



## Two-stage robust optimization strategy for spatially-temporally correlated data centers with data-driven uncertainty sets

Zhihao Yang<sup>a</sup>, Anupam Trivedi<sup>b</sup>, Haoming Liu<sup>a,\*</sup>, Ming Ni<sup>a</sup>, Dipti Srinivasan<sup>b</sup>

<sup>a</sup> College of Energy and Electrical Engineering, Hohai University, Nanjing, 211100, China

<sup>b</sup> Department of Electrical and Computer Engineering, National University of Singapore, 119260, Singapore



### ARTICLE INFO

**Keywords:**

Data center  
Demand response  
Electricity market  
Nested column-and-constraint generation (C&CG)  
Two-stage robust optimization

### ABSTRACT

Data center buildings (DCBs) in data center parks consume significant and ever-growing amounts of electric power for processing tremendous data workloads in a modern city. The workloads are divided into two types: interactive workloads (IWs) with spatially transferrable characteristics and batch workloads (BWs) with temporally shiftable characteristics. The unique spatially-temporally workloads show great demand response capability, enabling reducing the electricity cost of DCBs. However, the source-load uncertainties such as the number of arriving workloads, and photovoltaic output affect the cost-efficient operation of DCBs in the electricity market. Inadequate server numbers during worst-case scenarios may result in the inability to fulfill user contracts. This paper presents a model for DCBs that incorporates battery energy storage systems, photovoltaic generation, and electrical loads for processing spatially-temporally correlated workloads. Considering multiple uncertainties, a two-stage robust optimization strategy is proposed to coordinate workload and power for DCBs. To accurately determine the boundaries of the uncertainty sets, a data-driven method using the 1-Wasserstein metric is adopted. This method is reformulated into a distributionally robust chance-constrained programming model. The nested column-and-constraint generation (C&CG) method is used to resolve the established two-stage robust optimization problem with mixed-integer recourse. Finally, case studies verify the effectiveness of the proposed method.

### 1. Introduction

Data centers have been the critical infrastructure in modern society with the increasing demand for processing tremendous workloads in the scenarios of online applications and services, e.g., commercial transactions, industrial manufacturing, and residential living. To provide internet services with low latency and high reliability, the internet service companies (ISCs), such as Amazon, Google, and Microsoft, have built data center buildings (DCBs) globally in different data centers, which are usually placed in data center parks [1]. The necessary servers and support infrastructure in data centers consume large amounts of electric power. It is estimated that data centers account for around 1% of the worldwide electric power demands, and energy consumption will triple or quadruple within the next few decades [2]. The electricity costs of data centers have been heavy burdens for ISCs. It is urgent and significant for ISCs to manage demand-side resources within DCBs to reduce electricity costs.

Active servers are utilized to process workloads in a DCB, and a

corresponding cooling system is installed to maintain the indoor temperature at an appropriate level. The workload and power are coupled with each other, which means the increased number of active servers leads to an increase in electric power. The pricing signal in the deregulated electricity market is time-varying, which incentivizes end-users to manage the demand-side resources to reduce the electricity cost. Existing studies about the demand response practices of data centers mainly focus on internal energy management [3,4]. However, transferrable workloads, including batch workloads (BWs) and interactive workloads (IWs), are also valuable and widely studied recently. The BWs such as numerical computing can be scheduled across time scales due to higher tolerance for service delay [5]. This implies that any available time slot before the deadline can be selected for processing the BWs. By utilizing widely used fast fiber-optic equipment, certain IWs with lower latency requirements can be dispatched from one DCB to another quickly via the internet [6]. BWs and IWs are therefore spatially-temporally correlated demand-side resources [7], which limit the number of active servers that can be utilized. Workloads in DCBs are unique loads in comparison with general thermal loads, battery energy

\* Corresponding author.

E-mail address: [liuhao@hhu.edu.cn](mailto:liuhao@hhu.edu.cn) (H. Liu).

