denoted as χ_1, and the change of the global vulnerability index before and after the configuration of energy storage, represented by ΔH, are considered in this study. The overall constrains that are considered in middle-level optimization are summarized as follows.

(1) Improved constraints for enhancing the overall resilience of the power grid (19)

$$\begin{cases} H - H_0 > 0, \chi_1 = 999 \\ H - H_0 < 0 \ \ , \chi_1 = 1 \end{cases} \tag{19}$$

The overall vulnerability index of the distribution network before the configuration of energy storage is denoted as H_0. In order to ensure that the configuration scheme is discarded if it fails to meet the requirements for mitigating the overall vulnerability of the power grid, we introduce a penalty coefficient with a sufficiently large positive value.

(2) Capacity and investment constraints (20) -(22)

$$0 \le E_{ess} \le E_{ess}^{\max} \tag{20}$$

$$0 \le P_{ess} \le P_{ess}^{\max} \tag{21}$$

$$0 \le C_{inv} \le C_{inv}^{\max} \tag{22}$$

The variable $E_{ess}^{\max}$ in this equation represents the maximum capacity of the energy storage configuration. The variable $P_{ess}^{\max}$ represents the upper limit of power for the energy storage configuration, while $C_{inv}^{\max}$ represents the maximum investment funds.

2.3. Lower-level: maximizing economic benefits

The objective function of the Lower-level optimization model is expressed as:

$$\min F_{lw} = inc_{m1} + inc_{m2} + inc_{m3} \tag{23}$$

The first component of the objective function inc_{m1} represents the revenue generated from energy arbitrage through an energy storage system, which can be calculated by formulation (24).

$$inc_{m1} = \sum_{k \in K_{en}} \pi_k \sum_{t=1}^{T} \lambda_t^{en} \cdot P_t^{en} \cdot \Delta t \tag{24}$$

The second component encompasses the market income derived from frequency modulation auxiliary services, which comprises revenue from frequency modulation capacity and compensation for frequency modulation mileage. In terms of the latter, equivalent calculations are made based on average call mileage. Therefore, the representation of energy storage's participation in the market for frequency modulation auxiliary services is as follows.

$$inc_{m2} = \sum_{k \in K_{reg}} \phi_k \sum_{t=1}^{T} (f_t^{cap} + f_t^{\text{mil}}) \tag{25}$$

$$f_t^{cap} = \lambda_t^{\text{reg, cap}} \cdot P_t^{\text{reg}} \cdot A_{reg} \cdot \Delta t \tag{26}$$

$$f_t^{\text{mil}} = \lambda_t^{\text{reg, mil}} \cdot P_t^{\text{reg}} \cdot A_{reg} \cdot m \cdot \Delta t \tag{27}$$

The frequency modulation capacity price in time period t is denoted as $\lambda_t^{\text{reg, cap}}$, while the frequency modulation mileage price in time period t is represented by $\lambda_t^{\text{reg, mil}}$. P_t^{reg} signifies the declaration power of energy storage to participate in the frequency modulation assisted market during time period t. Index m denotes the average calling mileage coefficient of energy storage participating in frequency modulation, and A_{reg} describes the average performance index of energy storage for frequency modulation.

The third component pertains to the revenue generated from auxiliary services in peak shaving. This revenue is determined based on the energy storage's capacity for peak shaving during specific periods and the corresponding compensation price.

$$inc_{m3} = \sum_{k \in K_v} \sigma_k \sum_{t=1}^{T_v} \lambda_t^{vb} \cdot P_t^{vf} \cdot \Delta t \tag{28}$$

Where K_v represents the compensation price scenario for peak shaving, P_t^{vf} denotes the charging power of energy storage systems engaged in peak shaving during time period t, and λ_t^{vb} signifies the compensation price for

peak shaving at time t.

The electricity price scenario in the energy market can be represented as K_{en}, π_k represents the probability of an electricity price event, λ_t^{en} denotes the current electricity price at time t in the energy market, and P_t^{en} signifies the charging power during the energy storage period t.

The expression for the charge and discharge model of the energy storage component is as follows.

$$\begin{cases} SOC(t) = SOC(t-1) + \dfrac{P_t^{ch} \cdot \Delta t \cdot \eta_{ch}}{E_{ess}} \\ SOC(t) = SOC(t-1) + \dfrac{P_t^{dis} \cdot \Delta t}{E_{ess} \cdot \eta_{dis}} \end{cases} \tag{29}$$

Where $SOC(t)$ represents the state of charge of the energy storage element at time t, P_t^{ch} and P_t^{dis} represent the total charging and discharging power of the energy storage element at time t, respectively, η_{ch} and η_{dis} represent the charging and discharging efficiency of energy storage components, E_{ess} is the capacity of the energy storage element, and Δt is the charge and discharge interval which is set to one hour in this paper.

The charging state of the energy storage power station must be constrained within specified upper and lower limits to prevent excessive discharge depth from adversely impacting the service life of the energy storage battery. Thus, the charging state of the energy storage can be expressed as:

$$SOC_{\min} \le SOC(t) \le SOC_{\max} \tag{30}$$

$$SOC(1) = SOC(T) \tag{31}$$

In the formula, $SOC_{\min}$ and $SOC_{\max}$ are respectively the upper and lower limits of the charged state of the energy storage battery. $SOC(1)$ and $SOC(T)$ are the beginning and end charging states of the energy storage power station respectively. The setting constraint makes the amount of storage states of charge and discharge during the scheduling period the same as the initial state at the end of the scheduling period, so as to ensure that the energy storage system has sufficient power capacity in the next scheduling period. The bid power of the energy storage at time t must satisfy both upper and lower limits of the configured power for sum and difference, respectively, while also satisfying the power balance constraints. The specific constraint of power balance is listed below:

$$P_{\min}^{ess} \le P_t^{ess} = P_t^{dis} - P_t^{ch} = P_t^{en} + P_t^{reg} + P_t^{fg} \le P_{\max}^{ess} \tag{32}$$

The charging and discharging power of the energy storage station is constrained by its rated power, while the charge/discharge state's mutual exclusion constraint ensures that the energy storage can only operate in a single mode at any given time. Therefore, it becomes necessary to impose certain limitations on its operational modes, which are outlined as follows.

$$P_t^{ch} \cdot P_t^{dis} = 0 \tag{33}$$

In the energy market, energy storage systems can exist in three states: charging, discharging, or standby. Let f_1、f_2 represent the thresholds for charging and discharging respectively in the market where the energy storage system participates. The operational state of the energy storage station can be determined based on the following conditions.

$$f_1, f_2 \in [0,1], \quad f_1 + f_2 \le 1 \tag{34}$$

When the energy storage device engages in multi-market transactions, it should have limited market participation capabilities, specifically by not simultaneously participating in both the frequency modulation auxiliary market and the peak shard auxiliary service market. This limitation ensures the provision of high-quality auxiliary services. The restricted participation state is defined as follows：

$$P_t^{vb} \cdot P_t^{reg} = 0 \tag{35}$$

The energy storage system must reserve sufficient capacity to comply with its declared frequency modulation capacity in response to the AGC instruction for frequency modulation, failure of which would result in economic loss due to misalignment penalties. Additionally, the redundant capacity constraint of the energy storage system is subject to the following limitations.

$$E_t^{ess} \le SOC_{\max} E_{ess} + (P_t^{en} + P_t^{fg} - P_t^{reg}) \cdot \Delta t \cdot \eta_{ch} \tag{36}$$

$$E_t^{ess} - SOC_{\min} E_{ess} \ge (P_t^{en} + P_t^{fg} + P_t^{reg}) \cdot \Delta t / \eta_{dis} \tag{37}$$

Where, E_t^{ess} is the residual storage capacity at time t.

3. Solution methodology

The primary layer algorithm aims to comprehensively assess node vulnerabilities by constructing a comprehensive vulnerability index and initially screening energy storage access locations. The second layer model aims to explore the profitability potential of energy storage system under various initial capacity and power configurations, by pursuing diverse generation schemes.

