

## Stochastic optimization for capacity configuration of data center microgrid thermal energy management equipment considering flexible resources

Yang Cui<sup>a</sup>, Yufeng Cheng<sup>a</sup>, Han Zhu<sup>a,\*</sup>, Yuting Zhao<sup>a</sup>, Wuzhi Zhong<sup>b</sup>

<sup>a</sup> Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education (Northeast Electric Power University), Jilin 132012, China

<sup>b</sup> China Electric Power Research Institute, Beijing 100192, China



### ARTICLE INFO

#### Keywords:

Data center  
Flexible resources  
Wind power scenario generation  
Capacity configuration  
Stochastic optimization

### ABSTRACT

The rapid development of Internet and cloud computing technologies has led to the expansion of the scale of data centers, resulting in a continuous increase in the demand for data processing and storage. This not only results in higher energy consumption but also makes thermal management a key factor of energy management in data centers. A capacity configuration method is proposed for the core thermal management equipment of data center microgrids, electric boilers, cooling systems, and heat pumps. This method not only considers the uncertainty factors of the source and load but also effectively utilizes the characteristics of various flexible resources in data centers, such as the adjustable ability of different types of batch processing loads, air thermal inertia, and waste heat recovery, significantly improving the energy utilization efficiency and enhancing the system's economy and flexibility. In addition, a wind power scenario generation method based on conditional least squares generative adversarial networks was proposed to improve the generation quality of wind power scenarios, and a random optimization model was constructed. Considering a certain province's data center microgrid as a case study, the effectiveness of the model was verified, and a sensitivity analysis of the relevant parameters provided important references for microgrid planning. This study provides a new perspective and tool for the thermal management of data center microgrids and comprehensive utilization of renewable energy, which is of great significance for promoting the improvement of data center energy efficiency and sustainable development.

### 1. Introduction

In contemporary society, data centers play an increasingly important role as the infrastructure supporting the Internet and information technology services [1]. Utilizing the flexibility of data centers to effectively manage energy aligns with global initiatives, such as Sustainable Energy for All [2], Sustainable Development Goals (SDG) [3,4], and commitments under the Paris Agreement and Nationally Determined Contributions [5]. This approach contributes to energy efficiency and emission reduction efforts in accordance with international policies and agreements aimed at sustainable development. According to relevant data statistics [6–8], the United States currently has over 5,000 data centers, accounting for 33 % of the global total, and holds a leading position in terms of quantity. Germany ranks second with 521 data centers. China, with 449 data centers, is one of the top countries in terms of data center infrastructure. Other significant countries including France, Australia, the Netherlands, Russia, Japan, and Italy also possess

a large amount of data center infrastructure. Data center microgrids that integrate renewable and conventional energy resources, complemented by smart grid technologies and energy management strategies, offer an efficient solution for enhancing the energy efficiency and sustainability of data centers [9]. With the continual growth in data processing and storage demands, the energy consumption of data centers is rapidly increasing, drawing widespread global attention to their energy efficiency and environmental impact [10]. In this context, we focus on capacity configuration schemes for thermal energy management equipment to improve the energy utilization efficiency and the overall system sustainability, with special attention paid to various flexible resources in data centers, such as different types of batch processing loads, air thermal inertia, and waste heat recovery. This study considers different types of batch processing loads as direct flexible resources, whereas air thermal inertia and waste heat recovery are considered indirect flexible resources, whose effective utilization is crucial for optimizing thermal energy management in data centers.

\* Corresponding author.

E-mail addresses: [chengyufeng0522@163.com](mailto:chengyufeng0522@163.com) (Y. Cheng), [zhuhan\\_97@126.com](mailto:zhuhan_97@126.com) (H. Zhu).

For data center microgrids, the primary concern for operators is reducing investment and operational maintenance costs [11]. Current research focuses on operational optimization strategies for data center microgrids, with most studies leveraging the adjustability of data center loads. Data center loads are crucial adjustable resources and are generally categorized into interactive and batch-processing loads [12]. Interactive loads refer to workloads that require immediate or near-immediate responses, whereas batch processing loads have more flexible timing requirements, allowing tasks to be completed over longer periods. Reference [13] utilized the spatial transfer capability of interactive loads and established a multi-microgrid cooperative game scheduling model based on Nash bargaining theory. Reference [14] promoted interactive load and energy coordination among data centers in regional clusters and developed an online algorithm based on the Lyapunov optimization framework. Reference [15] establishes a collaborative online scheduling framework for local interconnected data centers considering interactive loads and shared energy storage. Reference [16] incentivized the potential of data center users, focusing on analyzing the flexible scheduling of data center batch loads to better utilize the possibilities of renewable energy and respond to grid demands. Reference [17] fully utilized the spatial transfer capability of interactive loads and the time transfer capability of batch loads to determine the location and capacity of interconnected Internet data centers and batteries. Reference [18] addressed the problem of high and fluctuating energy loads in data center servers, proposing a server optimization method that considers the spatial and temporal load distribution of data centers. Reference [19] proposed a novel decomposition method based on the adjustable capacity of batch loads, which effectively coordinates the spatial relationship between data centers and power systems. Reference [20] suggests that considering the adjustability of batch processing loads can help reduce network losses. Reference [21] argues that the adjustability of batch processing loads could mitigate market electricity price fluctuations to certain extent. These studies analyzed the impact of the data center load adjustability on the economic operation of the system and provided important references for this study. However, the classification of batch processing loads in previous studies has not been sufficiently detailed. In reality, batch processing loads have various types and different adjustability capabilities, such as handling multiple small-scale file transfer tasks at a certain moment or large-scale data processing tasks spread across several moments. Few studies further subdivided batch processing loads based on their adjustability characteristics and established a general mathematical model. Further, most existing research has concentrated on the scheduling level, with studies on capacity configuration still in their early stages of development.

Data center operators can not only fully leverage the adjustability of

loads to reduce their own costs, but also consider the inherent air thermal inertia and the recycling of a large amount of waste heat generated during the operation of data centers to further improve energy utilization and overall system sustainability. However, current research often overlooks these two points, particularly in terms of capacity configuration. To keep IT equipment in data centers operating efficiently, indoor temperatures must be maintained within a certain range, with air thermal inertia playing a key role in the temperature management of data centers [22]. Reference [23] states that although research on cooling systems has made certain progress, most studies have focused on the impact of a single factor on energy consumption and thermal management, lacking a systematic understanding of the comprehensive impact factors of air thermal inertia. The experiments in [24] revealed the potential of optimizing air distribution systems to improve thermal performance but also exposed the challenge of effectively utilizing air thermal inertia to further reduce energy consumption and improve efficiency in practical applications. Although the study in [25] does not directly mention air thermal inertia, the improved design of the air-cooled server chassis emphasizes the importance of considering air flow and heat dissipation paths in the thermal management of data centers.

To further reduce the energy consumption of data centers, [26] proposed the adoption of a circulating water-cooling system in data centers. Compared with traditional air cooling, water cooling has advantages such as high heat transfer efficiency, low noise, small footprint, low carbon, and environmental friendliness [27], promising a bright future. The water cooling system is not only an “energy consumer” but also a controllable cold load. The operation of data centers generates a significant amount of heat, posing challenges to water-cooling systems while offering opportunities for energy recovery [28]. Reference [29] describes the design of an integrated data center cooling and waste heat recovery system capable of switching operating modes based on the environmental temperature. Reference [30] integrates an organic Rankine cycle system into a data center cooling system for waste heat recovery and analyzed its thermodynamic performance and economic impact. The working principle of air source heat pumps, based on absorbing heat from low-temperature heat sources and transferring it to higher-temperature heat sinks, makes them excellent devices for waste heat recovery [31], and their working characteristics can be well coupled with the cooling system of the data center. If the reuse of waste heat is considered, the water cooling system also acts as a heat source, thus forming an electricity-heat-cold coupled data center microgrid.

These studies primarily focus on flexible resources within data centers but overlook another critical factor. Recent trend of integrating renewable energy, particularly wind power, into data center power supplies is becoming more apparent, with companies such as Google and Microsoft purchasing substantial amounts of wind power for their data centers [32]. The operation of data center microgrids demands high reliability in energy supply and demand, and any interruption in energy supply can lead to data loss or service disruption, causing significant losses to businesses and users [33]. Considering the flexible resources in data centers and applying the capacity configuration method for thermal energy management equipment to engineering practice will inevitably be influenced by uncertainties in source and load. This paper primarily discusses the uncertainty brought about by the random fluctuations of wind power output. Currently, in the field of optimization decision-making considering the uncertainty of new energy generation, the mainstream methods include robust optimization (RO), distributionally robust optimization (DRO), and stochastic optimization (SO) [34]. RO does not rely on precise probability distributions of uncertainty parameters but instead focuses on decision-making for the worst-case scenarios of renewable energy generation, often resulting in overly conservative planning schemes [35]. Compared to RO, DRO lacks a conservative assessment of the uncertainty of wind and solar power generation. The DRO creates a collection of fuzzy probability distributions of uncertain factors to identify the most adverse probability

**Table 1**  
Methods of generative scene generation.

Reference	Method	Applications
[4041]	C-WGAN-GP	Wind power and photovoltaic scenario generation
[42]	CGAN	Wind power scenario generation.
[43]	AM-GAN	Wind power scenario generation.
[4445]	WGAN-GP	Wind power and photovoltaic scenario generation
[46]	CGAN	Load scenario generation.
[47]	RAC-GAN	Wind power scenario generation.
[48]	WGAN-GP	Load scenario generation.
[49]	C-DCGAN	Wind power and photovoltaic scenario generation
[5051]	GAN	Load scenario generation.
[52]	GAN	Wind power scenario generation.
[53]	GAN	Wind power and photovoltaic scenario generation
[5455]	WGAN	Wind power scenario generation.
[56]	WGAN, CGAN	Wind power and photovoltaic scenario generation
[57]	LSGAN	Wind power and photovoltaic scenario generation
[58]	CGAN	Wind power and photovoltaic scenario generation
[59]	ctrl-GANs	Wind power and photovoltaic scenario generation
[60]	C <sup>2</sup> RGAN	Photovoltaic and load scenario generation
This paper	C-LSGANs	Wind power scenario generation

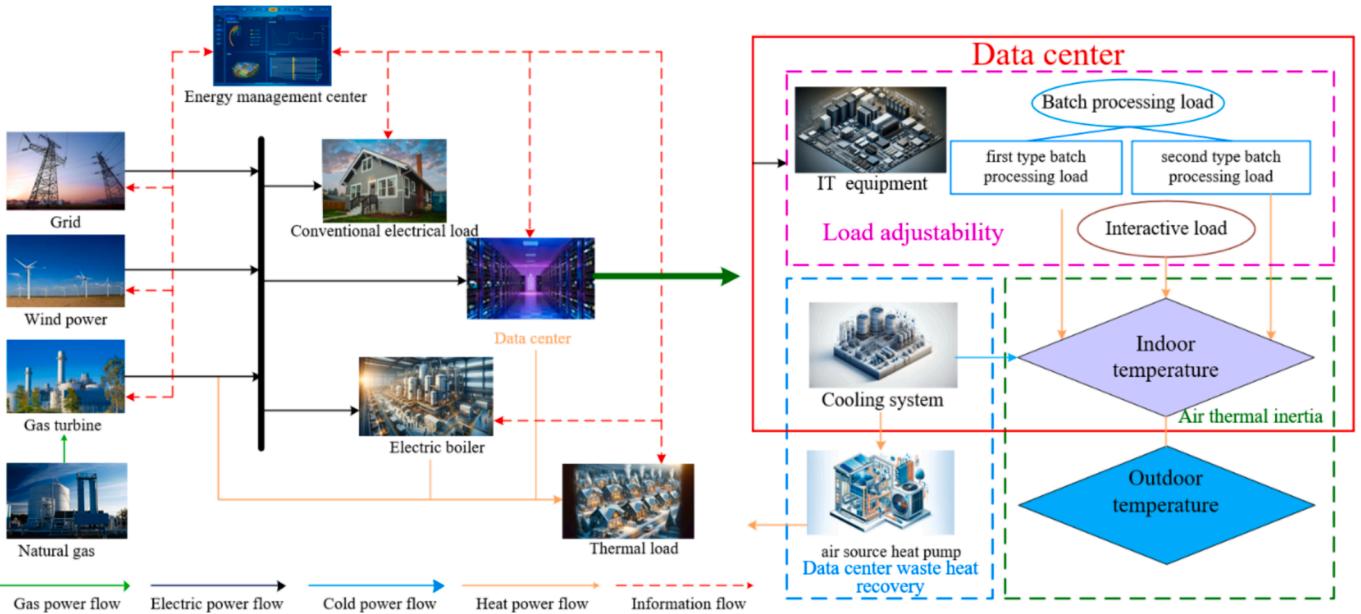


Fig. 1. Data center microgrid framework diagram.

distribution as the basis for decision making [36]. However, DRO faces challenges in modeling the complexity of the spatiotemporal correlation of random variables and solving them, and have limitations in constructing fuzzy probability distribution sets owing to subjective experience [37]. By establishing probability models of uncertain factors and generating discrete time-series scenarios, the SO transforms uncertainty optimization problems into deterministic problems. It can effectively represent time-series correlations and is easier to solve than other models [38]. In recent years, with the application of artificial intelligence technologies, particularly generative networks, to model the uncertainty of renewable energy, scenario generation models based on these technologies have been used to create wind and photovoltaic power generation scenarios with temporal and spatial correlations [39]. This progress not only addresses the issues of poor quality or insufficient quantity of original operational data but also provides a solid data foundation for stochastic optimization models based on scenario analysis. Various methods are available for scenario generation (Table 1). Although existing methods such as C-WGAN-GP, CGAN, and AM-GAN have achieved effectiveness to a certain extent in scenario generation, there is still scope for further improvement in accurately characterizing the uncertainty of wind power output.

In summary, compared to existing research, the main contributions of this study are as follows:

1. From the perspective of the adjustable characteristics of batch loads, this study divided batch loads into finer categories to quantify their adjustability more accurately, which facilitated the implementation of demand response strategies
2. From the viewpoint of thermal management, this study conducted capacity allocation for core thermal management equipment in data center microgrids, effectively utilizing flexible resources while considering uncertainty factors, thereby improving energy utilization efficiency and enhancing system flexibility and economy
3. To address the challenge posed by wind power uncertainty in capacity allocation, a wind power scenario generation method based on conditional least-squares generative adversarial networks (C-LSGANs) is proposed to improve the quality of scenario generation
4. In-depth discussions were conducted on parameters such as the proportion of the first type of batch load and the allowable delay to understand their impact on the system planning results. This helps

data center operators set the corresponding configuration schemes based on the relevant data

The remainder of this paper is organized as follows. Section 2 introduces the architecture of data center microgrids and establishes mathematical models for various flexible resources within the data center. Section 3 proposes a scenario generation method based on C-LSGANs. Section 4 establishes a stochastic optimization model for the capacity configuration of data center microgrids. Section 5 presents a case study that verifies the effectiveness of the proposed strategy. Finally, the conclusions are presented in Section 6.