Nomenclature		Constants
<i>Indices and Sets</i>		
$t$	Index of time slots	$\Delta t$ The time slot
$T$	Set of time slots	$\pi_t^{B1}, \pi_t^{B2}$ The buying electricity price from the grid
$j$	Index of data center buildings (DCBs)	$\pi_t^{S1}, \pi_t^{S2}$ The selling electricity price to the grid
$D$	Set of DCBs	$\pi_t^{\text{BESS}}$ The scheduling price of BESS
<i>Variables ('1': the first stage, '2': the second stage)</i>		$A_{j,\max}^{\text{DC}}$ The maximum number of servers
$P_{j,t}^{B1}, P_{j,t}^{B2}$	The electricity bought from the grid	$P_j^{\text{idle}}, P_j^{\text{peak}}$ The idle/ peak active power of a server
$P_{j,t}^{S1}, P_{j,t}^{S2}$	The electricity sold to the grid	$L_j^{\text{rate}}$ The service rate of a server
$P_{j,t}^{\text{DC1}}, P_{j,t}^{\text{DC2}}$	The total active power load of DCB	$C_j^{\text{PUE}}$ The designed PUE of a DCB
$A_{j,t}^{\text{DC1}}, A_{j,t}^{\text{DC2}}$	The number of active servers in the DCB	$C^{\text{DT}}$ The maximum delay time of IWs
$L_{j,t}^{\text{DC1}}, L_{j,t}^{\text{DC2}}$	The total arriving workloads in the DCB	$TD_j$ The maximum delay time of BWs
$A_{j,t}^{\text{BW1}}, A_{j,t}^{\text{BW2}}$	The number of active servers for batch workloads (BW)	$\eta_j^{\text{DC}}$ The percentage of the redundant servers
$A_{j,t}^{\text{IW1}}, A_{j,t}^{\text{IW2}}$	The number of active servers for interactive workloads (IW)	$\bar{L}_{j,t}^{\text{BW}}$ The predicted arriving BWs in the DCB
$A_{j,t}^{\text{R1}}, A_{j,t}^{\text{R2}}$	The number of reserved active servers	$\bar{L}_{j,t}^{\text{IW}}$ The predicted arriving IWs at the front-end server
$L_{j,t}^{\text{BW1}}, L_{j,t}^{\text{BW2}}$	The arriving BWs in the DCB	$\bar{P}_{j,t}^{\text{PV}}$ The predicted photovoltaic output power
$L_{j,t}^{\text{IW1}}, L_{j,t}^{\text{IW2}}$	The arriving IWs in the DCB	$E^{\text{RB}}$ The rated electric power of BESS
$P_{j,t}^{\text{BC1}}, P_{j,t}^{\text{BC2}}$	The charging power of the battery energy storage system (BESS)	$SOC_{j,\min}, SOC_{j,\max}$ The minimum and maximum state-of-charge (SOC) of BESS
$P_{j,t}^{\text{BD1}}, P_{j,t}^{\text{BD2}}$	The discharging power of the BESS	$\eta_j^{\text{BC}}, \eta_j^{\text{BD}}$ The charging/discharging efficiency of BESS
$E_{j,t}^{\text{B1}}, E_{j,t}^{\text{B2}}$	The stored electric power of BESS	$P_{\text{Grid}}^{\max}$ The maximum exchanging electricity between DCBs and the grid
$w_{j,t}^{\text{BC1}}, w_{j,t}^{\text{BD1}}, w_{j,t}^{\text{BC2}}, w_{j,t}^{\text{BD2}}$	The indicators of charging and discharging power of BESS	$P_{j,\max}^{\text{B}}$ The maximum active charging/discharging power of BESS
$w_{j,t}^{\text{B1}}, w_{j,t}^{\text{S1}}, w_{j,t}^{\text{B2}}, w_{j,t}^{\text{S2}}$	The indicators of buying and selling electricity	$P_{j,\max}^{\text{B2}}$ The maximum active charging/discharging power of BESS in the second stage

storage systems (BESS), and smart home devices which can only be controlled in one place based on the reducible and shiftable characteristics [3,8]. Electric vehicles in charging stations are also spatially-temporally correlated demand-side resources, but they cannot be scheduled from one place to another place immediately like workloads [9,10]. Therefore, the load characteristics of workloads processed in active servers differ significantly from typical electric loads. Leveraging the spatial-temporal demand response capability of DCBs will help ISCs reduce energy costs in the electricity market. However, some problems remain unsolved in the existing coordinated optimization of workloads and electric loads (ELs), which are summarized as follows:

- 1) Workload uncertainty is often overlooked in the electricity market, resulting in the uneconomical operation of DCBs [11]. Additionally, inadequate server numbers during worst-case scenarios may result in the inability to fulfill user contracts.
- 2) The unique workload model of DCBs imposes constraints on optimization that involve both electric power and active servers. The presence of integer variables, such as the number of active servers, creates difficulties in solving the uncertain optimization problem. Ji et al. [12] ignored the influence of the integer variables by equally distributing workloads to all servers. It is advisable to take into account the integer variables of the number of active servers to ensure efficient processing of workloads with a minimal number of active servers.

In previous work, energy management strategies of DCBs are usually modeled as deterministic optimization (DO) problems with a single data center [13] or multiple data centers [14]. Nevertheless, in the deregulated electricity market, uncertain factors such as the fluctuating on-site renewable generation like photovoltaic (PV), real-time electricity prices,

and arriving workloads could impact the cost of DCBs and introduce operation risk. Wang et al. [15] and Niu et al. [16], modeled the interaction between data centers and power grids as a two-stage stochastic optimization (SO) problem considering the uncertain workloads with a probability distribution. Ding et al. [17] considered the stochastic variables of renewable generation and dynamic electric and thermal loads in a two-stage problem. However, it is often challenging to obtain an accurate probability distribution for various uncertain factors in practical scenarios. Inaccurate estimations of these distributions can result in significant cost errors.

The robust optimization (RO) method is widely used for optimization under uncertainties without requiring the probability distribution but solely based on the uncertainty sets, compared with the SO method. The robust operation of DCBs has been exploited. Tian et al. [4] introduced a robust framework in the energy management of data centers to save costs by controlling the conservatism level of uncertain electricity prices. Chen et al. [18] proposed a robust approach for managing workloads and energy consumption considering data and power networks to deal with the uncertainties of renewable energy. Nevertheless, the aforementioned RO methods use interval-based uncertainty sets with predefined boundaries according to the subjective judgment of decision-makers, which affects the conservativeness of the RO problem. Therefore, Duan et al. [19] and Yuan et al. [20] proposed data-driven methods with the Wasserstein metric to determine the boundaries of uncertainty sets through historical data and risk preference. However, existing studies fail to address the combined impact of multiple uncertainties, particularly those related to demand-side resources such as spatial-temporal correlated workloads, which can significantly affect both the economic and reliable operation of data centers.

The column and constraint generation (C&CG) method is commonly used to tackle the aforementioned RO problem. This approach decomposes the problem into a master problem (MP) and a subproblem

(SP) [20,21]. The SP should be convex using the Karush-Kuhn-Tucker (KKT) conditions. Nevertheless, the RO problem becomes challenging to solve when the SP involves integer variables. Zhao et al. [22] have provided a nested C&CG method to deal with the two-stage RO problem with integer variables in the SP by further decomposing the SP.