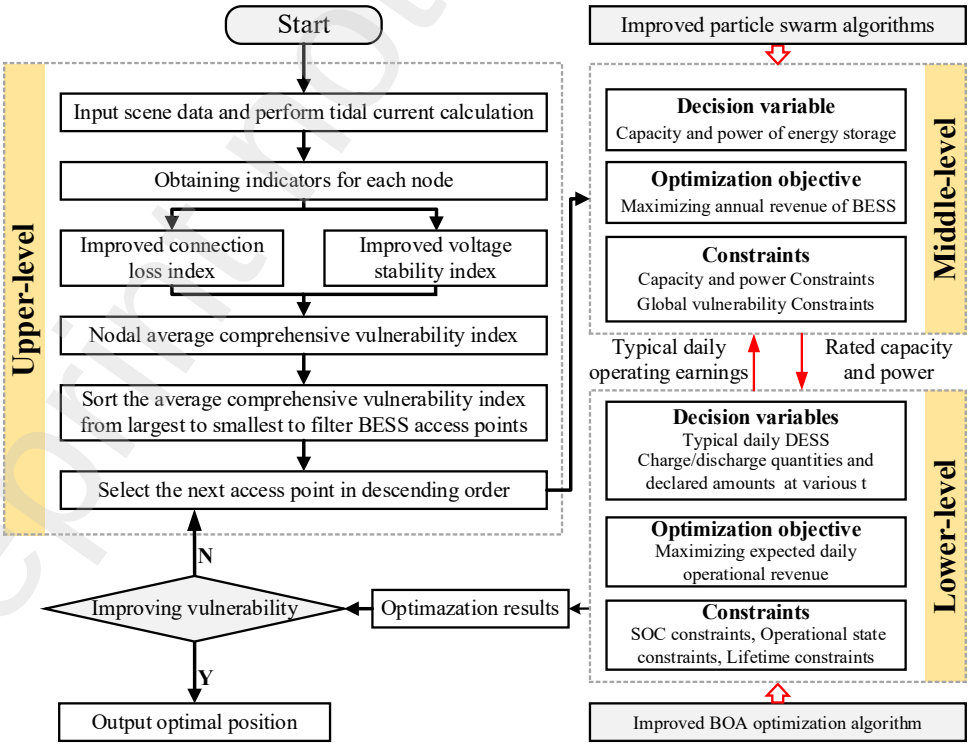


Fig.2 The tri-level BESS sizing and siting optimization flowchart

As shown in Fig.2, to enhance the global optimization ability of the algorithm and meet the solution requirements of the first layer model, we utilize a particle swarm optimization algorithm with fewer parameter settings and fast search speed. Additionally, we improve the generation method of particle initial population using Halton sequence to ensure both randomness and uniform distribution in population diversity. This paper aims to employ the enhanced butterfly optimization algorithm for addressing the multivariate optimization problem in the third layer model, which exhibits intricate nonlinearity.

3.1. DR-CA based approach for uncertainty

Integrating clustering analysis into scenario reduction preserves the distinct data characteristics of various clustered scenarios while maintaining their differences[39]. Therefore, this paper introduces the DR-CA based approach to address the challenges arising from uncertainties associated with renewable energies.

$$\min \sum_{u_i \in J} p_i \min_{u_j \in Q} d\left(u_i, u_j\right) \tag{38}$$

Firstly, traverse the initial scenario set J to identify the scenario u_i with the minimum sum of probabilistic distances to other scenarios and move it to the retention set Q. Repeat this process until the number of scenarios in Q meets the algorithm's requirements. However, unlike the backward reduction method, after each iteration of the fast forward selection method, probabilities of scenes in Q need to be redistributed based on their initial probability proportions so that total probability for each scene in Q always sums up to 1. The given data set is initially divided into subsets in the second step. Through iteration, data is moved between subsets. The distance between data within a subset serves as the similarity evaluation index, where close distances indicate high similarity and far distances indicate low similarity between different subsets. Continuous improvement of intra-subset similarities and reduction of inter-subset similarities lead to the formation of k subsets when the total square error function reaches its minimum value. Assuming n individual data points with p dimensions characteristic attributes represented by X_i features each (as shown in formula (39)). If the initial number of subsets is k, then the clustering center for each subset at the beginning can be represented by formula (40). The error sum of squares function C serves as the evaluation metric for assessing clustering convergence, and its calculation is defined by formula (41).

$$X_i = \left(x_{i1}, x_{i2}, \cdots, x_{ip}\right) \quad i = 1, 2, 3, \cdots, n \tag{39}$$

$$Z_j = \left(z_{j1}, z_{j2}, \cdots, z_{jp}\right) \quad j = 1, 2, 3, \cdots, k \tag{40}$$

$$C = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \left\| X_i - M_j \right\|^2 \tag{41}$$

Where n_j is the number of data individuals in the j subset; M_j is the individual mean of the j subset, which is the geometric centre of the subset. The calculation method for DR-CA based uncertainty analysis is as follows:

$$M_j = \frac{1}{n_j} \sum_{i=1}^{n_j} X_i \tag{42}$$

The method of updating subsets with repeated iteration is employed: in each iteration, the dataset is partitioned into

k subsets based on the previously determined k geometric centers, followed by recalculation of the geometric centers for each newly formed subset. These recalculated centers serve as reference points for subsequent iterations. Through multiple iterations, the obtained reference point gradually approaches the true geometric center, resulting in a decreasing value of C. If there is no change in the geometric center between two consecutive iterations, it indicates convergence of the error sum of squares function and outputs the resulting subsets.

3.2. Dynamic variance Gaussian variation strategy based BOA for optimization

3.2.1 Mathematical formulation for BOA

The conventional BOA has found extensive applications in solving multi-dimensional complex function problems, such as mechanical design, feature extraction, and clustering. It is widely acknowledged for its simplicity in operation, minimal parameter adjustment requirements, and robustness[40]. To enhance the computational efficiency and convergence accuracy while mitigating the risk of algorithmic convergence to local optima. This section investigates the application of an IBOA in addressing the BESS sizing and siting problem and enhances the conventional BOA approach to achieve improved accuracy in solving energy storage planning.

According to the aforementioned concepts, in the BOA, the calculation formula for the scent emitted by individual butterflies can be expressed as follows:

$$f = cI^a \tag{43}$$

In the formulation (43), f indicates the level of olfactory signal emitted by a butterfly, which can be detected by conspecific individuals, c represents the butterfly's perceptual pattern. In addition, I represents the stimulus intensity, which is affected by the objective function. The power index a is dependent on the perception pattern and serves to elucidate the extent of fragrance absorption. It is noted that the value of index c and a both fall within the range of [0,1].

The first step in the BOA is to initiate the generation of the butterfly population $X_i(i = 1, 2, \ldots N)$, where X_i represents the spatial position of the i_{th} butterfly in a butterfly population. The value of Initialized perceptual mode, power index and the switching probability are initiated as c, a and p. The stimulus intensity I of each butterfly was determined according to the objective function $f(X_i)$. The fitness value of each individual within the population should be assessed and arranged in ascending order, thereby identifying the butterfly occupying the optimal position. Subsequently, the optimal position of the record should be determined along with its corresponding fitness value, denoted as the "best fitness", and reveal the olfactory characteristics of butterflies through quantitative analysis. To mitigate the impact of external environmental factors, a random number p is employed to determine whether the butterfly's search strategy should entail local exploration or global exploration. The iteration stage serves as the primary solution stage in BOA, wherein local search and global search constitute the two pivotal steps. In each iteration, all butterflies within the solution space undergo movement to new positions, followed by re-evaluation of their fitness and generation of a updated scent at the updated location, the updated scent can be expressed by formulation (43). The global search is conducted by BOA when the value of p exceeds the predetermined random number. Each individual butterfly converges towards the optimal position g^*, the mentioned action as defined by the following formula (44).

$$x_i^{t+1} = x_i^t + (r^2 \times g^* - x_i^t) \times f_i, p > rand \tag{44}$$

The local search is performed when the value of p is less than the random number, and the butterfly's random movement process can be described by the formulation (45).

$$x_i^{t+1} = x_i^t + (r^2 \times x_j^t - x_k^t) \times f_i, p < rand \tag{45}$$

In the above expression, The vector x_i^t represents the solution vector x_i for the i_{th} butterfly in the t_{th} iteration. The solution g^* represents the optimal outcome among all the potential solutions in the current iteration. The solution vectors x_j^t and x_k^t correspond to butterfly j and butterfly k, respectively, within the same population in the solution space. The scent emitted by butterfly i is indicated by factor f_i.