## 2. Data center flexible resource modeling

Fig. 1 presents the typical architecture of a data center microgrid, which encompasses wind turbine units, conventional power generation units, and the data center itself, with capabilities for both purchasing and selling electricity to a superior grid. The flexible resources considered include the data center's batch processing loads, the thermal inertia of the data center's air, and the recovery and utilization of waste heat. The adjustable capacity of batch processing loads is regarded as a direct flexible resource, whereas the thermal inertia of the data center air and the recovery and utilization of waste heat are considered as indirect flexible resources. A detailed analysis and model of these types of flexible resources is provided below.

### 2.1. Data center load adjustability modeling

Data center loads can be categorized into interactive and batch processing loads based on their real-time requirements. Interactive loads are workloads that require immediate response and interaction, including tasks such as web browsing, real-time communication, online gaming, and video conferencing. Interactive loads necessitate that data centers process requests quickly and provide immediate results to ensure a smooth user experience. By contrast, batch processing loads do not require an immediate response and encompass tasks such as large-scale data processing, small-scale file transfer, and regular report generation with lower real-time requirements.

This study further subdivided batch processing loads into two categories based on their adjustable characteristics: bandwidth-variable, temporally variable, time-limited shiftable loads, and bandwidth-

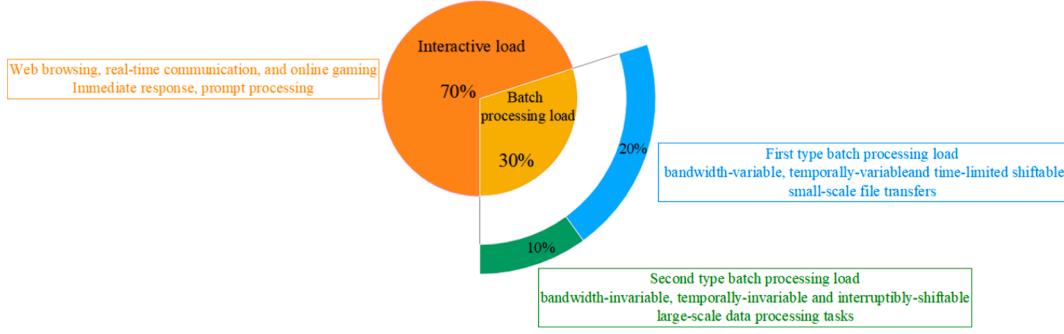


Fig. 2. Schematic of data center load characteristics.

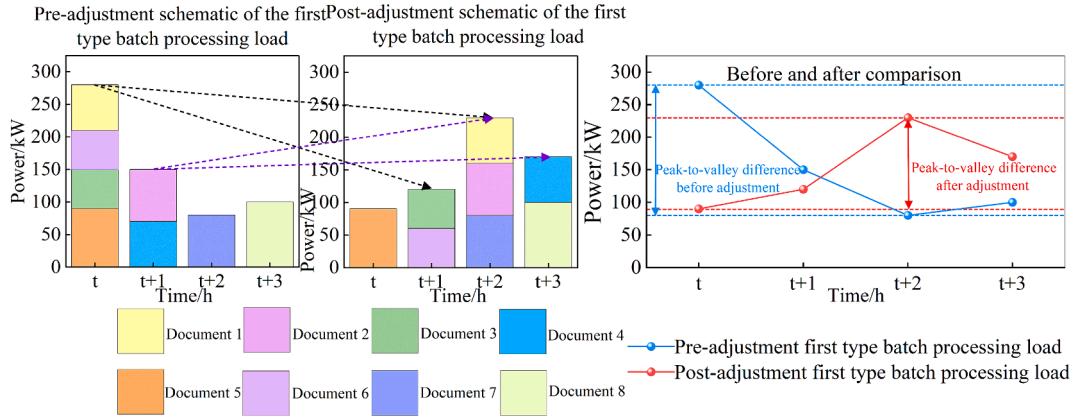


Fig. 3. Schematic of the adjustability of the first type batch processing load.

invariable, temporally invariable, and interruptible shiftable loads.

The former pertains to a variety of tasks waiting for processing before adjustment at time  $t$ , such as small-scale file transfers. This study refers to this as the first category of batch processing load. The latter relates to batch processing loads that are engaged in the same task and wait for processing from time  $t$  to  $t + y$ , similar to large-scale data processing tasks. This study refers to this as the second category of batch processing load.

The term 'bandwidth' is used to denote the capacity requirements for processing batch load tasks, essentially reflecting the size demands of the batch processing load, as elaborated in Reference [61]. The bandwidth variability or invariability refers to whether the size of the load at time  $t$  before the adjustment changes or remains the same after shifting to time  $t + x$ . Temporal variability or invariability refers to whether the sequence of processing loads from time  $t$  to  $t + y$  is altered or remained constant after adjustment.

In practice, interactive loads constitute approximately 70 % of the total data center load, whereas the first category of batch processing loads accounts for approximately 20 %, and the second category of batch processing loads comprises 10 %. The specific characteristics of the data center loads are shown in Fig. 2. A detailed analysis of these two types of batch processing load is provided below.

### 2.1.1. First type batch processing load adjustability

The first category of batch processing loads bears certain resemblance to traditional shiftable loads, providing a degree of flexibility to data centers and enabling them to be considered as demand response loads participating in the grid's operational optimization scheduling [62]. However, unlike traditional shiftable loads, they must be completely processed before their deadlines [63], embodying the characteristic of being time-limited and shiftable. This characteristic is schematically shown in Fig. 3.

As illustrated in Fig. 3, the adjustable capacity of the first category of batch-processing loads aids in peak shaving and valley filling, thereby reducing the peak-to-valley load difference. This study posits that the maximum delay time for processing the first category of batch processing loads is two hours, with modeling as follows:

First, we introduce the time-limited shiftable matrix  $\mathbf{A}$ , initial diagonal matrix  $\mathbf{P}_{dy,dzq}$  for the first category of batch processing loads before adjustment, and load matrix  $\mathbf{P}_{dy,pv}$  for the first category of batch processing loads after adjustment. The dimensions of all three matrices were  $24 \times 24$ , and the specific formula is as follows:

$$0 \leq \mathbf{P}_{dy,pv} \leq \mathbf{A} \cdot \mathbf{P}_{dy,dzq} = \mathbf{B} \quad (1)$$

where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & \dots & 1 & 1 \\ 1 & 1 & 0 & \dots & 0 & 1 \\ 1 & 1 & 1 & \dots & \vdots & 0 \\ 0 & 1 & \vdots & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 1 \end{pmatrix} \quad (2)$$

$$\mathbf{P}_{dy,dzq} = \begin{pmatrix} b_{1,1} & 0 & 0 & \dots & 0 & 0 \\ 0 & b_{2,2} & 0 & \dots & 0 & 0 \\ 0 & 0 & b_{3,3} & \dots & \vdots & 0 \\ 0 & 0 & \vdots & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \dots & b_{23,23} & 0 \\ 0 & 0 & 0 & \dots & 0 & b_{24,24} \end{pmatrix} \quad (3)$$

where the subscripts  $t_1$  and  $t_2$  of  $b_{t_1,t_2}$  ( $t_1, t_2 = 1, 2, \dots, 24$ ) represent the row  $t_1$  and column  $t_2$  of the matrix  $\mathbf{P}_{dy,dzq}$ , respectively.  $b_{t,t}$  denotes the size of the first category of batch processing load at time  $t$  before adjustment, distributed along the diagonal of  $\mathbf{P}_{dy,dzq}$ . The introduction of this matrix indicates that at each moment before adjustment, there is a

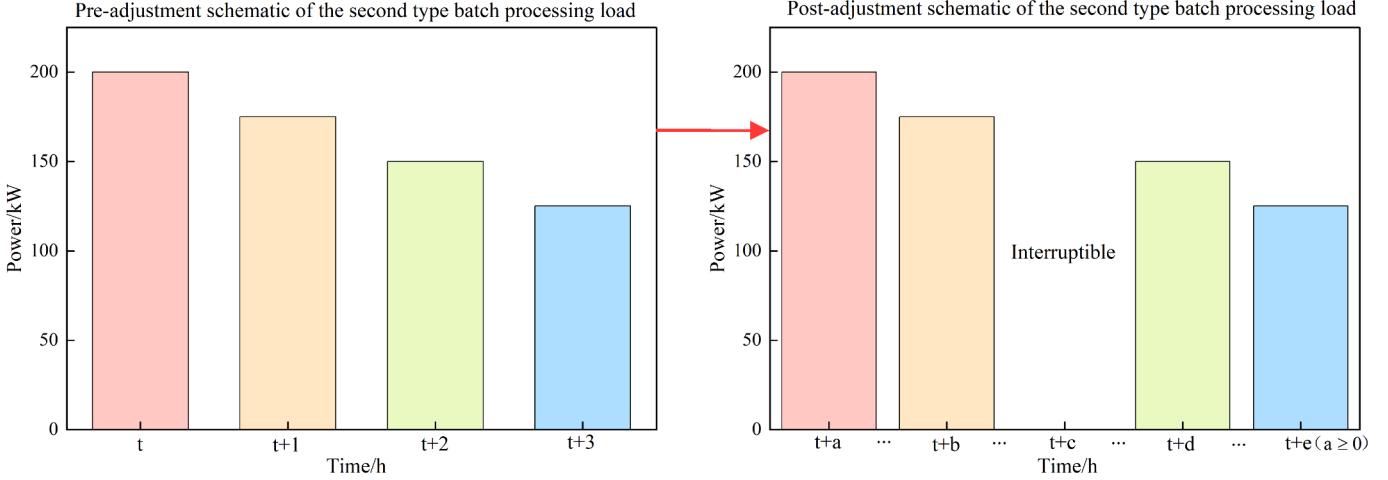


Fig. 4. Schematic diagram of the adjustability of the second type of batch processing load.

certain amount of the first category of batch processing load, leading to:

$$\mathbf{B} = \begin{pmatrix} b_{1,1} & 0 & 0 & \cdots & b_{23,23} & b_{24,24} \\ b_{1,1} & b_{2,2} & 0 & \cdots & 0 & b_{24,24} \\ b_{1,1} & b_{2,2} & b_{3,3} & \cdots & \vdots & 0 \\ 0 & b_{2,2} & \vdots & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \cdots & b_{23,23} & 0 \\ 0 & 0 & 0 & \cdots & b_{23,23} & b_{24,24} \end{pmatrix} \quad (4)$$

Due to the maximum delay processing time of two hours for the first category of batch processing loads, loads at hour 1 can only be moved to hours 2 or 3, and not to any other time. Thus,  $\mathbf{A}$  matrix is introduced to constrain the time-limited movement of the first category of batch processing loads, where the elements in row  $t$ , row  $t+1$ , and row  $t+2$  of each column  $t$  are 1, whereas all other elements are 0. Essentially, the elements in column  $t$  of matrix  $\mathbf{A}$  represent that the first category of batch processing load at time  $t$  before adjustment can be shifted to times  $t+1$  and  $t+2$ .

The structure of  $\mathbf{P}_{dy,py}$  is similar to that of  $\mathbf{A}$ , represented as follows:

$$\mathbf{P}_{dy,py} = \begin{pmatrix} c_{1,1} & 0 & 0 & \cdots & c_{1,23} & c_{1,24} \\ c_{2,1} & c_{2,2} & 0 & \cdots & 0 & c_{2,24} \\ c_{3,1} & c_{3,2} & c_{3,3} & \cdots & \vdots & 0 \\ 0 & c_{4,2} & \vdots & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \cdots & c_{23,23} & 0 \\ 0 & 0 & 0 & \cdots & c_{24,23} & c_{24,24} \end{pmatrix} \quad (5)$$

The purpose of introducing Equation (1) is to apply the upper and lower bounds to the size of each element within it, ensuring that the three elements in rows  $t$ ,  $t+1$ , and  $t+2$  of column  $t$  in the matrix  $\mathbf{P}_{dy,py}$  are not zero. Consequently, we have:

$$c_{t+n,t} \leq b_{t,t} \quad (t = 1, 2, \dots, 23, 24n = 0, 1, 2) \quad (6)$$

By shifting the first category of the batch processing load in row  $t$  of matrix  $\mathbf{P}_{dy,dzq}$  before adjustment to rows  $t+1$  and  $t+2$  after adjustment, the sum of column  $t$  in matrix  $\mathbf{P}_{dy,py}$  equals the size of the first category of the batch processing load at time  $t$  before the adjustment. Consequently, we have:

$$\mathbf{P}_{dy,zq} = \sum_{t_1=1}^{24} c_{t_1,t_2} = [b_{1,1}b_{2,2}\dots b_{23,23}b_{24,24}] \quad t_2 = 1, 2, \dots, 24 \quad (7)$$

where the subscripts  $t_1$  and  $t_2$  of  $c_{t_1,t_2}$  represent the row  $t_1$  and column  $t_2$  of the matrix  $\mathbf{P}_{dy,zq}$ , respectively;  $\mathbf{P}_{dy,zq}$  is the matrix of the first category of batch processing loads before adjustment, with dimensions of  $1 \times 24$ , where the element in the column  $t$  represents the size of the first category of batch processing load at time  $t$  before adjustment.

Given a maximum delay processing time of two hours for batch

processing loads, the batch processing load at time  $t$  after adjustment equals the total load shifted from times  $t-1$  and  $t-2$  to time  $t$  before adjustment plus any portion of the load at time  $t$  before adjustment that was not shifted. Essentially, the elements in row  $t$  of matrix  $\mathbf{A}$  represent the sum of the first category of batch processing load shifted to time  $t$  from times  $t-1$  and  $t-2$ , plus any load at time  $t$  that was not shifted after adjustment. Thus, summing the elements in row  $t$  of matrix  $\mathbf{P}_{dy,py}$  equals the size of the first category of the batch processing load at time  $t$  after adjustment. Consequently, we have:

$$\mathbf{P}_{dy,zh} = \sum_{t_2=1}^{24} c_{t_1,t_2} t_1 = 1, 2, \dots, 24 \quad (8)$$

where  $\mathbf{P}_{dy,zh}$  is the matrix of the first category of batch processing load after adjustment, with dimensions of  $24 \times 1$ , and the element in row  $t$  represents the size of the batch processing load at time  $t$  after the adjustment.

To reduce the system security risks and avoid the concentration of loads that could lead to new peak loads, we impose the following upper limit constraint:

$$\mathbf{P}_{dy,zh} \leq \mathbf{P}_{dy,zh\_up} \quad (9)$$

where the dimensions of  $\mathbf{P}_{dy,zh\_up}$  are  $24 \times 1$  and the element in row  $t$  represents the upper limit of the batch processing load size at time  $t$  after the adjustment.