To this end, this paper proposes a two-stage robust scheduling strategy for ISC that takes into account the distinctive features of workloads, aiming to minimize electricity costs. It coordinates workloads and electric power in the data center park considering multiple uncertainties. The main contributions of this paper are summarized as follows:

- 1) A two-stage RO strategy is proposed for ISC to manage multiple DCBs with spatial-temporal workloads in the data center park. The first stage is the non-adjustable stage with predicted values of uncertain variables, whereas the second stage is the adjustable stage with the updated uncertain variables based on the first-stage strategy. By accounting for multiple uncertainties, such as IWs, BWs, and PV generation, the proposed two-stage robust model determines day-ahead operating decisions.
- 2) A data-driven method is proposed to model uncertainty sets of workloads and PV generation. Specifically, the 1-Wasserstein metric is adopted to determine the boundaries of uncertainty sets, which are constructed by a distributionally robust chance-constrained programming model.
- 3) Given the second-stage integer recourse variables that capture the spatial-temporal demand response of DCBs, the nested C&CG algorithm is employed to decompose and solve the two-stage RO problem, which is converted into a tri-level optimization problem.

The rest of the paper is organized as follows. Section 2 presents the framework of the data center park and formulates the two-stage model with multiple uncertainties. Section 3 puts forwards the two-stage RO strategy with data-driven uncertainty sets. Section 4 performs case studies and analyses to verify the effectiveness of the proposed strategy. Finally, the conclusion is summarized in Section 5.

## 2. Problem formulation

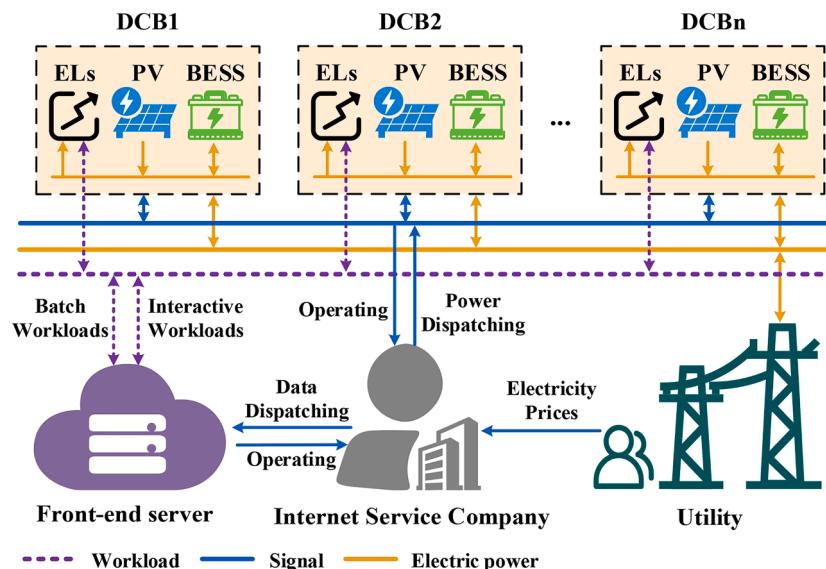
### 2.1. Problem statement

As depicted in Fig. 1, the data center park under consideration

assumes that multiple DCBs are managed by the same ISC. Each DCB is regarded as a prosumer with a BESS, a PV generation system, and ELs. The ISC with multiple DCBs can buy and sell electricity in the wholesale electricity market directly. To reduce the electricity cost, demand-side resources like IWs, BWs, and BESS in DCBs can be scheduled spatially and temporally by the ISC. Nevertheless, numerous factors such as the number of arriving BWs and IWs, as well as PV generation, introduce uncertainties that impact the operational state of active servers and determine the buying and selling of electricity by DCBs. The risk of balancing costs [23] in real-time operation can be reduced by taking into account the uncertain factors in day-ahead operation strategies. To this end, a day-ahead two-stage scheduling model is established for ISC considering uncertainties of the number of arriving BWs and IWs, and PV generation. Specifically, the model is formulated as follows:

- 1) The first stage is the non-adjustable stage with the optimal scheduling of workloads and ELs in light of uncertain PV generation, IWs, and BWs. Specifically, the buying and selling amount of electricity in the day-ahead market, as well as the charging/discharging status of BESS, is to be determined. Given the spatial and temporal correlation of IWs and BWs, any alterations made to the first-stage strategy during the decision-making process of the second stage would compromise the overall strategy, resulting in penalties for the ISC. Nonetheless, the impact of the penalty is not considered in this paper. Furthermore, the first-stage strategy should still be implemented in the second stage although the number of BWs and IWs could be reduced despite the existing uncertainties. Additionally, it is essential to reserve an adequate number of redundant servers for the second stage.
- 2) The second stage comprises the adjustable operation of the ISC under the worst-case scenario of uncertain parameters optimization taking into account uncertain variables. As a result of the difference between the predicted values and the ‘worst-case’ uncertain values, the redundant and inactive servers can be scheduled to process additional workloads. Furthermore, BESS can be adjusted based on the first-stage set points and the charging/discharging states. It shows that BESS also has some redundant capacity to be rescheduled to reduce the buying and selling electricity cost in the second stage [24, 25].

The strategies of the first non-adjustable stage and the second adjustable stage work in coordination with each other. For the sake of



**Fig. 1.** The framework of a data center park with DCBs.

distinction, the relevant variables of the first stage are marked with '1', and the relevant variables of the second stage are marked with '2'.