In order to enhance the applicability of the BOA in solving high-dimensional nonlinear model problems, prevent premature convergence to local optima, and achieve higher accuracy in convergence, we have improved the variation mode and optimization strategy of the algorithm by utilizing an enhanced dynamic switching probability and a Gaussian variation strategy with dynamic variance.

3.2.2 Improved dynamic switching probability

The conventional butterfly optimization algorithm is prone to local optima, thus this paper introduces the dynamic variance gaussian mutation strategy into the conventional BOA for enhancing the convergence rate of the algorithm. The basic butterfly algorithm relies on the switching probability of control parameters to determine the transition between global optimization and local optimization, effectively balancing the interplay between global search mechanisms and local search mechanisms. Hence, the selection of appropriate values for switching probabilities significantly impacts both the algorithm's ability to find global optimal solutions and its convergence speed. If the switching probability is small, the algorithm tends to prioritize local search during the solution exploration process, which may lead to a propensity for local optima and hinder the discovery of global optimal solutions. Conversely, if the switching probability is excessively large, the algorithm emphasizes global search in its solution exploration process. Thus, the switching probability is investigated in order to enhance the algorithm's performance. A random number p is employed to determine whether the butterfly's search strategy should entail local exploration or global exploration Thus, the proposed algorithm incorporates a globally optimal process during the search phase and a locally optimal process during the development phase, followed by integrating these two phases through parameter p. The index p in conventional BOA is commonly set to a fixed value of 0.6, which imposes limitations on the search capability of BOA. However, the dynamic variance Gaussian mutation strategy is employed in this study to enhance the search efficiency of BOA by utilizing a dynamically adjusted conversion probability, as expressed in the formula (46).

$$p = 0.6 - 0.1 \times \frac{Maxiter - t}{Maxiter} \tag{46}$$

Where, t represents the current iteration and $Maxiter$ denotes total iterations.

However, the current setting of the dynamic formula fails to enhance the algorithm's exploration ability during the

early stage of optimization. To address this issue, we propose a new dynamic switching probability p for application in the IBOA, with a value range of [0.7, 0.8]. By adjusting the value of p at the initial search phase of the algorithm, we aim to strengthen its exploratory capabilities. The proposed formulation is expressed as:

$$p = \alpha + \frac{Maxiter - \alpha \tan(t)}{Maxiter} * \beta * \alpha \tan(t) \tag{47}$$

The dynamic switching probability, denoted as , is determined by the initial switching probability value , controlled by factor β. The current iteration number is denoted as t and the maximum iteration number is denoted as $Maxiter$. For the diagram of improvements related to the dynamic switching probability, please refer to the appendix.

3.2.3 Dynamic variance Gaussian mutation strategy

To address the issue of BOA's susceptibility to local optima, an enhanced dynamic variance Gauss variation strategy is employed to enhance population diversity, improve the algorithm's global search capability, and expand the search space. The gaussian probability density function can be mathematically represented as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{48}$$

Where, μ represents the mean value; σ is the standard variance. The function curve depicted in Fig.3 corresponds to the values of x being 0.5, 1, and 2 respectively when u equals zero.

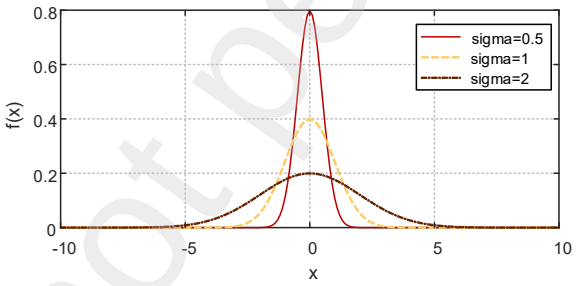


Fig.3 Gaussian distribution curve with different values

4. Case study

4.1. Framework

The simulation analysis was conducted based on the extended IEEE-33 node distribution network test system, as depicted in Fig.4. The common coupling point is connected to the upper power network and node 1 to facilitate power exchange between the main network and the distribution network, ensuring power balance within the distribution network. Two photovoltaic power stations are connected to nodes 21 and 28 respectively. The test system has a reference capacity of 10MVA, a reference voltage value of 12.66kV, a total load power of (3.715+j2.3) MVA, and allows for node voltage per unit values within the range [0.93,1.08] p.u.

The permissible installation nodes for energy storage components range from node 2 to node 33, with the restriction that BESS cannot be installed at the same location. By analyzing the load characteristics based on average and peak levels of typical output scenarios, we assess the region's load profile. The BESS has a maximum capacity of 10MWh, a rated power of 2MW, and a charge/discharge range between [-2,2] MW. The power grid operator adopts time-of-use pricing

when procuring electricity from the upper power grid and selling it to users.

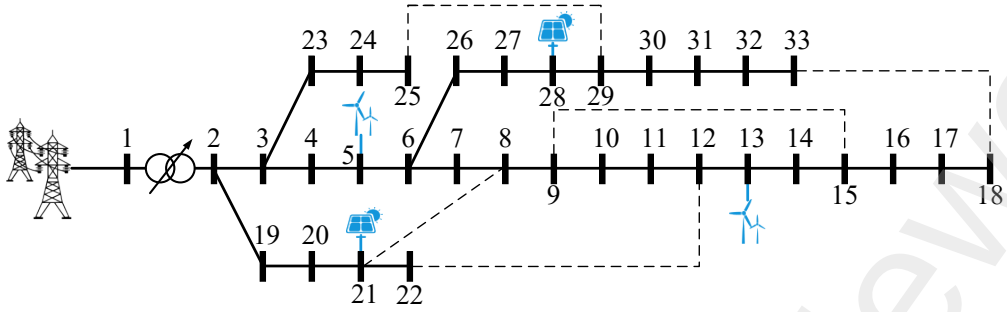


Fig.4 Extended IEEE-33 node system diagram

The price sold to distribution network users is set at 1.3 times higher than the purchase price. Table 1 presents prices when peak and off-peak periods are equally divided. We benchmarked DG electricity prices as well, considering that purchasing electricity from DG can reduce reliance on the superior grid while also benefiting from carbon price trading subsidies. For detailed information on BESS parameters and related economic indicators for simulation, please refer to the appendix.

Table 1

TOU electricity price

	Time	Electricity Price / CNY • KWh^{-1}
Peak	8:00-11:00, 13:00-15:00, 18:00-21:00	1.17
Off-Peak	6:00-8:00, 11:00-13:00, 15:00-18:00, 21:00-22:00	0.72
Normal	22:00- Next Day 6:00	0.58

4.2. Analysis of grid reduction improvement results

The time interval characterized by high load levels and significant peak-valley differences is selected for analysis. During this period, certain transmission lines exhibit overloading issues with a small safety margin, resulting in excessive overall node voltage deviation and terminal node voltages exceeding the lower limit. Consequently, the grid's overall resilience against disturbances decreases. An assessment of each node's average comprehensive vulnerability situation is conducted to obtain quantitative results regarding both structural vulnerability and state vulnerability. For normalized values of the vulnerability assessment indices for each node, please refer to the appendix.