Generally, if the maximum delay processing time for the first category of batch processing loads is  $e$  h, it is only necessary to transform matrix  $\mathbf{A}$  into matrix  $\bar{\mathbf{A}}$ . The transformations for the remaining formulas are similar, and  $\bar{\mathbf{A}}$  can be represented as

$$\bar{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 1_{e+1} & \vdots & \vdots & 1 \\ 1 & 1 & 0 & \cdots & 0 & 1_{e+1} & 1_e & \vdots \\ \vdots & 1 & 1 & \ddots & \vdots & 0 & 1_{e+1} & 1_e \\ 1_e & \vdots & 1 & \ddots & 0 & \vdots & 0 & 1_{e+1} \\ 1_{e+1} & 1_e & \vdots & \ddots & 1 & 0 & \vdots & 0 \\ 0 & 1_{e+1} & 1_e & \vdots & 1 & 1 & 0 & \vdots \\ \vdots & 0 & 1_{e+1} & \ddots & \vdots & 1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1_e & \vdots & 1 & 1 \end{pmatrix} \quad (10)$$

where the dimensions of the matrix  $\bar{\mathbf{A}}$  are  $24 \times 24$ , with the elements in row  $t$ , row  $t+1$ , up to row  $t+e$  of column  $t$ , all being 1. This implies that the first category of the batch processing load at time  $t$  before adjustment can be moved to times  $t+1$  to  $t+e$  after adjustment. In summary, the general modeling of the first category of batch processing loads was completed.

### 2.1.2. Second type batch processing load adjustability

The second category of batch processing load corresponds to large-scale tasks, such as massive data processing, which require several hours to complete. Subsequent tasks can be initiated only after completing the preceding tasks. The second type of batch processing load has lower real-time requirements, allowing data center operators to control it freely. It can be completed the following day at low cost. This study disregards the cost; its characteristics are illustrated in Fig. 4.

If the second category of batch processing load appears only from 2 AM to 7 AM before adjustment, and under the premise of maintaining an unchanged bandwidth and sequence, this portion of the load is shifted backward, with the specific modeling as follows:

First, we introduce the matrix  $P_{de,zq}$ , which represents the second category of the batch processing load from 2 AM to 7 AM before adjustment:

$$P_{de,zq} = [b_2, b_3, b_4, b_5, b_6, b_7] \quad (11)$$

where the dimensions of the matrix  $P_{de,zq}$  are  $1 \times 6$ , and  $b_t (t = 2, 3, \dots, 6, 7)$  represents the size of the second category of the batch processing load at time  $t$  before adjustment.

In addition, we introduce an auxiliary matrix  $P_{de,fz}$ , where all the elements are Boolean variables with dimensions of  $24 \times 6$ , specifically represented as

$$P_{de,fz} = \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,6} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,6} \\ d_{3,1} & \vdots & \cdots & d_{3,6} \\ \vdots & d_{23,2} & \ddots & \vdots \\ d_{24,1} & d_{24,2} & \cdots & d_{24,6} \end{pmatrix} \quad (12)$$

where the subscripts  $t_1$  and  $t_2$  of  $d_{t_1,t_2}$  represent the row  $t_1$  and column  $t_2$  of matrix  $P_{de,fz}$  respectively; Because the elements are Boolean variables,  $d_{t_1,t_2}$  equals 0 or 1.

To enforce the characteristic of unchanging bandwidth for the second category of batch processing loads, the load of each hour before adjustment can only be transferred to another specific hour after adjustment and cannot be distributed across multiple hours. Hence, a constraint is set on the matrix such that the sum of each column is 1, meaning that there is only one 1 in all elements of each column of the matrix  $P_{de,fz}$ , specifically represented as

$$\sum_{t_1=1}^{24} d_{t_1,t_2} = 1 \quad (13)$$

Next, it is necessary to determine the specific hours at which the second category of batch processing loads is transferred after the adjustment. Therefore, a time matrix  $T_{sk}$  is introduced with dimensions of  $1 \times 24$ , which is specifically represented as

$$T_{sk} = [1, 2, 3, 4, \dots, 23, 24] \quad (14)$$

where the element in the column  $t$  of matrix  $T_{sk}$  represents hour  $t$ .

Subsequently, we can determine the moments at which the second category of batch processing load is transferred after adjustment with the matrix  $T_{zysk}$ , specifically represented as

$$T_{zysk} = T_{sk} \cdot P_{de,fz} \quad (15)$$

where the dimensions of the matrix  $T_{zysk}$  are  $1 \times 6$ , and the element in column  $t$  signifies the time to which the second category of the batch processing load is transferred after adjustment. However, the initial processing start time for the second category of batch processing loads before adjustment is 2 AM, meaning that the earliest time these loads can be processed after adjustment must be greater than or equal to 2 AM and not 1 AM. Therefore, an additional constraint was required. That is,

$$d_{t_1,1} = 0 (t_1 = 1) \quad (16)$$

However, this type of batch processing load possesses an immutable temporal-order characteristic, implying that the batch processing load at

2 AM must be completed before processing the load at 3 AM (before adjustment). This sequence did not change after adjustment. Therefore, another constraint was established as follows:

$$T_{zysk}^t - T_{zysk}^{t-1} \geq 1 \quad (2 \leq t \leq 24) \quad (17)$$

where  $t$  represents the column  $t$  of matrix  $T_{zysk}$ . This implies that in matrix  $T_{zysk}$ , each subsequent column's element must be greater than that of the previous column, ensuring the immutable temporal order characteristic of the second category of batch processing load after adjustment.

In summary, the matrix  $P_{de,zh}$  for the second category of batch processing loads after adjustment can be represented as:

$$P_{de,zh} = P_{er,fz} \cdot P_{bb,zq}^T \quad (18)$$

where the dimensions of matrix  $P_{de,zh}$  are  $24 \times 1$ , containing 18 zeros and 6 non-zero elements. Superscript T indicates the transpose of a matrix. The second category of batch processing loads is successfully transferred while maintaining the same sequence and bandwidth. The row  $t$  in matrix  $P_{de,zh}$  corresponding to non-zero elements represent the specific hours  $t$  after adjustment for the second category of batch processing loads.

Generally, if the second category of batch processing loads appears between hours  $u$  and  $v$  before adjustment, it is only necessary to make appropriate adjustments to Equations (11)–(18). The specific expressions are as follows.

$$\left\{ \begin{array}{l} P_{de,zq} = [b_u, b_{u+1}, b_{u+2}, \dots, b_v] \\ P_{de,fz} = bv(24, v - u + 1) \\ \sum_{t_1=1}^{24} d_{t_1,t_2} = 1 (t_2 = 1, 2, \dots, v - u + 1) \\ d_{t_1,t_2} = 0 (t_1 = 1, 2, \dots, u - 1) \\ T_{sk} = [1, 2, \dots, u + 1, u + 2, \dots, 23, 24] \\ T_{zysk} = T_{sk} \cdot P_{de,fz} \\ T_{zysk}^t - T_{zysk}^{t-1} \geq 1 \quad (2 \leq t \leq 24) \\ P_{de,zh} = P_{de,fz} \cdot P_{de,zq}^T \end{array} \right. \quad (19)$$

In summary, the general modeling of the second category of batch processing loads was completed. If the interactive load matrix being adjusted is  $P_{jh}$  with dimensions of  $24 \times 1$ , the element in row  $t$  represents the size of the interactive load at time  $t$  before the adjustment. Therefore, the expression for the total data center load  $P_{data}$  at time  $t$  after adjustment is

$$P_{data} = P_{jh} + P_{dy,zh} + P_{de,zh} \quad (20)$$

where the dimensions of the matrix  $P_{data}$  are  $24 \times 1$ , where the element in row  $t$  represents the size of the total data center load at time  $t$  after adjustment.

### 2.2. Data center air thermal inertia modeling

The thermal inertia of air in data centers can also be considered a flexible resource. Without considering the air thermal inertia, all waste heat produced by the IT equipment would need to be absorbed by the cooling system. However, considering the air thermal inertia, fluctuations in the internal temperature of the data center can be maintained within tolerable limits, making the power adjustment of the cooling system more flexible. The thermal inertia of air is generally described by a first-order linear homogeneous differential equation [64]. After simplification, the following can be obtained:

$$\left\{ \begin{array}{l} (T_t^{\text{out}} - T_t^{\text{in}})H^{\text{DC}}S^{\text{DC}} + Q_{\text{ws}}^t - Q_{\text{cs}}^t = \rho_a C_a V^{\text{DC}} \delta_t^{\text{in}} \\ T_t^{\text{in}} - T_0^{\text{in}} = \delta_t^{\text{in}} \Delta t, t = 1 \\ T_t^{\text{in}} - T_{t-1}^{\text{in}} = \delta_t^{\text{in}} \Delta t, 2 \leq t \leq 24 \\ Q_{\text{cs}}^t \geq 0 \\ T_{\text{min}}^{\text{in}} \leq T_t^{\text{in}} \leq T_{\text{max}}^{\text{in}} \\ \delta_{\text{min}}^{\text{in}} \leq \delta_t^{\text{in}} \leq \delta_{\text{max}}^{\text{in}} \end{array} \right. \quad (21)$$

where the subscript  $t$  represents the moment at time  $t$ , which will not be reiterated further;  $T_t^{\text{out}}$ ,  $T_t^{\text{in}}$  and  $\delta_t^{\text{in}}$  represent the outdoor temperature, indoor temperature, and rate of change of the indoor temperature of the data center, respectively.  $H^{\text{DC}}$ ,  $S^{\text{DC}}$  and  $V^{\text{DC}}$  represent the heat dissipation coefficient, heat exchange area, and volume of the data center, respectively;  $Q_{\text{ws}}^t$  and  $Q_{\text{cs}}^t$  represent the waste heat power generated by the data center and the required cooling power, respectively;  $T_{\text{max}}^{\text{in}}$  and  $T_{\text{min}}^{\text{in}}$  denote the upper and lower limits of the indoor temperature, respectively;  $\delta_{\text{max}}^{\text{in}}$  and  $\delta_{\text{min}}^{\text{in}}$  denote the upper and lower limits of the rate of change in indoor temperature, respectively.

### 2.3. Data center waste heat recovery system modeling

For the cooling system of the data center, the following expression is given:

$$Q_{\text{cs}}^t = P_{\text{cs}}^t \eta_{\text{cs}} \quad (22)$$

where  $\eta_{\text{cs}}$  is the coefficient of performance of the cooling system's electric refrigeration and  $P_{\text{cs}}^t$  represents the cooling system's refrigeration power consumption.

Data center IT equipment generate a significant amount of waste heat  $Q_{\text{ws}}^t$  during operation. The specific expressions are as follows:

$$Q_{\text{ws}}^t = (\Upsilon - 1)P_{\text{data}}^t \eta_{\text{cs}} \quad (23)$$

where  $\Upsilon$  represents the power efficiency of the data center. Using heat pumps to recover this heat, it is possible to reduce the demand for new energy sources and enhance the overall energy utilization efficiency. The specific expressions are as follows:

$$Q_{\text{hp}}^t = P_{\text{hp}}^t \eta_{\text{hp}} \quad (24)$$

$$0 \leq Q_{\text{hp}}^t \leq \eta_{\text{re}} Q_{\text{cs}}^t \quad (25)$$

where  $P_{\text{hp}}^t$  and  $Q_{\text{hp}}^t$  represent the power consumption and heat production power of the air-source heat pump, respectively, and  $\eta_{\text{hp}}$  represents the conversion efficiency of electricity to heat.

### 3. Scenario generation method based on conditional least squares generative adversarial networks

As a renewable energy source, wind power offers significant energy-saving and emission-reduction advantages in data center microgrids. However, their intermittency and instability pose numerous challenges to the stable operation of microgrids. For instance, the variability in the wind power output requires microgrids to have flexible scheduling strategies to ensure the reliability and stability of the power supply. In addition, the uncertainty of wind power output affects the configuration strategy of the system. Therefore, improving the accurate characterization of wind power output uncertainty is particularly important. This study circumvents the complexities and limitations of traditional statistical models by employing conditional least squares generative adversarial networks (C-LSGANs) to generate scenario sets for renewable energy outputs. This approach, grounded in a deep learning framework, delves into the high-dimensional nonlinear characteristics of historical data on renewable energy output by predicting the complex

mapping relationship between the noise distribution under forecast information and a set of scenarios. Furthermore, this study establishes a comprehensive evaluation index system aimed at precisely comparing and analyzing the quality of the generated scenarios, ensuring that the data accurately reflect the uncertainty of the wind power output, thereby providing reliable support for the optimized operation and capacity configuration of integrated energy systems and microgrids.

#### 3.1. GANs method

Generative adversarial networks (GANs), proposed by Goodfellow in 2014, represent a revolutionary unsupervised learning model [65]. This model consists of two independently operating deep learning networks: a generator and a discriminator. The core function of the generator is to grasp and simulate the hidden distribution in historical data and transform random noise into samples with practical significance. The discriminator's task is to accurately identify the authenticity of the input samples. Through the continuous adversarial and learning processes of these two networks, the model could significantly enhance the authenticity of the generated samples.

For clarity, let  $x$  represent real data samples,  $p_{\text{data}}(x)$  denote the probability distribution of real data,  $z$  represent random noise data, and  $p_z(z)$  represent the probability distribution of noisy data. During the training phase, a batch of samples from  $z \sim p_z(z)$  is introduced into the generator, thereby producing simulated samples  $G(z)$  that follow the probability distribution  $p_g(z)$ . The training objective of the generator was to adjust  $p_g(z)$  to closely fit the real data distribution  $p_{\text{data}}(x)$ . The training objective of the discriminator was to identify accurately whether the input data are from a real sample set or manufactured by the generative model. Through this adversarial training process, the generator and discriminator eventually reach a Nash equilibrium state, in which the discriminator cannot effectively distinguish whether its inputs are real data samples or samples produced by the generator. Upon reaching this state, it can be inferred that the generator mastered the underlying probability distribution mechanism of the real samples.

Based on the training objectives of the generator and discriminator, the loss functions  $L_G$  for the generator and  $L_D$  for the discriminator are defined as follows.

$$L_G = \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \quad (26)$$

$$L_D = -\mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] - \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \quad (27)$$

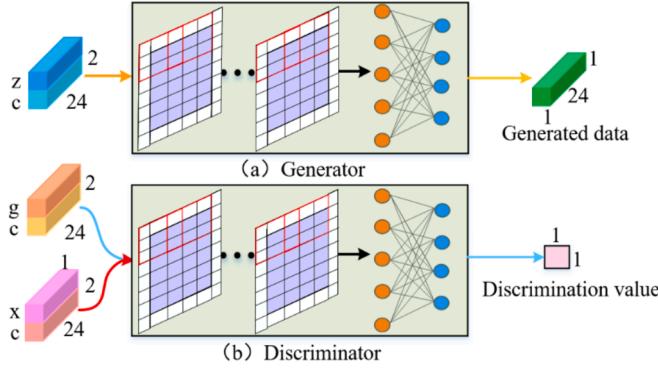
where  $\mathbb{E}$  represents the expected value of the batch samples.