## 2.2. Objective function

The objective function is to minimize the total day-ahead electricity cost  $F$ . It includes two parts:

$$\min F = \sum_{t \in T} \sum_{j \in D} \left( F_{j,t}^1 + F_{j,t}^2 \right) \quad (1)$$

where  $F_{j,t}^1$  and  $F_{j,t}^2$  are the first-stage and second-stage electricity costs of DCB  $j$  at time  $t$ , which are expressed as follows:

$$F_{j,t}^1 = \left( \pi_t^{B1} P_{j,t}^{B1} - \pi_t^{S1} P_{j,t}^{S1} \right) \Delta t + \left( P_{j,t}^{BC1} + P_{j,t}^{BD1} \right) \pi_t^{BESS} \Delta t, \forall j \in D, \forall t \in T \quad (2a)$$

$$F_{j,t}^2 = \left( \pi_t^{B2} P_{j,t}^{B2} - \pi_t^{S2} P_{j,t}^{S2} \right) \Delta t + \left( P_{j,t}^{BC2} + P_{j,t}^{BD2} \right) \pi_t^{BESS} \Delta t, \forall j \in D, \forall t \in T. \quad (2b)$$

The first-stage electricity cost includes the cost of buying and selling electricity with day-ahead electricity prices and the operation cost of BESS based on the non-adjustable strategy. The second-stage electricity cost consists of the cost of buying and selling electricity with real-time electricity prices, as well as the operation cost of BESS based on the adjustable strategy.

## 2.3. Constraints in the first stage

Demand-side resources, including workloads, BESS, and PV, can participate in both first-stage and second-stage decision-making. In the first stage,  $\bar{u}$  is used to represent the predicted value of the variable  $u$ . Constraints in this stage are listed as follows.

### 1) ELs constraints

The ELs in the DCB include active servers that process workloads and corresponding auxiliary devices like cooling devices. Power usage efficiency (PUE) is a commonly used metric for assessing the energy consumption of DCBs. It is calculated by dividing the total energy consumption by the IT energy consumption, as follows:

$$P_{j,t}^{DC1} = \left[ P_j^{\text{idle}} + \left( C_j^{\text{PUE}} - 1 \right) P_j^{\text{peak}} \right] A_{j,t}^{\text{DC1}} + L_{j,t}^{\text{DC1}} \left( P_j^{\text{peak}} - P_j^{\text{idle}} \right) / L_j^{\text{rate}}, \forall j \in D, \forall t \in T \quad (3a)$$

$$L_{j,t}^{\text{DC1}} = L_{j,t}^{\text{BW1}} + L_{j,t}^{\text{IW1}}, \forall j \in D, \forall t \in T \quad (3b)$$

$$A_{j,t}^{\text{DC1}} = A_{j,t}^{\text{BW1}} + A_{j,t}^{\text{IW1}} + A_{j,t}^{\text{RI}}, \forall j \in D, \forall t \in T \quad (3c)$$

where Eq. (3a) represents the ELs of DCB  $j$  at time  $t$ , while (3b) and (3c) represent the total workloads to be processed and the total number of active servers, respectively.

In the first stage, optimization is performed using the predicted values of IWs and BWs with spatial-temporal transferrable capabilities. BWs can be transferred within a DCB across different time slots, and should satisfy the following constraints [26]:

$$\sum_{\tau=1}^t L_{j,\tau}^{\text{BW1}} \leq \sum_{\tau=1}^t \bar{L}_{j,\tau}^{\text{BW}}, \forall j \in D, \forall t \in T \quad (3d)$$

$$\sum_{\tau=1}^{t+TD_j} L_{j,\tau}^{\text{BW1}} \geq \sum_{\tau=1}^t \bar{L}_{j,\tau}^{\text{BW}}, \forall j \in D, \forall t \in [1, |T| - TD_j] \quad (3e)$$

$$\sum_{\tau=1}^T L_{j,\tau}^{\text{BW1}} \geq \sum_{\tau=1}^t \bar{L}_{j,\tau}^{\text{BW}}, \forall j \in D, \forall t \in [|T| - TD_j + 1, |T|]. \quad (3f)$$

Constraint (3d) represents that the processed BWs are no more than the accumulated BWs before time  $t$ . Constraint (3e) indicates that the accumulated BWs should be processed within the maximum delay time, and constraint (3f) suggests that all BWs should be processed in a day. It is assumed that all the BWs processed in the homogeneous active servers should satisfy the total number of BWs as the constraint (3g). Constraint (3h) ensures that the number of BWs and the corresponding active servers should be positive.

$$L_j^{\text{rate}} A_{j,t}^{\text{BW1}} \geq L_{j,t}^{\text{BW1}}, \forall j \in D, \forall t \in T \quad (3g)$$

$$L_{j,t}^{\text{BW1}}, A_{j,t}^{\text{BW1}} \geq 0, \forall j \in D, \forall t \in T. \quad (3h)$$

On the other hand, the dispatched IWs at different DCBs should be more than the arriving IWs at the front-end server, which can be expressed as (3i):

$$\sum_{j \in D} L_{j,t}^{\text{IW1}} \geq \bar{L}_t^{\text{FIW}}, \forall t \in T. \quad (3i)$$

Due to the proximity of each DCB in a data center park, the transmission time is too short to be factored in, but the processing time cannot be disregarded. However, the active servers to process IWs should meet users' requirements for the time delay, which is the quality of service (QoS) constraint (3j). Constraint (3k) ensures that the number of IWs and the corresponding active servers should be positive.

$$\frac{1}{L_j^{\text{rate}} - L_{j,t}^{\text{IW1}} / A_{j,t}^{\text{IW1}}} \leq C^{\text{DT}}, \forall j \in D, \forall t \in T \quad (3j)$$

$$L_{j,t}^{\text{IW1}}, A_{j,t}^{\text{IW1}} \geq 0, \forall j \in D, \forall t \in T. \quad (3k)$$