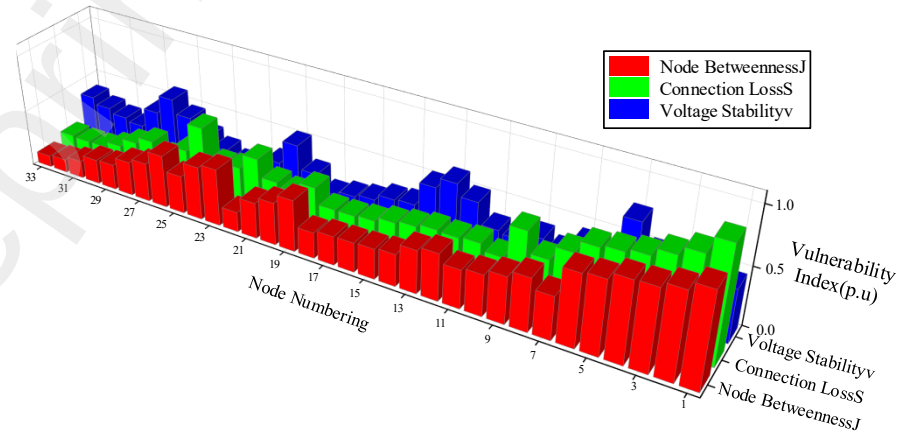


Fig.5 Node vulnerability index value

As depicted in Fig.5, the distribution network's radiation structure positions nodes 1 to 5 as central hubs connected to crucial power transmission channels, resulting in a significantly higher structural vulnerability index compared to other nodes. Conversely, nodes such as 18, 22, and 33 are situated at the periphery of the radiation network where their failure or withdrawal has minimal impact on the overall integrity of the network structure; thus exhibiting low structural vulnerability. However, due to substantial voltage fluctuations caused by operational modes, these peripheral nodes demonstrate a higher state vulnerability index than those located closer to power sources. Nodes 4, 27, and 12 are proximate to renewable energy sources and exhibit increased state vulnerability due to frequent switching of new energy sources leading to output fluctuation indices from fans. This suggests that these particular nodes are prone to failures. Notably observed is the significant disparity in both structural and state vulnerability indices among different nodes indicating varying effects on both grid structure and operational states. Although improved node numbers consider energy transmission and DG influences, they fail to account for local characteristics specific to each node. Connection loss primarily considers distribution network operation mode impacts on power supply stability while voltage stability index evaluates overall operational stability of the entire power network system. Each index possesses inherent limitations thereby rendering it insufficient for accurately assessing node vulnerabilities independently. To comprehensively assess node vulnerability from multiple perspectives, the subjective weights of the three indicators were determined using the analytic hierarchy process. This facilitated decision-making for policymakers by constructing comprehensive vulnerability indicators.

Table 2

Comparison matrix

evaluation index	J	S	v
J	1	1	1/3
S	1	1	1/3
v	3	3	1

When constructing the evaluation matrix, a trilevel scale is employed to determine the relative importance of each index. In considering the significance of node indices, equal importance is attributed to both the node number and connection loss index, while the voltage stability index holds utmost importance. Consequently, a comparison matrix can be formulated as depicted in Table 2 Subsequently, consistency test along with consistency ratio are conducted and normalized eigenvector corresponding to the maximum eigenvalue serves as weight values. The weights assigned to node interconnections, connection losses, and voltage stability indexes are 0.2, 0.2, and 0.6 respectively. Comprehensive vulnerability indicators for nodes are then calculated as illustrated in Fig.6.

The comprehensive vulnerability index of nodes is dimensionless, and a higher value indicates weaker anti-risk ability of the corresponding node and a greater possibility of collapse due to disturbances such as faults. This quantified index can serve as a reference for decision makers when configuring energy storage systems. As depicted in Fig.6, the average comprehensive vulnerability index ranks nodes as $5 \rightarrow 1 \rightarrow 13 \rightarrow 28 \rightarrow 21$ from high to low, highlighting the significant role these nodes play in the power grid. Deploying flexible resources at these critical locations facilitates timely regulation of power flow by grid managers, thereby improving operational status and reducing the risk of instability. Node 1 serves as a balancing node crucial for maintaining voltage and power equilibrium across the entire system; hence it is

not considered in energy storage configuration decisions. Consequently, nodes 5 and 13 are initially selected as potential sites for energy storage.

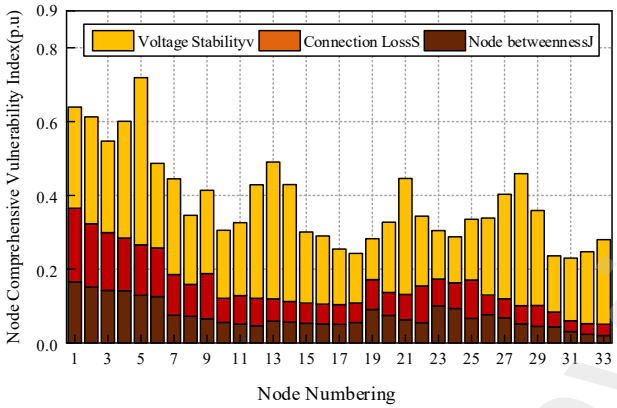


Fig.6 Average comprehensive vulnerability index of nodes

4.3. Analysis of annualized revenue and daily operation scheduling expected value results

The energy storage power stations participating in the electricity market employ day-ahead declaration, intra-day call, and real-time clearing as their trading modes. When formulating bidding strategies, energy storage plants need to predict the electricity prices of both the energy market and auxiliary market in order to guide energy allocation based on price signals. In this study, historical electricity price data from the PJM market for a randomly selected day is utilized to generate day-ahead predicted clearing electricity price scenarios for both the energy market and frequency modulation market (Fig.7). To simulate the stochastic characteristics of electricity market prices, it is assumed that errors in energy clearing prices, frequency modulation capacity prices, and frequency modulation distance prices follow a normal distribution with a mean of 0 and standard deviation of 0.1. Monte Carlo method is employed to conduct 50 sampling iterations.

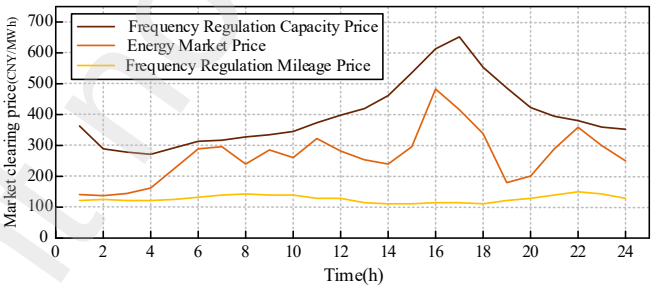


Fig.7 Market clearing price

Grid operators procure electricity at a low price from the energy market and store it for later use, thereby meeting peak-time electricity demand of distribution network users while also achieving profitability. The energy storage power station employs a strategy of securing bids by offering lower prices in day-ahead auctions. The compensation price for energy storage participating in auxiliary services is set at 30 yuan/MWh per unit period, with an additional compensation of 15 yuan/MW for frequency modulation mileage. Energy storage exhibits rapid response and excellent frequency modulation performance, as indicated by an average frequency modulation performance index of 2.79 for the storage station. Enterprises engaging in peak shaving transactions with energy storage are required to declare their electricity load curve and transaction price during the trading period prior to the day ahead, along with compensating the energy storage

facility involved in peak shaving at a rate of 300 yuan/MWh. Furthermore, during peak shaving periods, charging power from the energy storage equipment cannot be utilized elsewhere within the market. Additionally, upper and lower limits were imposed on state-of-charge (SOC) levels for energy storage systems at 0.95 and 0.05 respectively; starting SOC was set at 0.5%.

The operational strategy of energy storage can be categorized into three types: Scheme 1, which focuses solely on energy arbitrage; Scheme 2, involving both energy arbitrage and peak regulation assistance; and Scheme 3, encompassing energy arbitrage, peak regulation assistance, and frequency modulation assistance. Each scheme is analyzed to compare the changes in vulnerability and economic viability among them, leading to a final conclusion.

The results presented in Table 3 demonstrate that scheme three achieves the highest gain, providing preliminary evidence for the economic viability of participating in the multi-market trading capacity allocation strategy. When energy storage solely engages in energy arbitrage, the configured capacity is limited to 13.04MW, resulting in a constrained peak-filling effect on the regional net load curve. However, when participating in market trading, the allocated capacity of energy storage increases by 3.1 and 4.9 times compared to single energy trading schemes respectively. This indicates that higher levels of stored energy enable greater utilization during peak load periods and lead to more significant improvements in smoothing out fluctuations within the distribution network load curve's peak-to-valley difference.