Based on the aforementioned expressions, the objective function of GANs is as shown in Equation (28).

$$\min_{G} \max_{D} V_{\text{GANs}}(D, G) = \mathbb{E}_{x \sim p_{\text{data}}} [\log(D(x))] + \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))] \quad (28)$$

#### 3.2. C-lsgans method

In the traditional GANs framework, the loss function of the discriminator essentially corresponds to the computation of Jensen-Shannon (JS) divergence [66]. When the distribution of the generated samples did not fully overlap with that of the real samples across all dimensions, the JS divergence tended towards a fixed value. At this point, the JS divergence reaches a saturation state, causing the discriminator loss to approach zero. Owing to the use of cross-entropy as the loss function, traditional GANs face the challenge of vanishing gradients during the execution of the backpropagation algorithm, exacerbating the difficulty of network training, and potentially leading to mode collapse. To address these issues, least-squares generative adversarial networks (LSGANs) employing coding and least-squares loss functions were proposed [67]. The objective function of LSGANs is as follows:



**Fig. 5.** Generator and discriminator network structure of C-LSGANs.

$$\begin{cases} \min_D V_{\text{LSGANs}}(D) = \frac{1}{2} \mathbb{E}_{x \sim P_{\text{data}}} [(D(x) - b)^2] + \frac{1}{2} \mathbb{E}_{z \sim P_2} [(D(G(z)) - a)^2] \\ \min_G V_{\text{LSGANs}}(G) = \frac{1}{2} \mathbb{E}_{z \sim P_2} [(D(G(z)) - c)^2] \end{cases} \quad (29)$$

where  $a$  and  $b$  represent the labels for the generated and real data, respectively, and  $c$  denotes the confidence value that the generator desires the discriminator attribute to the generated data.

In LSGANs, both the generator and discriminator are optimized using a least-squares loss function. By penalizing samples that deviate from the decision boundary, this strategy provides a network with enhanced gradient signals, effectively mitigating the gradient-vanishing problem encountered in traditional GANs. Further, this loss function is extremely beneficial for improving the quality of generated samples. Penalizing unrealistic samples and “pushing” them towards the decision boundary encourages the generated samples to simulate the real data distribution more accurately.

To generate data under specific events, conditional information  $c$  is incorporated into the generator and discriminator of LSGANs. Consequently, the objective function of conditional LSGANs (C-LSGANs) is constructed as

$$\begin{cases} \min_D V_{\text{LSGANs}}(D) = \frac{1}{2} \mathbb{E}_{x \sim P_{\text{data}}} [(D(x|c) - 1)^2] + \frac{1}{2} \mathbb{E}_{z \sim P_2} [(D(G(z|c)))^2] \\ \min_G V_{\text{LSGANs}}(G) = \frac{1}{2} \mathbb{E}_{z \sim P_2} [(D(G(z|c)) - 1)^2] \end{cases} \quad (30)$$

In this study, renewable energy forecast data were introduced as conditional information, and both forecast and real data were inputted into the C-LSGANs for training. Through this training process, the generator of the C-LSGAN can grasp the transformation relationship between the noise distribution and the target output distribution under the given conditional information. To maximize the use of renewable energy forecast information, this study meticulously designed a deep neural network architecture for the generator and discriminator in C-LSGANs, ensuring that the network can effectively learn and generate data that satisfy conditional constraints, further enhancing the model's application value and performance. The architecture is illustrated in Fig. 5.

### 3.3. Scenario evaluation metrics

To ensure that the generated scenario sets reflected the uncertainty of renewable energy sources to the greatest extent possible, two key performance indicators were used for quality assessment: coverage [68] and envelope area [69]. The coverage aims to quantify the extent to which the generated scenario sets reflect the actual range of renewable energy generation capabilities, whereas the envelope area evaluates the effectiveness of the scenario sets in covering the actual generation data

with the aim of minimizing the geometric envelope area. It is important to emphasize that these two evaluation metrics are not used in isolation, but are combined to achieve a more comprehensive performance assessment. Specifically, greater coverage and a smaller envelope area indicate the superior performance of the proposed method. The specific calculation methods for these two indicators are as follows.

$$\left\{ \begin{array}{l} \text{index1} = \frac{1}{N} \sum_{n=1}^N B_n \\ B_n = \begin{cases} 1, & \text{if } P_n^{\text{real}} \in [\min(P_n^{\text{gene}}), \max(P_n^{\text{gene}})] \\ 0, & \text{otherwise} \end{cases} \end{array} \right. \quad (31)$$

$$\text{index2} = \frac{1}{N} \sum_{n=1}^N (\max(P_n^{\text{gene}}) - \min(P_n^{\text{gene}})) \quad (32)$$

where  $\text{index1}$  represents the coverage,  $\text{index2}$  denotes the envelope area,  $N$  signifies the number of time periods,  $P_n^{\text{gene}}$  represents the output of the generated scenario, and  $P_n^{\text{real}}$  denotes the actual output.

For the data center, conventional electrical, and thermal loads, a spatiotemporal clustering algorithm was employed to perform scenario clustering and reduction, as detailed in [70].

## 4. Data center microgrid capacity configuration stochastic optimization method

Data center microgrids often involve complex energy management issues, including, but not limited to, fluctuating electricity and cooling demands, variability in renewable energy supply, and energy costs. These factors pose significant challenges to the capacity configuration of data center microgrids. This study focuses on thermal energy management; hence, configuring cooling systems, electric boilers, and air source heat pumps without capacity configuration for other equipment.

### 4.1. Objective function

An objective function is formulated to minimize the sum of the investment and operational maintenance costs within the planning period as follows:

$$f = \min(f_1 + f_2) \quad (33)$$

where  $f$  represents the total cost, and  $f_1$  and  $f_2$  denote the annual investment cost and annual operational maintenance cost for data center operators, respectively.

#### 4.1.1. Investment cost

The specific expression for the annual investment cost is as follows:

$$f_1 = R_n \sum_{n=1}^N \alpha_n S_n \quad (34)$$

where  $S_n$  and  $\alpha_n$  represent the capacity size and per-unit capacity investment cost of the  $n$ th device to be configured, respectively, and  $R_n$  denotes the investment recovery factor for the  $n$ th device, with the specific expression as follows:

$$R = \frac{r(1+r)^s}{(1+r)^s - 1} \quad (35)$$

where  $r$  is the discount rate and  $s$  is the number of discount years. The final expression for the capacity size of each configured device is

$$S_n = \arg(\min[f_1 + f_2]) \quad (36)$$

#### 4.1.2. Operational and maintenance costs

Operational and maintenance costs are generally due to equipment wear, tear, and manual maintenance, which are related to the real-time

operating power of the equipment. In addition, the stochastic optimization model considers several typical scenarios; hence, the mathematical expectation is used to calculate the average of the typical scenario set. The specific expressions are as follows:

$$f_2 = M \sum_{i=1}^K k_i \sum_{t=1}^{24} (O_{gt}^t + O_{grid}^t + O_{wur}^t + O_{bu}^t + O_{eb}^t + O_{cs}^t + O_{hp}^t) \quad (37)$$

where  $M$  represents the number of days in a year, taken as 365 in this study;  $K$  denotes the number of scenarios generated by the scenario generation method based on C-LSGANs proposed in this paper;  $k_i$  is the probability corresponding to the  $i$ th typical day;  $t$  stands for the time moment with the same notation used below.  $O_{gt}^t, O_{grid}^t, O_{wur}^t, O_{bu}^t, O_{eb}^t, O_{cs}^t$  and  $O_{hp}^t$  represent the gas consumption cost of the gas turbines, interaction cost with the upper-level grid, penalty cost for wind curtailment, compensation cost for the first type of batch processing load, and maintenance costs of electric boilers, cooling systems, and air-source heat pumps, respectively. The expression for each cost is as follows:

1 ) Gas consumption cost of gas turbines  $O_{gt}^t$

$$O_{gt}^t = a(P_{gt}^t)^2 + bP_{gt}^t + c \quad (38)$$

where  $P_{gt}^t$  represents the power output of the gas turbine and  $a, b, c$  denote the cost coefficients of the gas turbine.

2 ) Interaction cost with the upper-level grid  $O_{grid}^t$

$$O_{grid}^t = P_{buy}^t C_{buy}^t - P_{sell}^t C_{sell}^t \quad (39)$$

where  $C_{buy}^t$  and  $C_{sell}^t$  denote the time-of-use electricity price and the feed-in tariff, respectively;  $P_{buy}^t, P_{sell}^t$  represent the power purchased from and sold to the grid, respectively.

3 ) Penalty cost for wind curtailment  $O_{wur}^t$

$$O_{wur}^t = C_{cur} P_{windcur}^t \quad (40)$$

where  $C_{cur}$  represents the unit cost of the wind curtailment penalty and  $P_{windcur}^t$  denotes the power of wind curtailment.

4 ) Compensation cost for batch processing load  $O_{bu}^t$

$$O_{bu}^t = C_{bc} P_{dyzh}^t \quad (41)$$

where  $C_{bc}$  represents the unit compensation cost for batch processing loads and its constraints are shown in Equations (1)–(18).

5 ) Operational and maintenance cost of electric boilers  $O_{eb}^t$

$$O_{eb}^t = C_{eb} Q_{eb}^t \quad (42)$$

where  $C_{eb}$  represents the unit maintenance cost of electric boilers and  $Q_{eb}^t$  denotes the heating power of electric boilers.

6 ) Operational and maintenance cost of cooling systems  $O_{cs}^t$

$$O_{cs}^t = C_{cs} Q_{cs}^t \quad (43)$$

where  $C_{cs}$  represents the unit maintenance cost and  $Q_{cs}^t$  denotes the cooling power of the cooling systems.

7 ) Operational and maintenance cost of air source heat pumps  $O_{hp}^t$

$$O_{hp}^t = C_{hp} Q_{hp}^t \quad (44)$$

where  $C_{hp}$  represents the unit maintenance cost of the cooling systems and  $Q_{hp}^t$  denotes the heating power of air-source heat pumps.

#### 4.2. Constraint conditions

The constraints for each equipment and power balance are as follows:

1 ) Gas turbine output constraint

$$\begin{cases} P_{gt}^{\min} \leq P_{gt}^t \leq P_{gt}^{\max} \\ -R_{gt} \leq P_{gt}^t - P_{gt}^{t-1} \leq R_{gt} \end{cases} \quad (45)$$

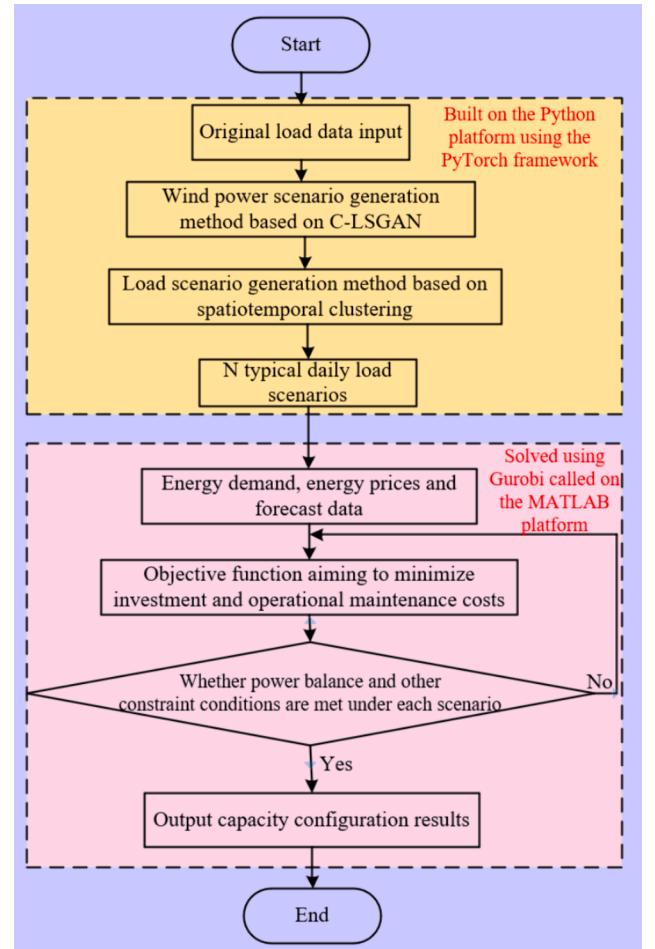


Fig. 6. Model solving process.

where  $P_{gt}^{\max}, P_{gt}^{\min}$  represents the upper and lower limits of the gas turbine output power, respectively, and  $R_{gt}$  is the ramp rate of the gas turbine.

$$Q_{gt}^t = \lambda P_{gt}^t \quad (46)$$

where  $Q_{gt}^t$  is the heat production power of the gas turbine and  $\lambda$  is the coefficient ratio of electric power production to heat production power of the gas turbine.

2 ) Power interchange constraint with the upper-level grid

$$\begin{cases} 0 \leq P_{buy}^t \leq P_{buy}^{\max} \\ 0 \leq P_{sell}^t \leq P_{sell}^{\max} \end{cases} \quad (47)$$

where  $P_{buy}^{\max}, P_{sell}^{\max}$  are the power upper limit values for buying and selling electricity between the microgrid and upper-level grid, respectively.

3 ) Wind power constraint

$$\begin{cases} 0 \leq P_{windcur}^t \leq P_{windyc}^t \\ 0 \leq P_{wind}^t \leq P_{windyc}^t \\ P_{windcur}^t + P_{wind}^t = P_{windyc}^t \end{cases} \quad (48)$$

where  $P_{windyc}^t, P_{wind}^t$  represent the forecasted power and actual power output of wind power, respectively.

4 ) Electric boiler constraint

$$Q_{eb}^t = P_{eb}^t \eta_{eb} \quad (49)$$

where  $P_{eb}^t$  is the power consumption of the electric boiler;  $\eta_{eb}$  is the

**Table 2**

Comparison of indicators by different methods.

Method	C-LSGANs	C-WGAN-GP	Monte-Carlo
Coverage rate (%)	92.39	90.69	92.09
Envelope area (%)	0.4301	0.6689	0.5493

electric-to-heat conversion efficiency of the electric boiler.

5 ) Equipment capacity upper limit constraint

$$0 \leq S_n \leq S_n^{\max} \quad (50)$$

where  $S_n^{\max}$  represents the upper limit of the capacity of the n-th equipment to be configured.