The constraint (3l) is that the number of active servers in the first stage should not exceed a certain limit. To ensure the reliable operation of DCBs and cope with the workload uncertainties, some redundant active servers are usually prepared in the first stage, which should satisfy the constraints (3 m):

$$A_{j,t}^{\text{DC1}} \leq A_{j,\text{max}}^{\text{DC}}, \forall j \in D, \forall t \in T \quad (3l)$$

$$0 \leq A_{j,t}^{\text{RI}} \leq \eta_j^{\text{DC}} A_{j,\text{max}}^{\text{DC}}, \forall j \in D, \forall t \in T. \quad (3m)$$

### 1) BESS constraints

The state of charge (SOC) of BESS at time  $t$  is related to the charging and discharging power at the time  $t - 1$ , which can be expressed as:

$$SOC_{j,t} = E_{j,t}^{\text{B1}} / E_j^{\text{RB}}, \forall j \in D, \forall t \in T \quad (4a)$$

$$E_{j,t}^{\text{B1}} = E_{j,t-1}^{\text{B1}} + \left( \eta_j^{\text{BC}} P_{j,t}^{\text{BC1}} - \frac{P_{j,t}^{\text{BD1}}}{\eta_j^{\text{BD}}} \right) \Delta t, \forall j \in D, \forall t \in T \setminus \{1\}. \quad (4b)$$

The constraints of BESS are as follows:

$$E_{j,1}^{\text{B1}} = E_{j,T}^{\text{B1}}, \forall j \in D \quad (4c)$$

$$SOC_{j,\text{min}} \leq SOC_{j,t} \leq SOC_{j,\text{max}}, \forall j \in D, \forall t \in T \quad (4d)$$

$$0 \leq \eta_j^{\text{BC}} P_{j,t}^{\text{BC1}} \leq w_{j,t}^{\text{BC1}} P_{j,\text{max}}^{\text{B}}, \forall j \in D, \forall t \in T \quad (4e)$$

$$0 \leq \frac{P_{j,t}^{\text{BD1}}}{\eta_j^{\text{BD}}} \leq w_{j,t}^{\text{BD1}} P_{j,\text{max}}^{\text{B}}, \forall j \in D, \forall t \in T \quad (4f)$$

$$0 \leq w_{j,t}^{\text{BC1}} + w_{j,t}^{\text{BD1}} \leq 1, \forall j \in D, \forall t \in T. \quad (4g)$$

Constraint (4c) indicates that the capacity of BESS at the end of a day is the same as the initial capacity in a day to prepare for the next scheduling day. Constraint (4d) limits the SOC of BESS. Constraint (4e)

and constraint (4f) limit the charging and discharging power of BESS. Constraint (4g) indicates that the BESS cannot simultaneously charge and discharge.

### 1) The power balance constraints

Both in the first stage and the second stage, the power balance constraints should be satisfied at each DCB. In the first stage, the values of PV and workloads all take the predicted value because the real values in the second stage are unknown. The power balance should be satisfied within each DCB:

$$P_{j,t}^{B1} + \bar{P}_{j,t}^{PV} + P_{j,t}^{BD1} = P_{j,t}^{DC1} + P_{j,t}^{BC1} + P_{j,t}^{S1}, \forall j \in D, \forall t \in T. \quad (5)$$

### 1) The electricity transaction constraints

Each DCB can buy and sell electricity in the day-ahead electricity market, but cannot buy and sell simultaneously. Furthermore, the maximum buying and selling of electric power cannot surpass the transmission limit of the connection line with the DCB. The constraints are (6a)-(6c) [25]:

$$0 \leq w_{j,t}^{B1} + w_{j,t}^{S1} \leq 1, \forall j \in D, \forall t \in T \quad (6a)$$

$$0 \leq P_{j,t}^{B1} \leq w_{j,t}^{B1} P_{\max}^{\text{Grid}}, \forall j \in D, \forall t \in T \quad (6b)$$

$$0 \leq P_{j,t}^{S1} \leq w_{j,t}^{S1} P_{\max}^{\text{Grid}}, \forall j \in D, \forall t \in T. \quad (6c)$$

## 2.4. Constraints in the second stage

In the first stage, the decision-making strategies for buying and selling electricity are based on the deterministic model that embodies an optimistic outlook. However, in the second stage, adjustments are made based on first-stage operational decisions and specific realizations of uncertain parameters.

Similar to the first stage, the uncertain workloads in the second stage can also be viewed as the demand-side resources, particularly in the spatial scheduling with IWs. The spatial-temporal correlation should be taken into account as well due to the limitation of servers.

### 1) ELs constraints

Since the strategies formulated in the first stage should remain unchanged, redundant servers can be utilized to process uncertain workloads. To prevent disorganized scheduling, it is vital to refrain from allocating delay-tolerant BWs across different time slots in the second stage. The additional BWs should be processed by redundant servers directly. Nevertheless, the arriving IWs can still be dispatched across multiple DCBs. The models and constraints of ELs are expressed as:

$$\begin{aligned} P_{j,t}^{DC2} &= \left[ P_j^{\text{idle}} + \left( C_j^{\text{PUE}} - 1 \right) P_j^{\text{peak}} \right] A_{j,t}^{R2} + L_{j,t}^{DC2} \left( P_j^{\text{peak}} - P_j^{\text{idle}} \right) / L_j^{\text{rate}}, \forall j \in D, \forall t \\ &\in T \end{aligned} \quad (7a)$$

$$L_{j,t}^{DC2} = \max \left\{ 0, \tilde{L}_{j,t}^{\text{BW}} - \bar{L}_{j,t}^{\text{BW}} \right\} + L_{j,t}^{\text{IW2}}, \forall j \in D, \forall t \in T \quad (7b)$$

$$L_j^{\text{rate}} A_{j,t}^{\text{BW2}} \geq \max \left\{ 0, \tilde{L}_{j,t}^{\text{BW}} - \bar{L}_{j,t}^{\text{BW}} \right\}, \forall j \in D, \forall t \in T \quad (7c)$$