Table 3

Optimal configuration results of different schemes

Operation strategies	Capacity of BESS/MWh	Power of BESS /MW	Objective value/10^5$	difference between peak and valley loads/MW
Scheme 1	13.04	1.78	92.47	1.01
Scheme 2	41.32	3.98	142.56	1.57
Scheme 3	64.54	4.83	235.41	1.92

Table 4

Influences of different configuration strategies on the overall vulnerability of power grid

Operation strategies	Location of BESS	Alterations in global vulnerability/p.u
Scheme 1	5	0.248
	13	0.206
Scheme 2	5	0.215
	13	0.169
Scheme 3	5	0.282
	13	0.233

The change in global vulnerability index is calculated using two different configurations of energy storage: one with five nodes and another with thirteen nodes as shown in Table 4. The grid vulnerability index gradually decreases across all three schemes, indicating enhanced security for overall grid operations; however, it is observed that changes at five nodes are more favorable than those at thirteen nodes. Moreover, larger capacity configurations yield better improvements regarding overall grid vulnerability effects. These findings provide evidence supporting that sufficient capacity plays a crucial role in enabling flexible adjustment capabilities for energy storage systems while reducing peak-to-valley differences benefits stable power grid operation significantly. Furthermore, installing energy storage systems at weak

links proves advantageous for mitigating power grid vulnerabilities. Therefore, based on these considerations and analysis outcomes outlined above, it was ultimately decided to deploy the energy storage system at node 5.

Using historical load data for a specific region, along with annual wind speed and solar irradiance data from a local wind farm and photovoltaic power station, seasonal clustering and analysis were performed to adequately prepare for energy storage planning and address the seasonal differences on both the supply and demand sides. The four typical source-load output scenarios after clustering are shown in Fig.8.

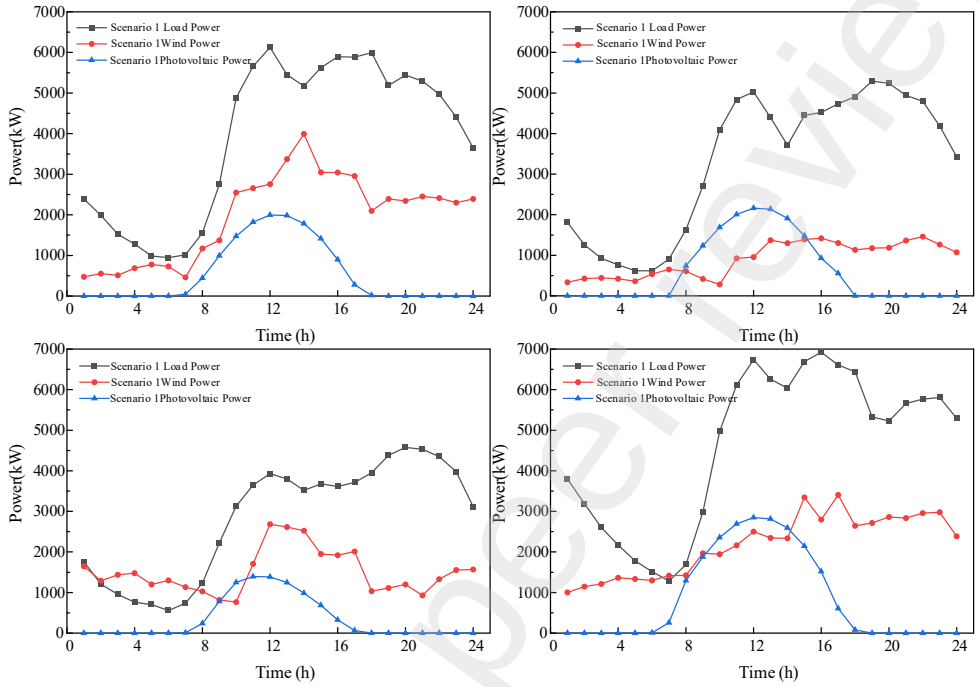
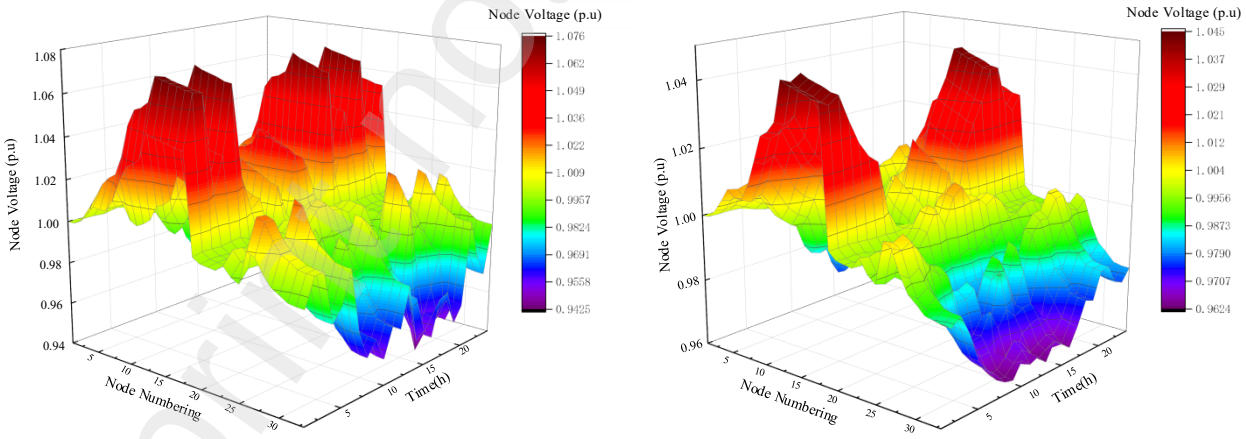


Fig.8 Four typical source-load power scenarios after clustering



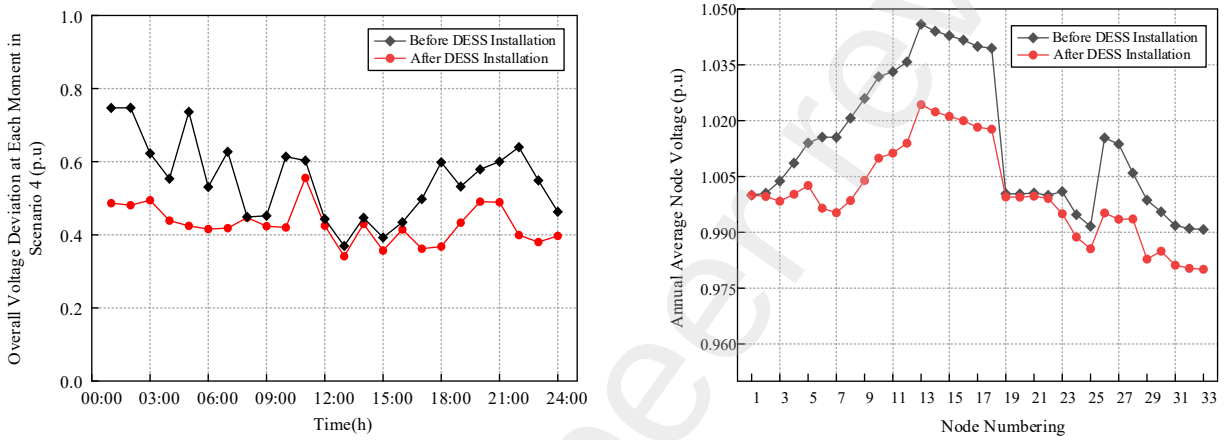
(a) Before BESS configuration (b) After BESS configuration

Fig.9 Voltage of each node

Scenario 4, representing winter, has a steep load curve, the largest peak-valley difference, and the highest load level. Therefore, winter was selected for detailed analysis. The voltage situation after the connection of the DG is depicted in Fig.9(a). Significant fluctuations are observed in the voltages of nodes 5, 13, 21, and 28 connected to the DG, resulting in an access node voltage reaching a peak value of 1.076p.u., which exceeds the system's voltage level requirements and