6 ) Power balance constraint

$$\begin{aligned} P_{\text{buy}}^t + P_{\text{gt}}^t + P_{\text{wind}}^t &= P_{\text{sell}}^t + P_{\text{load}}^t + P_{\text{data}}^t + P_{\text{eb}}^t + P_{\text{hp}}^t + P_{\text{cs}}^t \\ Q_{\text{gt}}^t + Q_{\text{eb}}^t + Q_{\text{hp}}^t &= Q_{\text{load}}^t \end{aligned} \quad (51)$$

The remaining constraints are discussed in Section 2. In summary, a flowchart for the capacity configuration method of the data center microgrid thermal energy management equipment, considering the flexible resources of the data center, is shown in Fig. 6.

## 5. Case study analysis

Based on the model and constraints described above, this study used a provincial data center microgrid as an example to configure its capacity to minimize total costs. The specific case setup and analysis of the results are detailed below.

### 5.1. Basic data

Using the load curve of 2020 as a sample with a data collection interval of 1 h, this study employs the proposed C-LSGAN method to generate five typical wind power scenarios. The neural-network structure and training pseudocode for this method are presented in Appendix Figure A1 and Table A1, respectively. The spatiotemporal clustering algorithm proposed in [70] is used to generate five typical scenarios for data center load, conventional electrical load, and thermal load, with the final five typical daily curves presented in Appendix B; Time-of-use electricity prices are listed in Appendix C, Table C1; Economic parameters of relevant equipment and other parameters are found in Appendix C, Table C2.

### 5.2. Comparison of scenario generation methods

To prove the effectiveness and superiority of the proposed wind power scenario generation technique, this study compared and analyzed it against the advanced C-WGAN-GP and the classical Monte Carlo method as control groups. Based on 365 days of historical wind power data, the evaluation was conducted using two key performance indicators: coverage and envelope area, with the results summarized in Table 2. The analysis shows that the wind power data generation strategy using the C-LSGAN technology significantly outperforms the C-WGAN-GP and Monte Carlo methods in terms of coverage, with increases of 1.70 % and 0.30 %, respectively. In terms of the envelope area indicator, C-LSGANs also demonstrated superiority, reducing it by 23.88 % and 11.92 % compared with the C-WGAN-GP and Monte Carlo methods, respectively. This confirms the efficiency and accuracy of the C-LSGAN strategy in generating sets of wind-power scenarios.

### 5.3. Simulation result analysis

#### 5.3.1. Operational scenario settings

A comparative analysis is conducted by establishing six operational scenarios to verify the effectiveness of the proposed method. The

**Table 3**

Comparison of different scenarios.

Scenario	1	2	3	4	5	6
Considering waste heat recovery	✗	✓	✗	✗	✗	✓
Considering air thermal inertia	✗	✗	✓	✗	✗	✓
Considering the adjustability of the first type of batch processing load	✗	✗	✗	✓	✗	✓
Considering the adjustability of the second type of batch processing load	✗	✗	✗	✗	✓	✓

settings of these scenarios are listed in Table 3, with Scenario 6 representing the method proposed in this study.

#### 5.3.2. Capacity configuration result analysis

The configuration results under the six operational scenarios are presented in Table 4 and Fig. 7, and a quantitative analysis is provided below.

Scenario 2 considers waste heat recovery from the data center, resulting in a 5.17 % reduction in investment costs, a 3.0 % reduction in operational costs, and a 3.12 % reduction in total cost. The capacity of the electric boiler decreased by 70.30 % and the wind curtailment rate decreased by 1.96 %, highlighting the significant role of waste heat recovery. However, the increased electricity consumption from installing heat pumps leads to higher electricity purchase costs. The thermal power balance charts for scenarios 1 and 2 on a typical Day 3 are shown in Figs. 8 and 9, respectively. After considering the thermal power recovered from the data center, the thermal power generated by the gas turbines and electric boilers was reduced accordingly, thereby lowering the operational cost of the overall system.

Scenario 3 considers the thermal inertia of the data center air, resulting in a 6.68 % decrease in investment costs, a 0.28 % decrease in operational costs, and a reduction of 0.63 % in total cost. The capacities of the electric boiler and cooling system decreased by 3.7 and 8.0 %, respectively, and the wind curtailment rate decreased by 1.17 %. The indoor and outdoor temperature changes on a typical Day 2 are shown in Fig. 10. The indoor temperature can be adjusted within a wider range with smaller fluctuations, demonstrating the advantages of considering air thermal inertia.

Scenario 4 considers the adjustable capability of the first type of batch processing load, with investment costs reduced by 28.0 %, operational and maintenance costs reduced by 0.84 %, and total costs reduced by 2.33 %. The capacities of the electric boilers and cooling systems were reduced by 2.39 and 39.5 %, respectively. Considering Day 5 as an example, the load distribution of the first type of batch processing load before and after the adjustment is shown in Fig. 11. It was found that the first type of batch processing load significantly decreased at 10, 11, 21, and 22 h, and significantly increased at 12, 13, 23, and 24 h, concentrating the response in periods of high to low electricity prices under the time-shiftable mode.

Scenario 5 considers the adjustable capability of the second type of batch processing load, with investment costs reduced by 22.3 %, operational and maintenance costs reduced by 1.51 %, and total costs reduced by 2.60 %. The capacity of the electric boilers increased by 11.6 %, whereas the capacity of the cooling systems decreased by 38.0 %, and the wind curtailment rate decreased by 4.86 %. Considering Day 5 as an example, the load distribution of the second type of batch processing load before and after the adjustment is shown in Fig. 12. It was observed that after considering its adjustable capability, the second type of batch processing load was sequentially shifted to 3, 5, and 21–24 h, influenced by electricity prices. Importantly, these time periods had sufficient wind power, and moving the second type of batch processing load to these times benefits wind power consumption.

Scenario 6 represents the strategy proposed in this study, considering the waste heat recovery of the data center, thermal inertia of the air, and adjustability of the two types of batch processing loads. The final

**Table 4**

Comparison of configuration results under different scenarios.

Scenario	Natural gas cost/\$	Electricity purchase cost/\$	Investment cost/\$	Operating cost/\$	Wind curtailment rate/%	Total cost/\$	Electric boiler/kW	Cooling system/kW	Heat pump/kW
1	331,811	186,649	30,163	518,461	9.42 %	548,623	670	1232	0
2	329,544	201,975	28,602	502,918	7.46 %	531,520	199	1232	223
3	359,809	185,345	28,148	517,006	8.25 %	545,155	645	1133	0
4	329,650	174,142	21,732	514,088	7.21 %	535,819	654	745	0
5	344,612	165,226	23,428	510,599	4.56 %	534,381	748	764	0
6	325,588	167,706	20,145	495,129	2.24 %	515,274	483	534	188

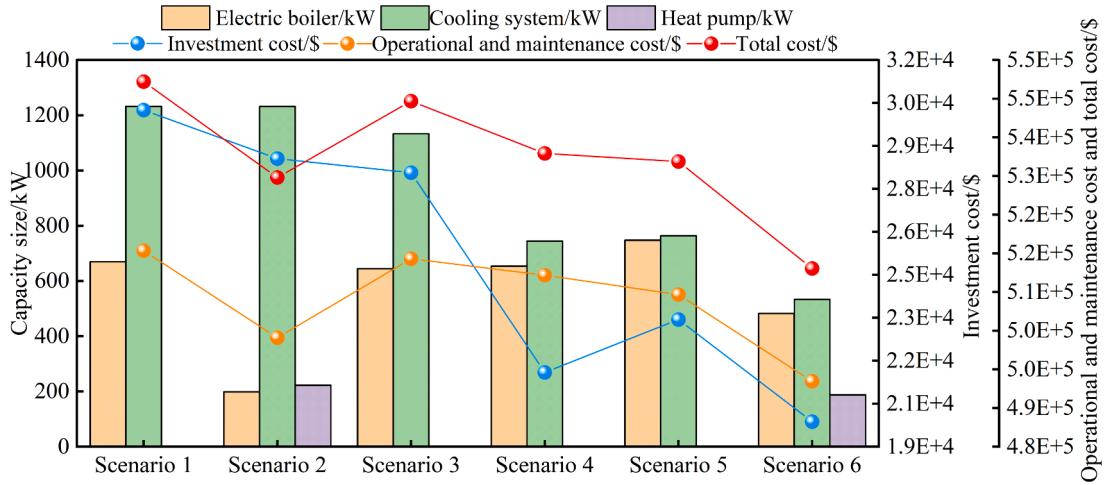


Fig. 7. Comparison of configuration results under different scenarios.

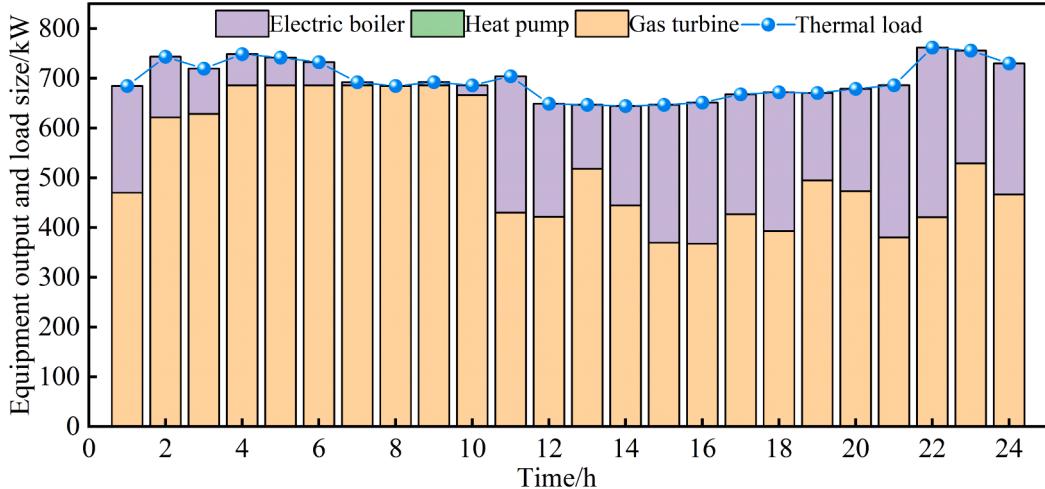


Fig. 8. Heat power balance chart under Scenario 1 for a typical Day 3.

investment cost was reduced by 33.2 %, operating cost was reduced by 4.5 %, and total cost was reduced by 6.08 %. The capacities of the electric boilers and cooling systems are reduced by 27.9 % and 56.7 %, respectively, and the wind curtailment rate was reduced by 7.18 %, demonstrating the effectiveness of the proposed strategy.

#### 5.4. Sensitivity analysis

This section conducts an analysis based on Scenario 6.

##### 5.4.1. Impact of the first type batch processing load proportion change on system configuration results

By leveraging historical operational data, data center operators can

ascertain the relationship between task volume and electrical energy consumption. During data center operation, when a user submits a task associated with the first type of batch processing load, the data center operator estimates the task volume and calculates the corresponding energy consumption, providing a certain incentive to the user. This incentive offered per unit of electricity is referred to as price compensation. When data center users receive a certain level of price compensation, there is an increase in the proportion of interactive loads, reported as batch processing loads. As this proportion and its corresponding variable increase, the data center can schedule the first type of batch processing load more flexibly to reduce operating costs. The relationship between the proportion of the first type of batch processing load and the compensation price is illustrated in Fig. 13, with the

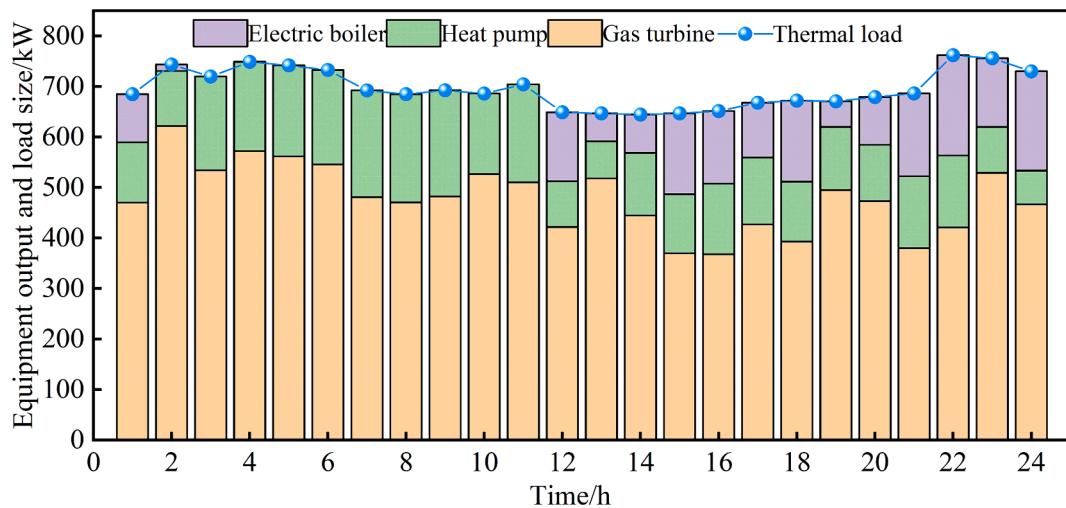


Fig. 9. Heat power balance chart under Scenario 2 for a typical Day 3.

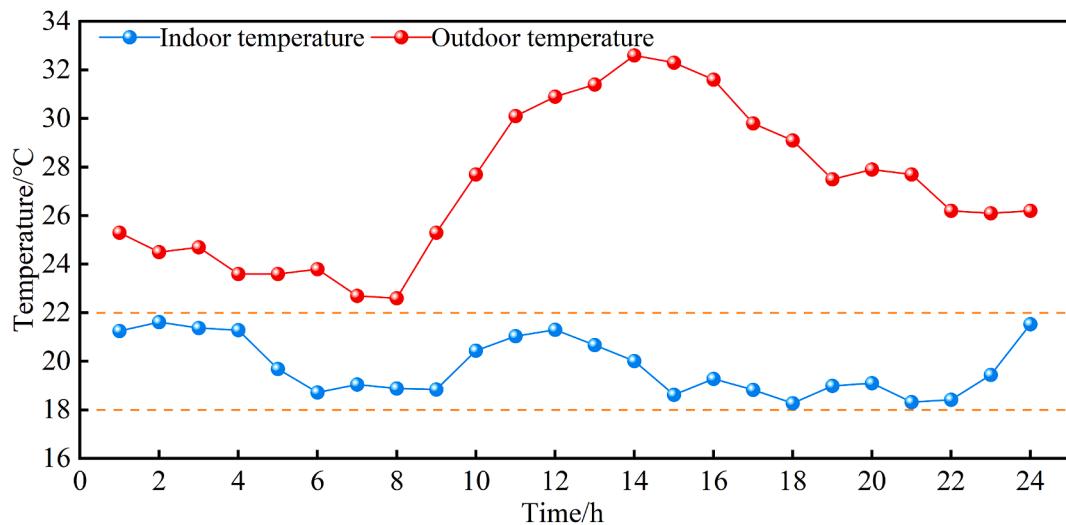


Fig. 10. Indoor and outdoor temperature change chart under Scenario 3 for a typical Day 2.