$$\sum_{j \in D} \left( L_{j,t}^{\text{IW1}} + L_{j,t}^{\text{IW2}} \right) \geq \tilde{L}_t^{\text{FIW}}, \forall t \in T \quad (7d)$$

$$\frac{1}{L_j^{\text{rate}} - \left( L_{j,t}^{\text{IW1}} + L_{j,t}^{\text{IW2}} \right) / \left( A_{j,t}^{\text{IW1}} + A_{j,t}^{\text{IW2}} \right)} \leq C^{\text{DT}}, \forall j \in D, \forall t \in T \quad (7e)$$

$$L_{j,t}^{\text{IW1}} + L_{j,t}^{\text{IW2}} \geq 0, \forall j \in D, \forall t \in T \quad (7f)$$

$$A_{j,t}^{\text{IW1}} + A_{j,t}^{\text{IW2}} \geq 0, \forall j \in D, \forall t \in T \quad (7g)$$

$$0 \leq A_{j,t}^{\text{BW2}} + A_{j,t}^{\text{IW2}} \leq A_{j,t}^{\text{R1}} + A_{j,t}^{\text{R2}}, \forall j \in D, \forall t \in T \quad (7h)$$

$$0 \leq A_{j,t}^{\text{R2}} \leq \eta_j^{\text{DC}} A_{j,\max}^{\text{DC}} - A_{j,t}^{\text{R1}}, \forall j \in D, \forall t \in T. \quad (7i)$$

[Eq. \(7a\)](#) is the ELs model in the second stage, which only considers the additional redundant servers. [Eq. \(7b\)](#) denotes the number of workloads that need to be processed in the second stage. It should be noted that the deviation amount of IWs may turn out to be negative due to uncertain factors, leading to the shutdown of the corresponding active servers. Constraint (7c) limits the number of active servers to process BWs. Constraints (7d)-(7g) represent QoS constraints combined with two stages' IWs. Constraints (7h) and (7i) limit the total redundant servers.

### 1) BESS constraints

To prevent frequent changes in the charging and discharging states of the BESS in the second stage, it is necessary to ensure that the states remain consistent with the first stage. Nevertheless, it is acceptable to make reasonable adjustments to the charging and discharging power within a certain range. However, it is imperative to ensure that the BESS models and constraints are still satisfied [27].

$$SOC_{j,t} = E_{j,t}^{\text{B2}} / E_j^{\text{RB}}, \forall j \in D, \forall t \in T \quad (8a)$$

$$\begin{aligned} E_{j,t}^{\text{B2}} &= E_{j,t-1}^{\text{B2}} \\ &+ \left[ \eta_j^{\text{BC}} \left( P_{j,t}^{\text{BC1}} + P_{j,t}^{\text{BC2}} \right) - \left( P_{j,t}^{\text{BD1}} + P_{j,t}^{\text{BD2}} \right) / \eta_j^{\text{BD}} \right] \Delta t, \forall j \in D, \forall t \in T \setminus \{1\} \end{aligned} \quad (8b)$$

$$E_{j,1}^{\text{B2}} = E_{j,T}^{\text{B2}}, \forall j \in D \quad (8c)$$

$$SOC_{j,\min} \leq SOC_{j,t} \leq SOC_{j,\max}, \forall j \in D, \forall t \in T \quad (8d)$$

$$0 \leq \eta_j^{\text{BC}} \left( P_{j,t}^{\text{BC1}} + P_{j,t}^{\text{BC2}} \right) \leq P_{j,\max}^{\text{B}}, \forall j \in D, \forall t \in T \quad (8e)$$

$$0 \leq \left( P_{j,t}^{\text{BD1}} + P_{j,t}^{\text{BD2}} \right) / \eta_j^{\text{BD}} \leq P_{j,\max}^{\text{B}}, \forall j \in D, \forall t \in T \quad (8f)$$

$$-w_{j,t}^{\text{BC1}} P_{j,\max}^{\text{B2}} \leq P_{j,t}^{\text{BC2}} \leq w_{j,t}^{\text{BC1}} P_{j,\max}^{\text{B2}}, \forall j \in D, \forall t \in T \quad (8g)$$

$$-w_{j,t}^{\text{BD1}} P_{j,\max}^{\text{B2}} \leq P_{j,t}^{\text{BD2}} \leq w_{j,t}^{\text{BD1}} P_{j,\max}^{\text{B2}}, \forall j \in D, \forall t \in T. \quad (8h)$$

The meanings of the models and constraints (8a)-(8f) in this stage are similar to the first stage, except that the total charging and discharging power must factor in the second stage charging and discharging power, which can be adjusted upwards or downwards, as specified in (8g)-(8h).

### 1) The power balance constraints

In the second stage, both PV and workloads assume uncertain values as per the ambiguity uncertainty sets. The power balance constraints of the first stage are modified with adjustable variables in this stage, which can be expressed as (9):

$$\begin{aligned} P_{j,t}^{\text{B1}} + P_{j,t}^{\text{B2}} + \bar{P}_{j,t}^{\text{PV}} + P_{j,t}^{\text{BD1}} + P_{j,t}^{\text{BD2}} &= P_{j,t}^{\text{DC1}} + P_{j,t}^{\text{DC2}} + P_{j,t}^{\text{BC1}} + P_{j,t}^{\text{BC2}} + P_{j,t}^{\text{S1}} \\ &+ P_{j,t}^{\text{S2}}, \forall j \\ &\in D, \forall t \in T. \end{aligned} \quad (9)$$