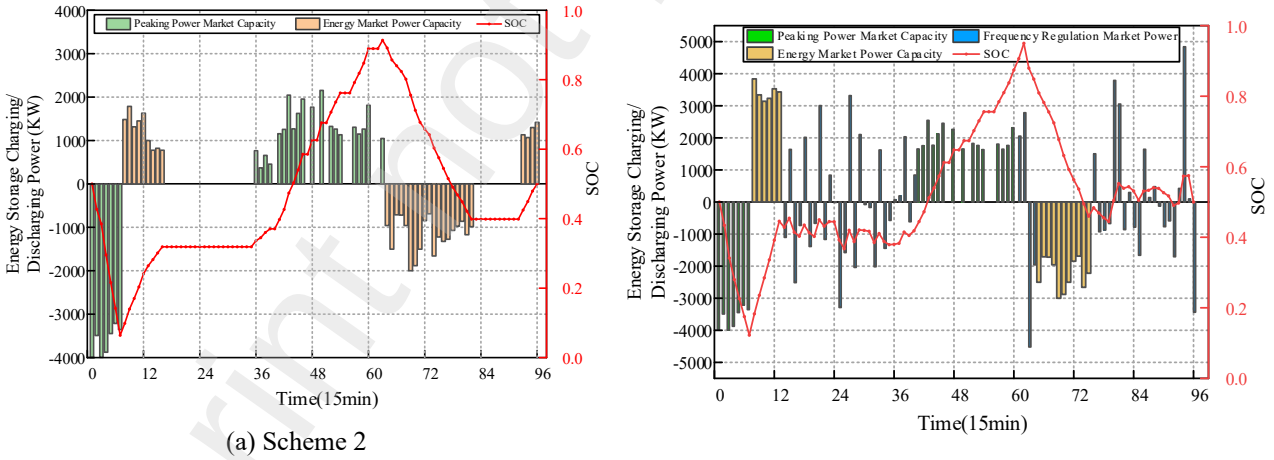
poses an increased risk to distribution network operation. Subsequently, efforts have been made to reduce this voltage to 0.942p.u., however, ensuring power quality for end users along this line remains challenging. When node 5 is configured with BESS, the voltage of each node is depicted in Fig.9(b). The voltage at peak load decreases to 1.045 p.u., while the voltage at the end node rises to 0.962 p.u., both satisfying the quality requirements for power supply voltage of the grid.

As illustrated in Fig.10, scenario 4 exhibits a consistent reduction in overall grid voltage deviation throughout all time periods, with a decrease from 0.68 to 0.34 p.u during peak hours and a stable trend observed in average node voltages over the course of a year, particularly evident by a decline from 1.04 to 1.02 p.u for node 13 connected to the wind farm grid connection point. These results demonstrate that BESS effectively mitigates voltage fluctuations caused by uncertainties associated with wind farm integration into the grid network.



(a) (b)

Fig.10 Scenario 4 overall voltage deviation and annual average voltage



(a) Scheme 2

(b) Scheme 3

Fig.11 Operation strategy of different scheme

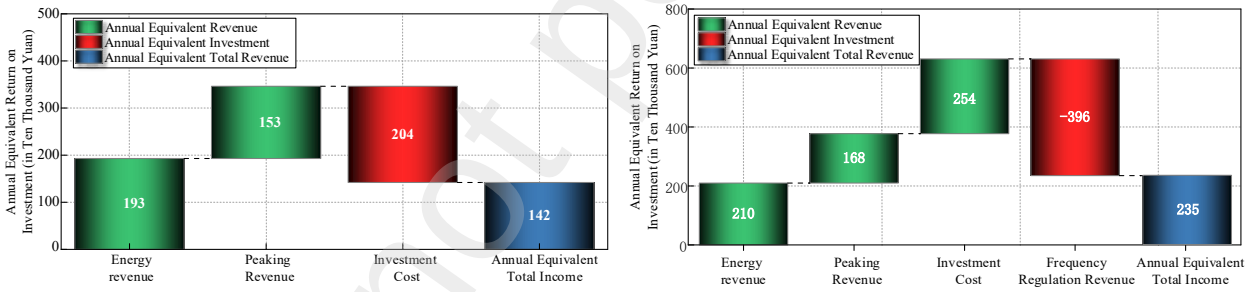
The simulation results of the operation strategy of scheme 2 are presented in Fig.11(a), demonstrating a comprehensive utilization of energy storage capacity in daily operations through nearly full charging and discharge cycles during the 25-hour operation process. However, idle periods occur from 4:00 to 8:00 and 21:00 to 23:00 due to the need for energy storage to reserve sufficient capacity for peak load absorption, limiting its application potential and additional income generation during these time intervals. Consequently, the daily utilization rate of energy storage is calculated at

71%. From Table 5, it can be seen that the peak shaving output utilization rate is 3.2 times that of energy arbitrage in scheme 2. In scheme 2, the annualized investment returns are shown in Fig.12(a). With increasing capacity allocation, solely relying on energy arbitrage cannot fully recover investment costs. Furthermore, limitations on operating days for investors' participation in the peak balancing market result in slightly lower income compared to low storage and high generation scenarios; nevertheless, the total annual value-based income is estimated at approximately 1.54 times that of Plan 1 with a payback period projected at around 6.42 years.

Table 5

Operation strategy results

		Scheme 2	Scheme 3
Daily Operation Plan (KW)	Power Purchased	15950.92	20521.62
	Power Sold	15950.92	20521.62
	Peak Shaving Output	51133.63	54343.93
	Frequency Regulation Output	N/A	77000.34
Daily Operation Revenue (CNY)	Power Purchase Cost	0.529	0.612
	Power Sales Revenue	1.172	1.378
	Peak Shaving Revenue	1.533	1.670
	Frequency Regulation Revenue	N/A	2.536
Total Investment Cost (10,000 CNY)	Construction Cost	882.38	1328.76
	Operation and Maintenance Cost	1340.32	2974.8



(a) Scheme 2 (b) Scheme 3

Fig.12 Annual value income composition

The energy storage operation strategy of scheme 3 is illustrated in Fig.11(b), demonstrating a utilization rate of 100% within a day. This validates the effectiveness of the proposed second- and third-layer operation optimization model in coordinating energy storage to participate in all three markets, thereby maximizing its daily operational profit. However, frequent charging and discharging activities lead to an increase in the number of daily cycles. As a result, the trading period for energy storage participating in peak shaving and energy market is reduced by a total of 4 hours. Considering the high clearing price and comprehensive compensation policy for frequency modulation (FM) in the market, it becomes necessary to reserve additional capacity to achieve optimal FM performance. From Table 5 scheme 3 and Fig.12(b), it is evident that the output from frequency modulation reaches 77 MW, which is 3.85 times higher than that achieved through arbitrage mode in the energy market. This indicates a significant utilization rate for capacity when participating in frequency modulation market operations. Although scheme 3 incurs an equal-annual investment cost that is 1.94 times higher than scheme 2, it generates an annual income that surpasses scheme 2 by ¥910,000 yuan ($140,000). Notably,

revenue from frequency modulation exceeds both peak shaving capability revenue and energy revenue combined with a contribution of approximately 40% towards equal-annual income; thus highlighting how the frequency modulation market can serve as one of the primary sources of income for grid-side energy storage power stations. Furthermore, compared to scheme 2 configuration scenario, the investment recovery period is shortened by 0.4 years signifying improved economic viability.

5. Conclusion

To address the security and economic challenges associated with energy storage allocation in active distribution networks, this paper proposes a trilevel optimization method for energy storage allocation. The key findings and conclusions are as follows:

1) Upper-level: Different vulnerability indices reflecting the importance of nodes in terms of power grid structural safety and operational state are established. By extending the IEEE-33 bus system, vulnerability indices for all nodes are calculated. The results demonstrate that nodes have varying impacts on both structural safety and operational state safety. A comprehensive vulnerability index, based on a weighted combination, effectively reflects node vulnerabilities and identifies the most vulnerable nodes efficiently.

2) Middle-level and lower-level: Based on the upper-level assessment of grid vulnerability, it is concluded that placing energy storage at vulnerable nodes can effectively enhance overall grid resilience. Furthermore, increasing capacity leads to more significant reduces in overall grid vulnerability, thereby enhancing power grid security. The capacity configuration strategy of the energy storage power station and its operation strategy are coupled and iteratively optimized.

3) Economic Perspective: By comparing the proposed trilevel constant volume energy storage power station siting model with different market planning strategies, it is concluded that simultaneous involvement in energy supply-demand balancing, peak shaving, frequency modulation, and other market strategies enables full utilization of energy storage operations. This approach not only enhances grid stability but also improves the economic benefits of grid-side energy storage power stations through multi-purpose planning principles.

In summary, the proposed trilevel optimization model addresses the dual challenges of security and economic efficiency in energy storage allocation, providing a comprehensive and effective solution.

Acknowledgment

This work was supported by Zhenjiang Key R&D Plan (Industry Foresight and Common Key Technology), Project Award Number: GY2023001.