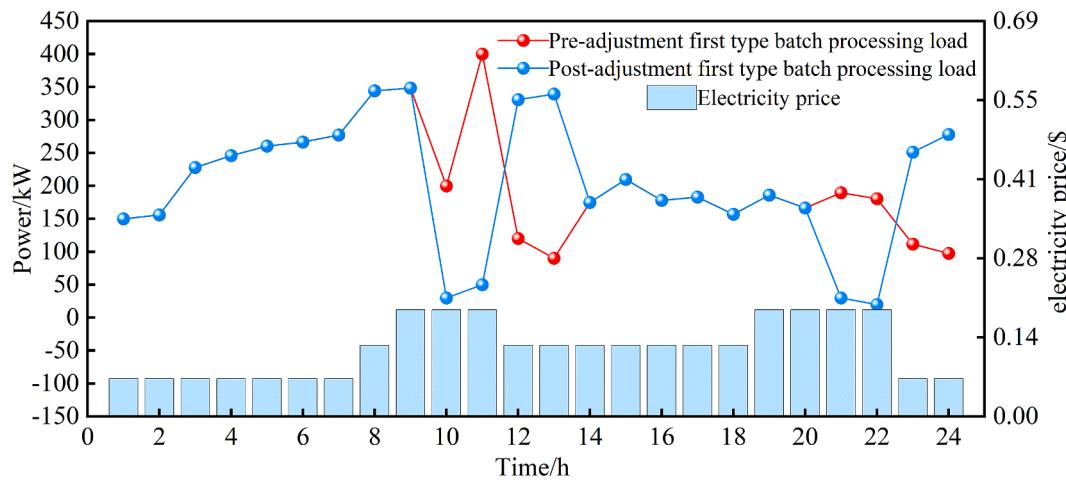
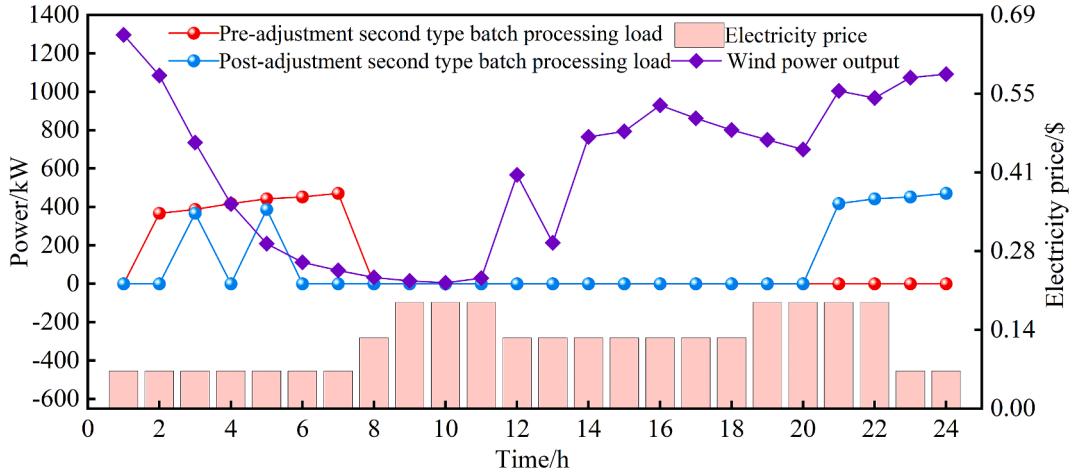
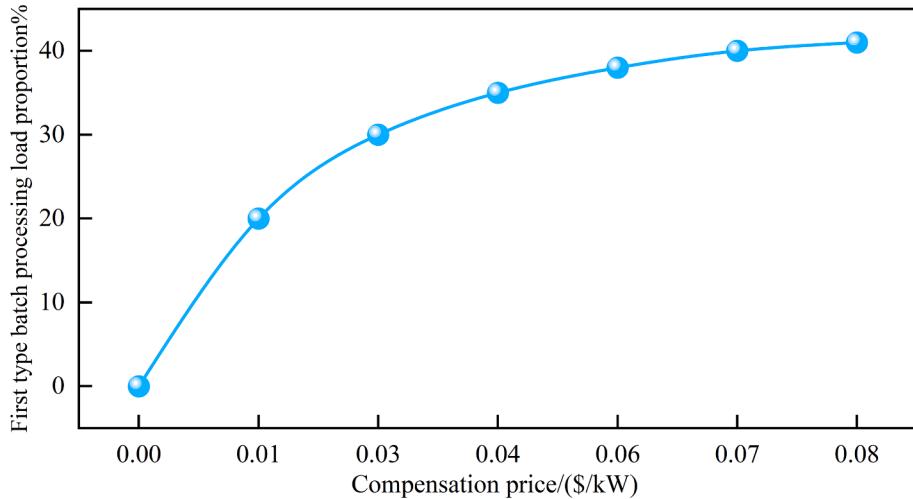


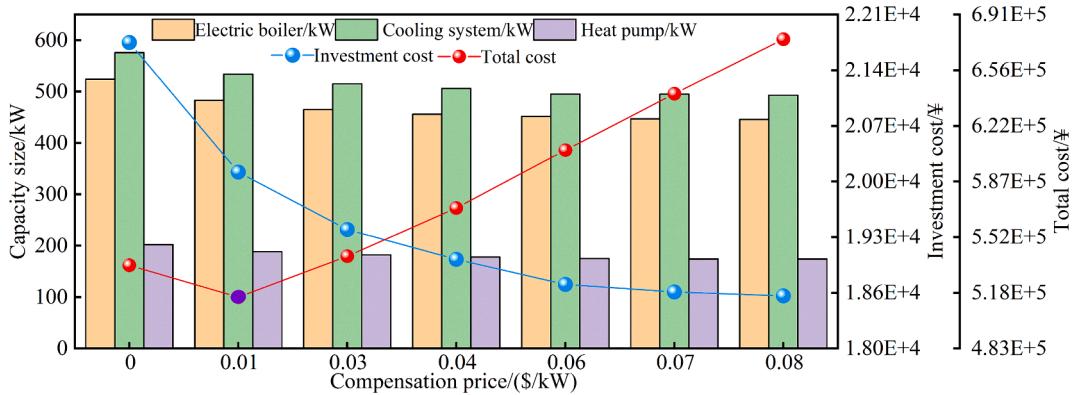
Fig. 11. Pre- and post-adjustment change chart of the first type of batch processing load under Scenario 4 for a typical Day 5.



**Fig. 12.** Pre- and post-adjustment change chart of the second type of batch processing load under Scenario 5 for a typical Day 5.



**Fig 13.** Relationship between compensation price and the proportion of the first type of batch processing load.



**Fig. 14.** Relationship between compensation price and planning results.

detailed proof found in Reference [71].

We further analyzed the relationship between the compensation price and planning results, as shown in Fig. 14 and Table 5. The higher the compensation price, the larger the response volume of the first type of batch processing load, and the more evident the effect of peak shaving and valley filling, thus reducing the upper limit of the output required by the system units. Consequently, the capacity configuration of each

device was smaller, leading to reduced investment costs. It was also observed that with an increase in the compensation price, the rate of decrease in investment costs gradually diminished; specifically, the reduction rates were 8.2, 3.8, 2.0, 1.2, 0.81, and 0.36 %. Because the total cost equals the investment cost plus operational and maintenance costs, when a certain incentive is provided to users, the user demand response significantly reduces operational and maintenance costs,

**Table 5**

Relationship between compensation price and planning results.

Compensation price/\$	Investment cost/\$	Total cost/\$	Electric boiler/kW	Cooling system/kW	Heat pump/kW
0	21,754	535,045	524	576	202
0.01	20,145	515,274	483	534	188
0.03	19,433	540,641	465	515	182
0.04	19,062	570,622	456	506	178
0.06	18,751	606,543	452	495	175
0.07	18,657	641,524	447	495	174
0.08	18,610	675,357	446	493	174

thereby initially decreasing the total cost. Compared to a compensation price of 0.01 \$/kW, that of 0.01 \$/kW reduced the total cost by 3.8 %. However, as the compensation price increased, the total cost showed an upward trend with increases of 4.7, 5.2, 5.9, 5.5, and 5.0 %. This is because the time-shiftable characteristic of the first type of batch processing load implies that the real economic benefits can only be generated during periods when electricity prices decrease from high to low, with limited economic benefits generated. However, the larger the compensation price, the more the compensation cost that needs to be paid; therefore, the total cost first decreases and then increases. This suggests that data center operators must set a reasonable proportion of the first type of batch processing load and compensation price to reduce their costs.

#### 5.4.2. Impact of different allowable delays of the first type batch processing load on costs

In this section, we assume that the proportion of the first type of batch processing load remains unchanged at 20 % and explore the impact of different delay times on the planning results. The simulation results, as shown in Fig. 15 and Table 6, indicate that both the investment cost and total cost decrease as the delay of the first type of batch processing load increases. Over the delay range of 0 to 12 h, the reductions in investment costs were 7.5, 2.4, 1.5, 0.84, 0.64, and 0.38 %, respectively, whereas the reductions in total costs were 3.8, 3.3, 2.4, 1.3, 0.74, and 0.37 %, respectively. This is because an increase in delay offers the system more flexibility; however, the magnitude of the cost reduction becomes progressively smaller owing to the fixed amount of transferrable load, which diminishes in variability with the increased delay.

#### 5.4.3. Impact of different delays and compensation prices of the first type batch processing load on planning results

This section discusses the effects of different delays and compensation prices on planning results. The investment cost simulation results are displayed in Fig. 16 and Table 7; for instance, at a 4 h delay, as the compensation price increases, the reduction rates of investment costs are

3.9, 1.5, 1.1, and 0.69 %, respectively. Similarly, with a compensation price of 0.04 \$/kW, as the delay increases, the investment cost reduction rates are 2.57, 1.45, 1.2, 0.43, and 0.37 % respectively; the trend of these results can be explained by sections 5.4.1 and 5.4.2, which are not reiterated here.

The simulation results for the total cost are shown in Fig. 17 and Table 8, indicating that for the same compensation price, that is, the same proportion of the first type of batch processing load, the total cost decreases with an increase in its delay. The larger the proportion of the first type of batch processing load, the more significant the trend of the total cost reduction, as shown in Fig. 18 and Table 9. For instance, with a compensation price of 0.04 \$/kW, as the delay increases from 2 h to 12 h, the total cost reduction rates are 11.2, 8.0, 5.1, 3.2, and 1.8 %, respectively. With a compensation price of 0.07 \$/kW, the total cost

**Table 6**  
Relationship between allowable delay and planning results.

Allowable delay/h	Investment cost/\$	Total cost/\$	Electric boiler/kW	Cooling system/kW	Heat pump/kW
0	21,754	535,045	524	576	202
2	20,145	515,274	483	534	188
4	19,173	497,338	463	504	180
6	18,951	485,210	455	503	176
8	18,863	479,083	450	502	176
10	18,792	475,432	449	500	175
12	18,759	473,687	449	498	175

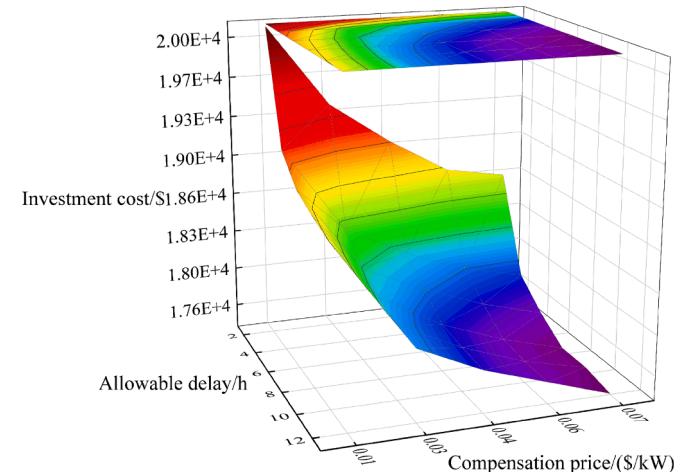


Fig. 16. Investment costs of the microgrid under different delays and compensation prices.

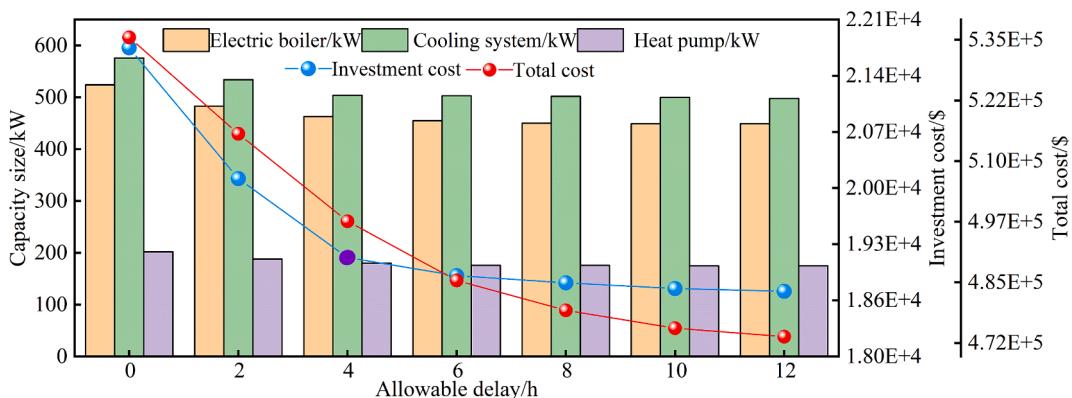
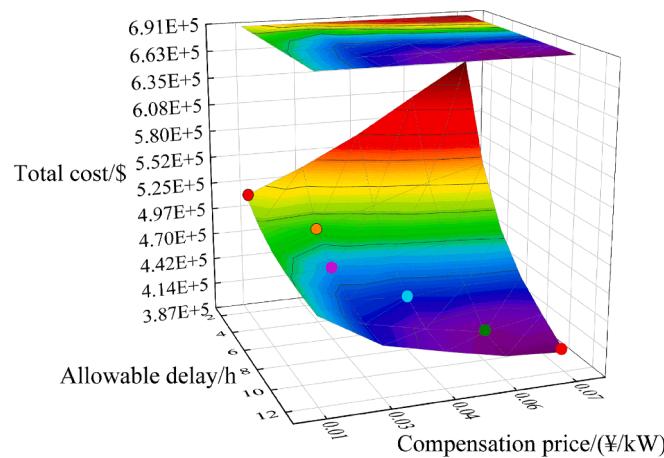


Fig. 15. Relationship between allowable delay and planning results.