### 1) The electricity transaction constraints

In the second stage, each DCB cannot buy and sell electricity simultaneously as well. It should meet the following constraints:

$$0 \leq w_{j,t}^{B2} + w_{j,t}^{S2} \leq 1, \forall j \in D, \forall t \in T \quad (10a)$$

$$0 \leq P_{j,t}^{B2} \leq w_{j,t}^{B2} P_{\max}^{\text{Grid}}, \forall j \in D, \forall t \in T \quad (10b)$$

$$0 \leq P_{j,t}^{S2} \leq w_{j,t}^{S2} P_{\max}^{\text{Grid}}, \forall j \in D, \forall t \in T. \quad (10c)$$

In addition, the total day-ahead buying and selling of electricity in two stages should also be within the transmission limit:

$$-P_{\max}^{\text{Grid}} \leq P_{j,t}^{B1} + P_{j,t}^{B2} - P_{j,t}^{S1} - P_{j,t}^{S2} \leq P_{\max}^{\text{Grid}}, \forall j \in D, \forall t \in T. \quad (10d)$$

### 3. Two-stage robust optimization strategy based on data-driven uncertainty sets

#### 3.1. Data-driven uncertainty sets

The PV generation is subject to uncertainty, mainly because of the imprecise prediction of solar radiation and other related environmental factors. Furthermore, the deviations in user behaviors also contribute to workload uncertainties. These various uncertainties could impact the decision-making for scheduling at each DCB. As the probability distributions of these multiple uncertainties are usually difficult to be obtained, RO is employed to model these uncertainties. Various types of uncertainty sets, such as box [15] and polyhedral [4] uncertainty sets, have been utilized to model the uncertainties associated with PV generation and workloads in a DCB. To avoid the over-conservativeness of traditional RO with box uncertainty sets, this paper adopts the polyhedral uncertainty sets. It can regulate the conservativeness of uncertain variables by the budget parameters  $\Gamma$  ( $\Gamma_j^{\text{PV}}$ ,  $\Gamma_j^{\text{BW}}$ , and  $\Gamma^{\text{FIW}}$ ) [10]:

$$\mathbf{U} := \begin{cases} \mathbf{u} = [\tilde{P}_{j,t}^{\text{PV}}, \tilde{L}_{j,t}^{\text{BW}}, \tilde{L}_t^{\text{FIW}}]^T, \forall j \in D, \forall t \in T \\ \begin{aligned} \tilde{P}_{j,t}^{\text{PV}} &= \bar{P}_{j,t}^{\text{PV}} - v_{j,t}^{\text{PV}-} \Delta P_{j,t}^{\text{PV}} + v_{j,t}^{\text{PV}+} \Delta P_{j,t}^{\text{PV}} \\ \sum_{t \in T} (v_{j,t}^{\text{PV}-} + v_{j,t}^{\text{PV}+}) &\leq \Gamma_j^{\text{PV}} \\ \tilde{L}_{j,t}^{\text{BW}} &= \bar{L}_{j,t}^{\text{BW}} - v_{j,t}^{\text{BW}-} \Delta L_{j,t}^{\text{BW}} + v_{j,t}^{\text{BW}+} \Delta L_{j,t}^{\text{BW}} \\ \sum_{t \in T} (v_{j,t}^{\text{BW}-} + v_{j,t}^{\text{BW}+}) &\leq \Gamma_j^{\text{BW}} \\ \tilde{L}_t^{\text{FIW}} &= \bar{L}_t^{\text{FIW}} - v_t^{\text{FIW}-} \Delta L_t^{\text{FIW}} + v_t^{\text{FIW}+} \Delta L_t^{\text{FIW}} \\ \sum_{t \in T} (v_t^{\text{FIW}-} + v_t^{\text{FIW}+}) &\leq \Gamma^{\text{FIW}} \end{aligned} \end{cases}. \quad (11)$$

Nonetheless, in some studies, the boundaries for the predicted errors  $\Delta u$  (i.e.  $\Delta P_{j,t}^{\text{PV}}$ ,  $\Delta L_{j,t}^{\text{BW}}$ , and  $\Delta L_t^{\text{FIW}}$ ) are set subjectively and the use of inappropriate values could cause significant deviation in the optimization results. This paper proposes a data-driven approach with the 1-Wasserstein metric to decide the ambiguity sets of these uncertain variables [20].

Given the historical sample set of predicted error  $\widehat{\mathcal{U}} := (\Delta \hat{u}^{(1)}, \Delta \hat{u}^{(2)}, \dots, \Delta \hat{u}^{(N)})$  with the mean value  $\Delta \hat{\mu}$  and the variance  $\widehat{\Sigma}$ , the standardized format of any data  $n$  can be calculated by Eq. (12a).

$$\Delta \hat{u}^{(n)} = \widehat{\Sigma}^{-1/2} (\Delta \hat{u}^{(n)} - \Delta \hat{\mu}). \quad (12a)$$

Then the standardized sample set of the predicted error is  $\mathcal{U} := \{\Delta \hat{u} \in \mathbb{R}, n \in [1, N] | -\Delta \bar{u} \leq \Delta \hat{u}^{(n)} \leq \Delta \bar{u}\}$  with the mean value  $\Delta \hat{\mu} = 0$ . The ambiguity boundary  $\Delta \bar{u}$  can be solved by the following risk-based chance-constrained programming problem with the confidence level  $\rho$ :

$$\min_{0 \leq \Delta \bar{u} \leq \Delta u_{\max}} \Delta \bar{u} \quad (12b)$$

s.t.  $\mathbb{Q}(\Delta \hat{u} \notin \mathcal{U}) \leq 1 - \rho$

where  $\mathbb{Q}$  represents the probability distribution of the predicted error of uncertain variables. In practice, however, the precise probability distribution  $\mathbb{Q}$  cannot be obtained. To this end, the empirical distribution  $\widehat{\mathbb{Q}}$  is set as the approximation of the true probability distribution  $\mathbb{Q}$ , that is:

$$\widehat{\mathbb{Q}} = \frac{1}{N} \sum_{n=1}^N \delta_{\Delta \hat{u}^{(n)}} \quad (12c)$$

where  $\delta_{\Delta \hat{u}^{(n)}}$  represents the Dirac function of  $\Delta \hat{u}^{(n)}$ .