Appendix A

The improved dynamic probability switching diagram.

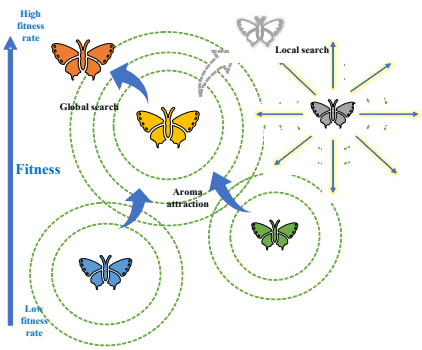


Fig.A.1 Optimization principle of butterfly algorithm

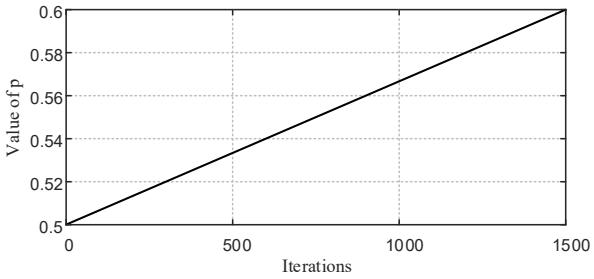


Fig.A.2 Conventional dynamic probability switching diagram

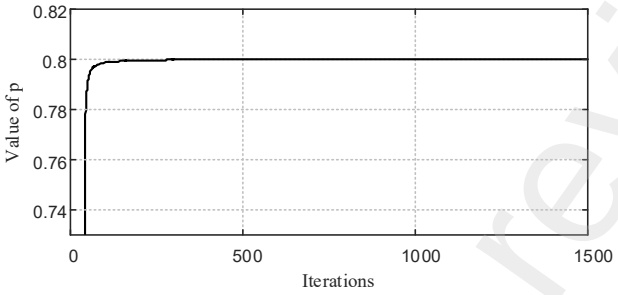


Fig.A.3 Improved dynamic probability switching diagram

The dynamic switching probability increases linearly with the number of iterations, as depicted in Fig.A.2. Based on the proposed formation, the improved dynamic probability switching process is shown in Fig.A.3.

Appendix B

Details on BESS parameters and simulation-related economic indicators.

Table B.1

Relevant parameters of BESS

	Technical Performance Index	Value and Unit
1	Unit Energy Cost	20(10000CNY /MWh)
2	Unit Power Cost	12(10000CNY /MWh)
3	Charge/Discharge Efficiency	95(%)
4	Unit O&M Cost	2000(CNY /MW)
5	Power Conversion Cost	5000(CNY)
6	Construction/System Cost	5(%)
7	Recovery Factor	18(%)
8	Service Life	10(Year)
9	Discount Rate	7(%)
10	Inflation Rate	1.5(%)
11	$SOC_{\max}$	95(%)
12	$SOC_{\min}$	20(%)

Table B.2

Income-related economic indicators

	Economic Indicators	Values and Units
1	Active Loss Cost	2000(CNY /MW)
2	Distribution Network Upgrade Cost	20(10000CNY /MW)

3	Original Capacity of Distribution Network Main Transformer	3.5(MVA)
4	Benchmark Electricity Price	1.08(CNY /KWh)
5	Carbon Trading Price	0.15(CNY /KWh)

Appendix C

The normalized values of vulnerability assessment indicators for each node.

Table C.1

Normalized values of average node vulnerability indicators

Node	*J*	*S*	*v*	Node	*J*	*S*	*v*
1	0.826	1	0.457	18	0.277	0.265	0.224
2	0.758	0.966	0.483	19	0.453	0.405	0.185
3	0.712	0.784	0.413	20	0.373	0.312	0.317
4	0.706	0.716	0.527	21	0.315	0.343	0.524
5	0.646	0.684	0.755	22	0.276	0.497	0.315
6	0.629	0.659	0.382	23	0.502	0.364	0.219
7	0.375	0.553	0.432	24	0.467	0.348	0.208
8	0.362	0.432	0.312	25	0.333	0.521	0.274
9	0.327	0.613	0.376	26	0.384	0.265	0.348
10	0.278	0.328	0.307	27	0.342	0.257	0.472
11	0.256	0.387	0.329	28	0.261	0.245	0.596
12	0.232	0.373	0.513	29	0.225	0.283	0.429
13	0.297	0.301	0.618	30	0.217	0.202	0.254
14	0.281	0.281	0.528	31	0.154	0.147	0.283
15	0.266	0.275	0.321	32	0.117	0.145	0.325
16	0.258	0.268	0.308	33	0.101	0.153	0.382
17	0.253	0.267	0.251				

References

[1] Ghaffari A, Askarzadeh A, Fadaeinedjad R. Optimal allocation of energy storage systems, wind turbines and photovoltaic systems in distribution network considering flicker mitigation. Appl Energy. 2022;319:119253.https://doi.org/10.1016/j.apenergy.2022.119253.

[2] Alnawafah H, Sarrias-Mena R, Harb A, Fernández-Ramírez LM, Llorens-Iborra F. Evaluating the inertia of the Jordanian power grid. Comput Electr Eng. 2023;109:108748.https://doi.org/10.1016/j.compeleceng.2023.108748.

[3] Zhang H, Fu S, Jiang J. Multi-period network equilibrium in power system with energy storage on generation side. Electr Power Syst Res. 2023;223:109533.https://doi.org/10.1016/j.epsr.2023.109533.

[4] Dozein MG, Gomis-Bellmunt O, Mancarella P. Simultaneous provision of dynamic active and reactive power response from utility-scale battery energy storage systems in weak grids. IEEE Trans Power Syst. 2021;36:5548-57.https://doi.org/10.1109/TPWRS.2021.3076218.

[5] Hou T, Fang R, Yang D, Zhang W, Tang J. Energy storage system optimization based on a multi-time scale decomposition-coordination algorithm for wind-ESS systems. Sustainable Energy Technol Assess. 2022;49:101645.https://doi.org/10.1016/j.seta.2021.101645.

[6] Yan HW, Liang G, Beniwal N, Rodriguez E, Farivar GG, Pou J. Flexible power point tracking aided power ramp rate control for photovoltaic systems with small energy storage capacity. IEEE Trans Power Electron. 2024;39:2798-810.https://doi.org/10.1109/TPEL.2023.3333534.

[7] Zheng L, Wang J, Chen J, Ye C, Gong Y. Two-stage co-optimization of a park-level integrated energy system considering grid interaction. IEEE Access. 2023;11:66400-14.https://doi.org/10.1109/ACCESS.2023.3290160.

[8] Javadi MS, Gough M, Mansouri SA, Ahmarinejad A, Nematbakhsh E, Santos SF, et al. A two-stage join