**Table 7**

Investment costs of the microgrid under different delays and compensation prices.

Allowable delay/h	Compensation price/\$	0.01	0.03	0.04	0.06	0.07
		Investment cost/\$				
2		20,145	19,432	19,061	18,751	18,657
4		19,174	18,429	18,156	17,958	17,833
6		18,951	18,221	17,972	17,773	17,624
8		18,863	18,112	17,812	17,613	17,463
10		18,792	18,038	17,780	17,587	17,462
12		18,759	18,014	17,762	17,587	17,411

**Fig. 17.** Total costs of the microgrid under different delays and compensation prices.

reduction rates were 15.6, 10.8, 7.2, 4.9, and 4.6 %, respectively. Although the reduction rates decreased continuously, the reduction rate for a compensation price of 0.07 \$/kW is greater than that for 0.04

\$/kW.

For the same delay, the trend of the total cost changed inconsistently with the variation in the workload proportion of the first type of batch processing load, as shown in Fig. 19 and Table 10. With a longer delay, the total cost decreases as the compensation price increases. For example, with a 12 h delay, as the compensation price increases from 0.01 to 0.07 \$/kW, the reduction rates are 7.6, 3.2, 3.0, and 0.95 %. With a moderate delay, the total cost first decreases and then increases as the compensation price increases. For example, with a delay of 4 to 10 h, and specifically at 6 h, as the compensation price increases from 0.01 to 0.07 \$/kW, the total cost first decreases with reduction rates of 6.9 and 1.1 %, and then starts to increase with increase rates of 1.2 and

**Table 9**

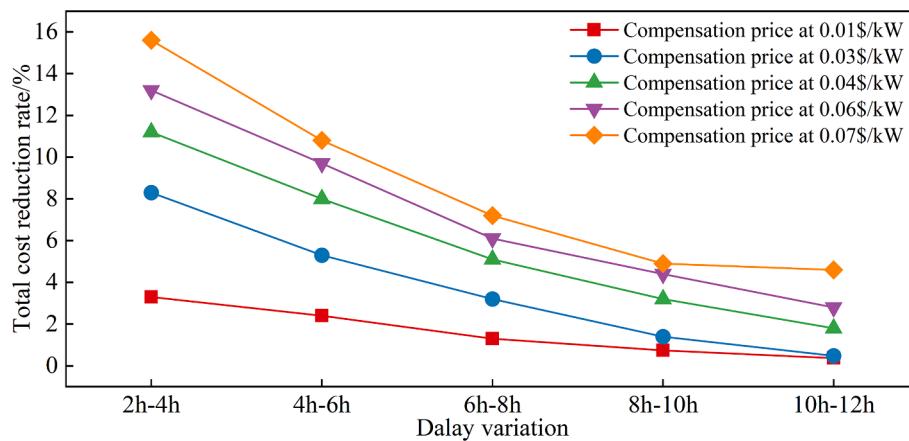
Relationship between delay variation and total cost reduction rate.

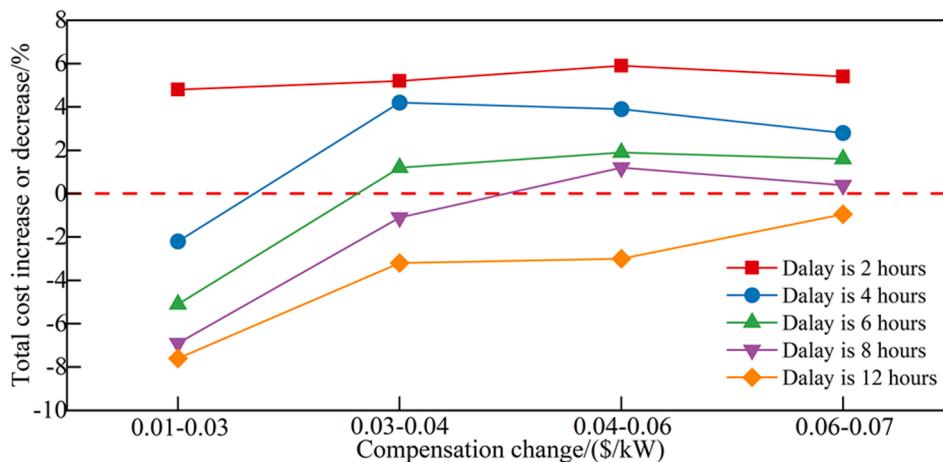
Allowable delay variation/h	Compensation price/(\$/kW)	0.01	0.03	0.04	0.06	0.07
		Total cost reduction rate/%				
[2,4]		3.3	8.3	11.2	13.2	15.6
[4,6]		2.4	5.3	8.0	9.7	10.8
[6,8]		1.3	3.2	5.1	6.1	7.2
[8,10]		0.74	1.4	3.2	4.4	4.9
[10,12]		0.37	0.48	1.8	2.8	4.6

**Table 8**

Total costs of the microgrid under different delays and compensation prices.

Allowable delay/h	Compensation price/\$	0.01	0.03	0.04	0.06	0.07
		Total cost /\$				
2	515,274	540,641	570,622	606,543	641,524	
4	497,338	486,458	506,712	526,468	541,443	
6	485,210	460,681	466,180	475,406	482,972	
8	479,083	446,042	441,073	446,474	448,266	
10	475,432	439,727	426,888	415,825	426,231	
12	473,687	437,663	423,476	410,573	406,640	

**Fig. 18.** Relationship between delay variation and total cost reduction rate.



**Fig. 19.** Relationship between compensation price change and the increase or decrease rate of total cost.

**Table 10**

Relationship between compensation price change and the increase or decrease rate of total cost.

Compensation price change/(\$/kW)	Allowable delay/h	Total cost increase or decrease rate/%					
		2	4	6	8	10	12
[0.01,0.03]	+4.8	-2.2	-5.1	-6.9	-7.5	-7.6	
[0.03,0.04]	+5.2	+4.2	+1.2	-1.1	-2.9	-3.2	
[0.04,0.06]	+5.9	+3.9	+1.9	+1.2	-1.1	-3.0	
[0.06,0.07]	+5.4	+2.8	+1.6	+0.39	+0.24	-0.95	

0.39 %, respectively. With a shorter delay, the total cost increased as its proportion increased. For instance, with a 2 h delay, as the compensation price increases from 0.01 to 0.07 \$/kW, the increase rates in the total cost are 4.8, 5.2, 5.9, and 5.4 %. Further study revealed that satisfying both a large delay and a high proportion of the first type of batch processing load provides the system with greater flexibility and economics.

## 6. Conclusions

This paper proposes a capacity allocation stochastic optimization method that considers flexible resources for core thermal management equipment such as electric boilers, cooling systems, and heat pumps in datacenter microgrids. Through a case study of a data center microgrid in a certain province, the following conclusions were drawn:

1) To address wind power uncertainty, a C-LGANs scenario generation method is proposed in this paper, with a coverage rate of 92.39 % and an envelope area of 0.43 %, outperforming C-WGAN-GP and Monte Carlo methods.

2) After considering various flexible resources, investment, operational, and total costs decreased by 33.2 %, 4.5 %, and 6.08 %, respectively. The capacities of the electric boiler and cooling system decreased by 27.9 % and 56.7 %, respectively, and the wind curtailment rate decreased by 7.18 %, demonstrating the effectiveness of the proposed strategies.

3) Sensitivity analysis reveals the impact of compensation prices for the first type of batch load and the allowable delay on the configuration results, aiding data center operators in designing corresponding configuration schemes based on relevant data.

Future research will delve deeper into the role of data center microgrids in broader energy networks, particularly their energy interaction and cooperation mechanisms with surrounding microgrids. Plans

include establishing more complex models to consider environmental changes, multi-energy system integration, efficient utilization of renewable energy, and studying how to optimize energy-sharing and load-balancing strategies among data center microgrids. By sharing our research findings and best practices with international partners, we expect to advance global progress in data center microgrid management, promote cooperation and exchange in this field among countries, collectively address global energy challenges, and achieve sustainable development goals.

## CRediT authorship contribution statement

**Yang Cui:** Validation, Software, Methodology, Conceptualization. **Yufeng Cheng:** Writing – original draft, Software, Formal analysis, Data curation. **Han Zhu:** Visualization, Formal analysis, Data curation. **Yuting Zhao:** Writing – review & editing. **Wuzhi Zhong:** Visualization, Data curation.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

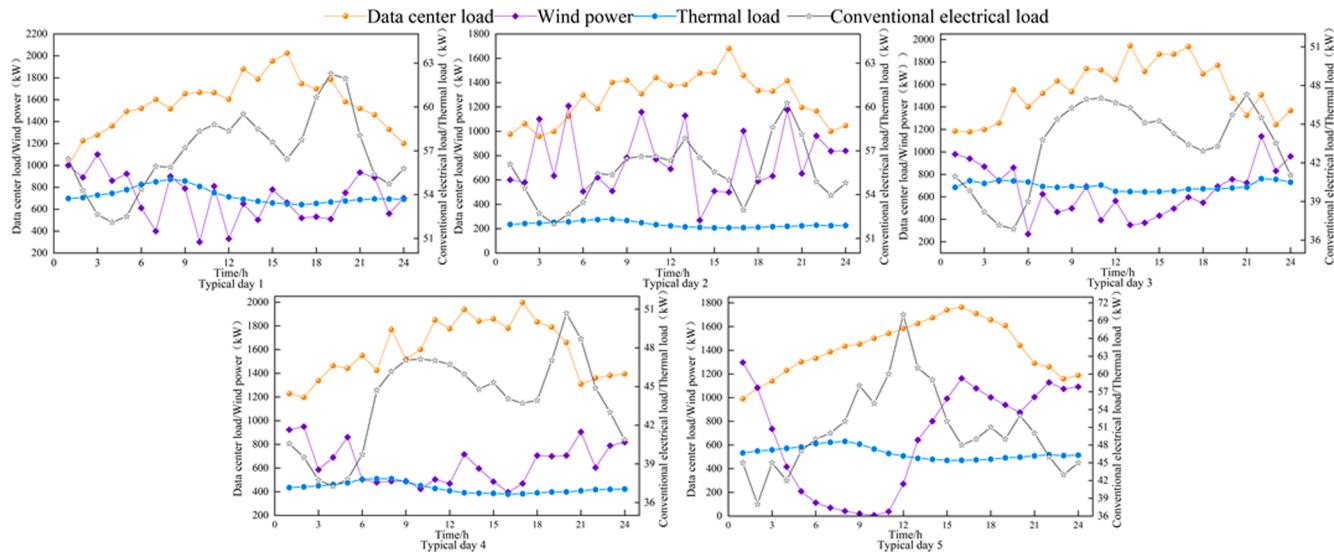
## Data availability

The authors do not have permission to share data.

## Acknowledgement

This study was supported by the National Natural Science Foundation of China (Grant no. 51777027).

## Appendix A



**Fig. A1.** Neural network architecture

**Table A1**

C-LSGANs training process.

**Algorithm 1: C-LSGANs**

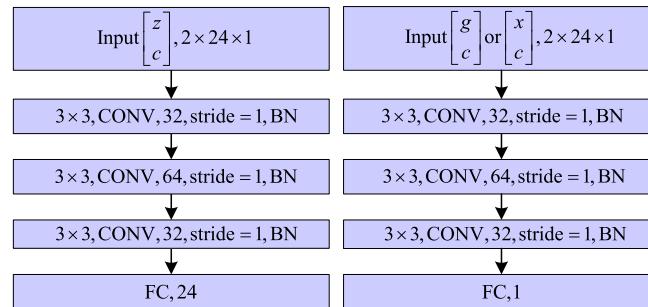
**Initialization:**  $\theta_d, \theta_g$ —Generator and discriminator network parameters.  $\alpha$ —Learning rate,  $\beta_1, \beta_2$ —Optimizer hyperparameters,  $m$ —Mini-batch size.

1:**for** episode = 1 to  $E$  **do**  
# Update discriminator network parameters  
2: Sample a batch of random noise.  
3: Randomly sample a batch of real samples from the training set.  
4: Randomly sample a batch of conditional information from the training set.  
5: Update the discriminator network parameters using gradient descent.  

$$\begin{cases} g_{\theta_d} \leftarrow \nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[ \frac{1}{2} (D(x^{(i)}|c^{(i)}) - 1)^2 + \frac{1}{2} (D(G(z^{(i)}|c^{(i)})) - 1)^2 \right] & \text{\# Update generator parameters} \\ \theta_d \leftarrow \theta_d - \alpha \cdot \text{Adam}(\theta_d, g_{\theta_d}, \beta_1, \beta_2) \end{cases}$$
  
6: Sample a batch of random noise.  
7: Randomly sample a batch of conditional information from the training set.  
8: Update the generator network parameters using gradient descent:  

$$\begin{cases} g_{\theta_g} \leftarrow \nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^m \left[ \frac{1}{2} (D(G(z^{(i)}|c^{(i)})) - 1)^2 \right] & 9: \text{end for} \\ \theta_g \leftarrow \theta_g - \alpha \cdot \text{Adam}(\theta_g, g_{\theta_g}, \beta_1, \beta_2) \end{cases}$$

## Appendix B



**Fig. B1.** Load curve of various typical days

## Appendix C

**Table C1**  
Electricity Pricing Situation.

Type of Electricity Price	Time period	Time-of-use pricing
Off-peak period	01:00–07:00	0.066 \$(kW•h)
	23:00–24:00	
Peak period	09:00–11:00	0.186 \$(kW•h)
	19:00–22:00	
Normal period	08:00	0.124 \$(kW•h)
	12:00–18:00	

**Table C2**  
Related equipment and other parameters.

Relevant parameters	Value/Unit	Relevant parameters	Value/Unit
$r$	0.05	$s$	20
$m$	365	$\alpha_{eb}$	177 \$/kW
$\alpha_{cs}$	209 \$/kW	$\alpha_{hp}$	287 \$/kW
$P_{sell}^{\max}$	800 kW	$P_{buy}^{\max}$	800 kW
$C_{cur}$	0.07 \$/kW	$C_{pc}$	0.014 \$/kW
$C_{hp}$	0.0002 \$/kW	$C_{es}$	0.014 \$/kW
$C_{eb}$	0.0003 \$/kW	$\eta_{eb}$	0.95
$\lambda$	2	$\eta_{cs}$	2
$\Upsilon$	1.5	$\eta_{hp}$	3
$\eta_{re}$	0.7	$S_{eb}^{\max}$	800 kW
$S_{cs}^{\max}$	1300 kW	$S_{hp}^{\max}$	300 kW
$\rho_a$	$1000/3600 \times 10^{-6}$ MW h/kg/°C	$C_a$	1.2 kg/m³
$H^{DC}$	$1.25 \times 10^{-6}$ MW/m²/°C	$T_{\min}^{\text{in}}, T_{\max}^{\text{in}}$	18,22/°C