t operation and planning model for sizing and siting of electrical energy storage devices considering dem and response programs. Int J Electr Power Energy Syst. 2022;138:107912.https://doi.org/10.1016/j.ijepes.2021.107912.
[9] Yan C, Geng X, Bie Z, Xie L. Two-stage robust energy storage planning with probabilistic guarantees: A data-driven approach. Appl Energy. 2022;313:118623.https://doi.org/10.1016/j.apenergy.2022.118623.
[10] Jiang X, Chen J, Zhang W, Wu Q, Zhang Y, Liu J. Two-step optimal allocation of stationary and mobil e energy storage systems in resilient distribution networks. J Modern Power Syst Clean Energy. 2021;9:7 88-99.https://doi.org/10.35833/MPCE.2020.000910.
[11] Jani A, Karimi H, Jadid S. Multi-time scale energy management of multi-microgrid systems considering e nergy storage systems: a multi-objective two-stage optimization framework. J Energy Storage. 2022;51:104 554.https://doi.org/10.1016/j.est.2022.104554.
[12] Ren Z, Guo H, Yang P, Zuo G, Zhao Z. Bi-Level optimal allocation of flexible resources for distribution network considering different energy storage operation strategies in slectricity market. IEEE Access. 2020; 8:58497-508.https://doi.org/10.1109/ACCESS.2020.2983042.
[13] Fakih S, Mabrouk MT, Batton-Hubert M, Lacarriere B. Bi-level and multi-objective optimization of renew able energy sources and storage planning to support existing overloaded electricity grids. Energy Rep. 20 23;10:1450-66.https://doi.org/10.1016/j.egyr.2023.08.015.
[14] Zhou S, Han Y, Chen S, Yang P, Mahmoud K, Darwish MMF, et al. A multiple uncertainty-based bi-le vel expansion planning paradigm for distribution networks complying with energy storage system function alities. Energy. 2023;275:127511.https://doi.org/10.1016/j.energy.2023.127511.
[15] Ma M, Huang H, Song X, Peña-Mora F, Zhang Z, Chen J. Optimal sizing and operations of shared ener gy storage systems in distribution networks: a bi-level programming approach. Appl Energy. 2022;307:118 170.https://doi.org/10.1016/j.apenergy.2021.118170.
[16] Gough M, Santos SF, Javadi MS, Home-Ortiz JM, Castro R, Catalão JPS. Bi-level stochastic energy tradi ng model for technical virtual power plants considering various renewable energy sources, energy storage systems and electric vehicles. J Energy Storage. 2023;68:107742.https://doi.org/10.1016/j.est.2023.107742.
[17] Wan T, Tao Y, Qiu J, Lai S. Data-driven hierarchical optimal allocation of battery energy storage system. IEEE Trans Sustainable Energy. 2021;12:2097-109.https://doi.org/10.1109/TSTE.2021.3080311.
[18] Toubeau JF, Bottieau J, Grève ZD, Vallée F, Bruninx K. Data-driven scheduling of energy storage in da y-ahead energy and reserve markets with probabilistic guarantees on real-time delivery. IEEE Trans Power Syst. 2021;36:2815-28.https://doi.org/10.1109/TPWRS.2020.3046710.
[19] [Yu WANG, Ke LI, Shuzhen LI, Xin MA, Chenghui ZHANG. A bi-level scheduling strategy for integrat ed energy systems considering integrated demand response and energy storage co-optimization. J Energy Storage. 2023;66:107508.https://doi.org/10.1016/j.est.2023.107508.
[20] Kelly JJ, Leahy PG. Optimal investment timing and sizing for battery energy storage systems. J Energy Storage. 2020;28:101272.https://doi.org/10.1016/j.est.2020.101272.
[21] Rancilio G, Merlo M, Lucas A, Kotsakis E, Delfanti M. BESS modeling: investigating the role of auxilia ry system consumption in efficiency derating. 2020 International Symposium on Power Electronics, Electri cal Drives, Automation and Motion (SPEEDAM)2020. p. 189-94.https://doi.org/10.1109/SPEEDAM48782.2020.9161875.
Luo Q, Yu D, Cheng X, Sheng H, Lyu W. Exploring truss maintenance in fully dynamic graphs: A mix ed structure-based approach. IEEE Trans Comput. 2023;72:707-18.https://doi.org/10.1109/TC.2022.3174594.
[22] Xu X, Hu W, Cao D, Huang Q, Chen C, Chen Z. Optimized sizing of a standalone PV-wind-hydropowe r station with pumped-storage installation hybrid energy system. Renewable Energy. 2020;147:1418-31.https://doi.org/10.1016/j.renene.2019.09.099.
[23] Mansuwan K, Jirapong P, Thararak P. Optimal battery energy storage planning and control strategy for g rid modernization using improved genetic algorithm. Energy Rep. 2023;9:236-41.https://doi.org/10.1016/j.egyr.2023.09.017.
[24] Calloquispe-Huallpa R, Huaman-Rivera A, Ordoñez-Benavides AF, Garcia-Garcia YV, Andrade-Rengifo F, Aponte-Bezares EE, et al. A comparison between genetic algorithm and particle swarm optimization for e conomic dispatch in a microgrid. 2023 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT-LA)2023. p. 415-9.https://doi.org/10.1109/ISGT-LA56058.2023.10328280.
[25] Hao P, Sobhani B. Application of the improved chaotic grey wolf optimization algorithm as a novel and efficient method for parameter estimation of solid oxide fuel cells model. Int J Hydrogen Energy. 2021;46:36454-65.https://doi.org/10.1016/j.ijhydene.2021.08.174.
[26] Song X, Zhang H, Fan L, Zhang Z, Peña-Mora F. Planning shared energy storage systems for the spatio -temporal coordination of multi-site renewable energy sources on the power generation side. Energy. 2023;282:128976.https://doi.org/10.1016/j.energy.2023.128976.
[27] Kang D, Kang D, Hwangbo S, Niaz H, Lee WB, Liu JJ, et al. Optimal planning of hybrid energy stora

ge systems using curtailed renewable energy through deep reinforcement learning. Energy. 2023;284:128623.https://doi.org/10.1016/j.energy.2023.128623.
[28] Feng B, Fu Y, Huang Q, Ma C, Sun Q, Wennersten R. Multi-objective optimization of an integrated energy system with high proportion of renewable energy under multiple uncertainties. Energy Rep. 2023;9:695-701.https://doi.org/10.1016/j.egyr.2023.05.194.
[29] Han O, Ding T, Zhang X, Mu C, He X, Zhang H, et al. A shared energy storage business model for data center clusters considering renewable energy uncertainties. Renewable Energy. 2023;202:1273-90.https://doi.org/10.1016/j.renene.2022.12.013.
[30] Wang Y, Tang L, Yang Y, Sun W, Zhao H. A stochastic-robust coordinated optimization model for CCHP micro-grid considering multi-energy operation and power trading with electricity markets under uncertainties. Energy. 2020;198:117273.https://doi.org/10.1016/j.energy.2020.117273.
[31] Hedayati-Mehdiabadi M, Hedman KW, Zhang J. Reserve policy optimization for scheduling wind energy and reserve. IEEE Trans Power Syst. 2018;33:19-31.https://doi.org/10.1109/TPWRS.2017.2707568.
[32] Zhang H, Ding T, Liu Y, Zhang X, Li L, Zhang Q, et al. Two-stage stochastic unit commitment for renewable energy integrated power systems considering dynamic capacity-increase technologies of transmission lines. Energy Rep. 2023;9:129-33.https://doi.org/10.1016/j.egyr.2023.09.122.
[33] Huang W, Li H, Yin Y, Zhang Z, Xie A, Zhang Y, et al. Node importance identification of unweighted urban rail transit network: An adjacency information entropy based approach. Reliab Eng Syst Saf. 2024;242:109766.https://doi.org/10.1016/j.ress.2023.109766.
[34] Wu W, Hou H, Wu X, Wei R, He H, Shi Y. Emergency repair strategy of distribution network under wind disaster based on multidimensional importance evaluation. Energy Rep. 2023;9:275-82.https://doi.org/10.1016/j.egyr.2023.04.302.
[35] Liu Y, Tang M, Zhou T, Do Y. Identify influential spreaders in complex networks, the role of neighborhood. Physica A. 2016;452:289-98.https://doi.org/10.1016/j.physa.2016.02.028.
[36] Zhang Z, Wang Z, Chen Z, Wang G, Shen N, Guo C. Study on grid-connected strategy of distribution network with high hydropower penetration rate in isolated operation. Processes. 2019.https://doi.org/10.3390/pr7060328.
[37] Wang X, Kang Q, Gao J, Zhang F, Wang X, Qu X, et al. Distribution network restoration supply method considers 5G base station energy storage participation. Energy. 2024;289:129825.https://doi.org/10.1016/j.energy.2023.129825.
[38] Zhang S, Liu W, Wan H, Bai Y, Yang Y, Ma Y, et al. Combing data-driven and model-driven methods for high proportion renewable energy distribution network reliability evaluation. Int J Electr Power Energy Syst. 2023;149:108941.https://doi.org/10.1016/j.ijepes.2022.108941.
[39] Makhadmeh SN, Al-Betar MA, Abasi AK, Awadallah MA, Doush IA, Alyasseri ZAA, et al. Recent advances in butterfly optimization algorithm, its versions and applications. Arch Comput Methods Eng. 2023;30:1399-420.https://doi.org/10.1007/s11831-022-09843-3.
[40] Makhadmeh SN, Al-Betar MA, Abasi AK, Awadallah MA, Doush IA, Alyasseri ZAA, et al. Recent advances in butterfly optimization algorithm, its versions and applications. Arch Comput Methods Eng. 2023;30:1399-420.https://doi.org/10.1007/s11831-022-09843-3.