## References

- [1] Gutiérrez SA, Botero JF, Branch-Bedoya JW. An adaptable and agnostic flow scheduling approach for data center networks. *J Netw Syst Manag* 2023;31(1):12.
- [2] Sovacool BK. A qualitative factor analysis of renewable energy and Sustainable Energy for All (SE4ALL) in the Asia-Pacific. *Energy Policy* 2013;59:393–403.
- [3] Gebara C, Laurent A. National SDG-7 performance assessment to support achieving sustainable energy for all within planetary limits. *Renew Sustain Energy Rev* 2023;173:112934.
- [4] Campbell BM, Hansen J, Rioux J, Stirling CM, Twomlow S, Wollenberg EL, et al. Urgent action to combat climate change and its impacts (SDG 13): transforming agriculture and food systems. *Curr Opin Environ Sustain* 2018;34:13–20.
- [5] Liu PR, Raftery AE. Country-based rate of emissions reductions should increase by 80% beyond nationally determined contributions to meet the 2 C target. *Commun Earth Environ* 2021;2(1):29.
- [6] Statista. Data centers worldwide by country 2024. Retrieved from: <https://www.statista.com/statistics/1228433/data-centers-worldwide-by-country/>.
- [7] Brightlio. 115 Data Center Stats You Should Know In 2024. Retrieved from: <https://brightlio.com/data-center-stats/>.
- [8] United States International Trade Commission (USITC). “Data Centers Around the World: A Quick Look”. Retrieved from: [https://www.usitc.gov/publications/332/executive\\_briefings/ebot\\_data\\_centers\\_around\\_the\\_world.pdf](https://www.usitc.gov/publications/332/executive_briefings/ebot_data_centers_around_the_world.pdf).
- [9] Yu L, Jiang T, Zou Y. Distributed real-time energy management in data center microgrids. *IEEE Trans Smart Grid* 2016;9(4):3748–62.
- [10] Li G, Sun Z, Wang Q, Wang S, Huang K. China’s green data center development: Policies and carbon reduction technology path. *Environ Res* 2023;116248.
- [11] Fu Y, Han X, Stershic J, Zuo W, Baker K, Lian J. Multi-stage Power Scheduling Framework for Data Center with Chilled Water Storage in Energy and Regulation Markets. *arXiv preprint arXiv:2007.09770*, 2020.
- [12] Lindberg J, Abdennadher Y, Chen J, Lesieutre B, Roald L. A guide to reducing carbon emissions through data center geographical load shifting. In: Proceedings of the Twelfth ACM International Conference on Future Energy Systems. 2021: 430–436.
- [13] Ye G, Gao F. Coordinated optimization scheduling of data center and electricity retailer based on cooperative game theory. *CPSS Trans Power Electron Appl* 2022;7(3):273–82.
- [14] Huang S, Yan D, Chen Y. An online algorithm for combined computing workload and energy coordination within a regional data center cluster. *Int J Electr Power Energy Syst* 2024;158:109971.
- [15] Xiao JW, Yang YB, Cui S, Wang Y. Cooperative online schedule of interconnected data center microgrids with shared energy storage. *Energy* 2023;285:129522.
- [16] Wu Z, Chen L, Wang J, Zhou M, Li G, Xia Q. Incentivizing the spatiotemporal flexibility of data centers toward power system coordination. *IEEE Trans Network Sci Eng* 2023.
- [17] Caishan G, Fengji L, Zexiang C, Zhao Yang D, Rui Z. Integrated planning of internet data centers and battery energy storage systems in smart grids. *Appl Energy* 2021; 281116093.
- [18] Zhang L, Deng J, Li Q, Yang Y, Lei M. Research on data center computing resources and energy load Co-optimization considering spatial-temporal allocation. *Comput Electr Eng* 2024;116:109206.
- [19] Niu T, Hu B, Xie K, Pan C, Jin H, Li C. Spacial coordination between data centers and power system considering uncertainties of both source and load sides. *Int J Electr Power Energy Syst* 2021;124:106358.
- [20] Wang J, Deng H, Liu Y, Guo Z, Wang Y. Coordinated optimal scheduling of integrated energy system for data center based on computing load shifting. *Energy* 2023;267:126585.
- [21] Li Y, Han M, Shahidehpour M, Li J, Long C. Data-driven distributionally robust scheduling of community integrated energy systems with uncertain renewable generations considering integrated demand response. *Appl Energy* 2023;335: 120749.
- [22] Araiz J, Martinez S. Contribution of the thermal inertia of trains to the primary frequency control of electric power systems. *Sustainable Energy Grids Networks* 2023;34:100988.
- [23] Xu S, Zhang H, Wang Z. Thermal management and energy consumption in air, liquid, and free cooling systems for data centers: a review. *Energies* 2023;16(3): 1279.
- [24] Marshall ZM, Duquette J. A techno-economic evaluation of low global warming potential heat pump assisted organic Rankine cycle systems for data center waste heat recovery. *Energy* 2022;242:122528.
- [25] Chauhan T, Bhandari R, Sivaraju KB, Chowdhury A, Agonafer D. Impact of immersion cooling on thermomechanical properties of low-loss material printed circuit boards. *J Enhanced Heat Transfer* 2021;28(7).
- [26] Khosravi A, Laukkonen T, Vuorinen V, Syri S. Waste heat recovery from a data centre and 5G smart poles for low-temperature district heating network. *Energy* 2021;218:119468.
- [27] Huang P, Copertaro B, Zhang X, Shen J, Lofgren I, Ronnelid M, et al. A review of data centers as prosumers in district energy systems: Renewable energy integration and waste heat reuse for district heating. *Appl Energy* 2020;258:114109.
- [28] Li Y, Huang J, Liu Y, Wang H, Wang Y, Ai X. A multicriteria optimal operation framework for a data center microgrid considering renewable energy and waste heat recovery: use of balanced decision making. *IEEE Ind Appl Mag* 2023.
- [29] Pakere I, Goncarovs K, Grävelsips A, Zirne M. Dynamic modelling of data center waste heat potential integration in district heating in Latvia. *Energies* 2024;17(2): 445.
- [30] Araya S, Wemhoff AP, Jones GF, Fleischer AS. Study of a lab-scale organic rankine cycle for the ultra-low-temperature waste heat recovery associated with data centers. *J Electron Packag* 2021;143(2):021001.

- [31] Nie J, Wang K, Kong X, Zhang H, Zhang S. Theoretical study and experimental validation on the applicable refrigerant for space heating air source heat pump. *Sustainability* 2023;15(12):9420.
- [32] Liu X, Hou G, Yang L. Combination optimization of green energy supply in data center based on simulated annealing particle swarm optimization algorithm. *Front Earth Sci*, 11: 1134523.
- [33] Li J, Qi W. Toward optimal operation of internet data center microgrid. *IEEE Trans Smart Grid* 2016;9(2):971–9.
- [34] Bomze IM, Gabl M. Optimization under uncertainty and risk: Quadratic and copositive approaches. *Eur J Oper Res* 2023;310(2):449–76.
- [35] Chen L, Yang D, Cai J, Yan Y. Robust optimization based coordinated network and source planning of integrated energy systems. *Int J Electr Power Energy Syst* 2024; 157:109864.
- [36] Chen F, Chen Y, Deng H, Lin W, Shao Z. Distributed robust operation of integrated energy system considering gas inertia and biogas-wind renewables. *Int J Electr Power Energy Syst* 2023;151:109123.
- [37] Yin X, Chen H, Liang Z, Zeng X, Zhu Y, Chen J. Robust transmission network expansion planning based on a data-driven uncertainty set considering spatio-temporal correlation. *Sustain Energy Grids Networks* 2023;33:100965.
- [38] Thang VV, Ha T, Li Q, Zhao Y. Stochastic optimization in multi-energy hub system operation considering solar energy resource and demand response. *Int J Electr Power Energy Syst* 2022;141:108132.
- [39] Dias JAS. Long-Term Hourly Scenario Generation for Correlated Wind and Solar Power combining Variational Autoencoders with Radial Basis Function Kernels. arXiv preprint arXiv:2306.16427, 2023.
- [40] Li S, Xu C, Wei L, Li R, Scenario AX. Generation of renewable energy based on improved diffusion model. In: 2023 IEEE Sustainable Power and Energy Conference (ISPEC). IEEE; 2023. p. 1–7.
- [41] Zhang Z, Zhu R. A distributionally robust optimization strategy for a wind-photovoltaic thermal storage power system considering deep peak load balancing of thermal power units. *Processes* 2024;12(3):534.
- [42] Li F, Zheng H, Li X. A novel hybrid model for multi-step ahead photovoltaic power prediction based on conditional time series generative adversarial networks. *Renew Energy* 2022;199:560–86.
- [43] Li Y, Li P, Yang T, Chen Z, Song X, Zhao Y, et al. A C-DCGAN-based method for generating extreme risk scenarios of high percentage new energy systems. In: 2023 3rd International Conference on Intelligent Power and Systems (ICIPS). IEEE, 2023: 588–593.
- [44] Wang Y, Liu Y, Yang Q. Operational scenario generation and forecasting for integrated energy systems. *IEEE Trans Ind Inf* 2023.
- [45] Ma X, Liu Y, Yan J, et al. A WGAN-GP-based scenarios generation method for wind and solar power complementary study. *Energies* 2023;16(7):3114.
- [46] Liang Z, Mieth R, Dvorkin Y. Operation-adversarial scenario generation. *Electr Pow Syst Res* 2022;212:108451.
- [47] Tang J, Liu J, Wu J, Jin G, Kang H. RAC-GAN-based scenario generation for newly built wind farm. *Energies* 2023;16(5):2447.
- [48] Huang N, He Q, Qi J, Hu Q, Wang R. Multinodes interval electric vehicle day-ahead charging load forecasting based on joint adversarial generation. *Int J Electr Power Energy Syst* 2022;143:108404.
- [49] Wang H, Liu Y, Yin C, Bai X, Sun B, Li M, et al. Day-ahead scenario generation method for renewable energy based on historical data analysis. In: International Conference on Big Data and Security. Singapore: Springer Nature Singapore; 2022. p. 37–46.
- [50] Wang Z, Hong T. Generating realistic building electrical load profiles through the Generative Adversarial Network (GAN). *Energ Buildings* 2020;224:110299.
- [51] Liao W, Wang Y, Wang Y, Powell K, Liu Q, Yang Z. Scenario generation for cooling, heating, and power loads using generative moment matching networks. *CSEE J Power Energy Syst* 2022;8(6):1730–40.
- [52] Jiang C, Mao Y, Chai Y, Yu M, Tao S. Scenario generation for wind power using improved generative adversarial networks. *IEEE Access* 2018;6:62193–203.
- [53] Yuan R, Wang B, Sun Y, Song X, Watada J. Conditional style-based generative adversarial networks for renewable scenario generation. *IEEE Trans Power Syst* 2022;38(2):1281–96.
- [54] Zhang Y, Ai Q, Xiao F, Hao R, Lu T. Typical wind power scenario generation for multiple wind farms using conditional improved Wasserstein generative adversarial network. *Int J Electr Power Energy Syst* 2020;114:105388.
- [55] Zhang X, Li D, Fu X. A novel wasserstein generative adversarial network for stochastic wind power output scenario generation. *IET Renew Power Gener* 2024.
- [56] Chen Y, Wang Y, Kirschen D, Zhang B. Model-free renewable scenario generation using generative adversarial networks. *IEEE Trans Power Syst* 2018;33(3): 3265–75.
- [57] Li Y, Li J, Wang Y. Privacy-preserving spatiotemporal scenario generation of renewable energies: A federated deep generative learning approach. *IEEE Trans Ind Inf* 2021;18(4):2310–20.
- [58] Dong W, Chen X, Yang Q. Data-driven scenario generation of renewable energy production based on controllable generative adversarial networks with interpretability. *Appl Energy* 2022;308:118387.
- [59] Qiao J, Pu T, Wang X. Renewable scenario generation using controllable generative adversarial networks with transparent latent space. *CSEE J Power Energy Syst* 2020;7(1):66–77.
- [60] Dalal D, Bilal M, Shah H, Anwarul IS, Anamitra P, Philip A. Cross-correlated scenario generation for renewable-rich power systems using implicit generative models. *Energies* 2023;16(4):1636.
- [61] Lu X, Kong F, Liu X, Yin J, Xiang Q, Yu H. Bulk savings for bulk transfers: Minimizing the energy-cost for geo-distributed data centers. *IEEE Trans Cloud Comput* 2017;8(1):73–85.
- [62] Yang P, Jiang H, Liu C, Kang L, Wang C. Coordinated optimization scheduling operation of integrated energy system considering demand response and carbon trading mechanism. *Int J Electr Power Energy Syst* 2023;147:108902.
- [63] Sun Q, Wu C, Li Z, Ren S. Colocation demand response: Joint online mechanisms for individual utility and social welfare maximization. *IEEE J Sel Areas Commun* 2016;34(12):3978–92.
- [64] Zhang Y, Qian T, Tang W. Buildings-to-distribution-network integration considering power transformer loading capability and distribution network reconfiguration. *Energy* 2022;244:123104.
- [65] Wang C, Xu C, Yao X, et al. Evolutionary generative adversarial networks. *IEEE Trans Evol Comput* 2019;23(6):921–34.
- [66] Veiner J, Alajaji F, Gharesifard B. A unifying generator loss function for generative adversarial networks. arXiv preprint arXiv:2308.07233, 2023.
- [67] Mao X, Li Q, Xie H, Lau RYK, Wang Z, Smolley SP. Least squares generative adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision; 2017. p. 2794–802.
- [68] Zhang M, Ai X, Fang J, Yao W, Zuo W, Chen Z. A systematic approach for the joint dispatch of energy and reserve incorporating demand response. *Appl Energy* 2018; 230:1279–91.
- [69] Zhang M, Wu Q, Wen J, Pan B, Qi S. Two-stage stochastic optimal operation of integrated electricity and heat system considering reserve of flexible devices and spatial-temporal correlation of wind power. *Appl Energy* 2020;275:115357.
- [70] Lv S, Wang H, Meng X, Yang C, Wang M. Optimal capacity configuration model of power-to-gas equipment in wind-solar sustainable energy systems based on a novel spatiotemporal clustering algorithm: A pathway towards sustainable development. *Renew Energy* 2022;201:240–55.
- [71] Tran NH, Oo TZ, Ren S, Zhu H, Eui-Nam H. Reward-to-reduce: An incentive mechanism for economic demand response of colocation datacenters. *IEEE J Sel Areas Commun* 2016;34(12):3941–53.