With more sampled data, the distance between  $\widehat{\mathbb{Q}}$  and  $\mathbb{Q}$  will be narrowed and  $\widehat{\mathbb{Q}}$  would finally converge  $\mathbb{Q}$ . In this paper, the 1-Wasserstein metric is used to characterize the distance.

**Definition 1.** For any probability distribution  $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{C}$ , the Wasserstein Metric can be defined as:

$$D_w(\mathbb{Q}_1, \mathbb{Q}_2) = \inf_{\Pi} \left\{ \int_{\mathcal{U}^2} \|u_1 - u_2\|_{\gamma} \Pi(du_1, du_2) \right\} \quad (12d)$$

where  $\mathcal{C}$  represents the ambiguity set of all probability distributions with support  $\mathcal{U}$ ,  $\Pi$  is the joint distribution of  $u_1$  and  $u_2$  with marginal distributions  $\mathbb{Q}_1$  and  $\mathbb{Q}_2$ ,  $\|\cdot\|_{\gamma}$  denotes the  $\gamma$ -norm in  $\mathbb{R}$ . When  $\gamma = 1$ , the distance represents the 1-Wasserstein metric.

Hence, (12b) can be rewritten as: where  $\widehat{\mathcal{C}}$  represents the 1-Wasserstein ball with the radius  $\epsilon$  and the center  $\widehat{\mathbb{Q}}$ , which is defined as:

$$\widehat{\mathcal{C}} := \{\mathbb{Q} \in \mathcal{C} | D_w(\mathbb{Q}, \widehat{\mathbb{Q}}) \leq \epsilon\}. \quad (12f)$$

Further, according to the derivation of references [19,20], (12e) can be converted to:

$$\begin{aligned} \min_{\lambda \geq 0, 0 \leq \Delta \bar{u} \leq \Delta u_{\max}} \Delta \bar{u} \\ \text{s.t. } \lambda \cdot \epsilon + \frac{1}{N} \sum_{n=1}^N (1 - \lambda (\Delta \bar{u} - \|\Delta \hat{u}^{(n)}\|_{\infty}))^+ \leq 1 - \rho \end{aligned} \quad (12g)$$

where  $(\vartheta)^+ = \max(\vartheta, 0)$ . The optimization problem can be easily solved by the nested bisection search method [28].

To avoid the over-conservative situation of the radius  $\epsilon$  in the real numerical experience, reference [19] provides the following equations to get the ball radius:

$$\epsilon = C \sqrt{\frac{1}{N} \ln \frac{1}{1 - \kappa}} \quad (12h)$$

$$C = 2 \inf_{\phi > 0} \sqrt{\frac{1}{2\phi} \left[ 1 + \ln \left( \frac{1}{N} \sum_{n=1}^N e^{\phi \|\Delta \hat{u}^{(n)} - \Delta \hat{\mu}\|_1^2} \right) \right]} \quad (12i)$$

where  $\kappa$  denotes a confidence level. The decision variable  $\phi$  can be easily solved using the Fmincon function provided in the MATLAB optimization toolbox by calling the interior-point method [28].

To this end, the real value of the ambiguity boundary  $\Delta \bar{u}$  can be derived by using the inverse standardization method, that is:

$$\Delta u = \widehat{\Sigma}^{1/2} \Delta \bar{u} + \Delta \hat{\mu} \quad (12j)$$

#### 3.2. The day-ahead two-stage RO strategy

The proposed day-ahead strategy for buying and selling electricity coordinating workloads and electric power with multiple uncertainties is a two-stage RO problem. For the convenience of description, the model is written in a compact form:

$$F : \min_{\mathbf{x}} \mathbf{a}^T \mathbf{x} + \max_{\mathbf{u}} \min_{\mathbf{y}, \mathbf{z}} \mathbf{b}^T \mathbf{y}$$

s.t. 
$$\begin{cases} \mathbf{Ax} = \mathbf{c} \\ \mathbf{Dx} \geq \mathbf{d} \\ \mathbf{Ex} + \mathbf{Fy} + \mathbf{Gz} = \mathbf{f} - \mathbf{Ju} \\ \mathbf{Lx} + \mathbf{Hy} + \mathbf{Iz} \geq \mathbf{h} - \mathbf{Ku} \\ \mathbf{x} \in \mathbf{X}, \mathbf{u} \in \mathbf{U}, \mathbf{y} \in \mathbf{Y}, \mathbf{z} \in \mathbf{Z} \end{cases} \quad (13)$$

$$\mathbf{x} = \begin{bmatrix} P_{j,t}^{B1}, P_{j,t}^{S1}, w_{j,t}^{B1}, w_{j,t}^{S1}, A_{j,t}^{R1}, A_{j,t}^{BW1}, A_{j,t}^{IW1}, \\ L_{j,t}^{BW1}, L_{j,t}^{IW1}, P_{j,t}^{BC1}, P_{j,t}^{BD1}, w_{j,t}^{BC1}, w_{j,t}^{BD1} \end{bmatrix}^T, \forall j \in D, \forall t \in T \quad (14a)$$

$$\mathbf{y} = \left[ P_{j,t}^{\text{B2}}, P_{j,t}^{\text{S2}}, L_{j,t}^{\text{IW2}}, P_{j,t}^{\text{BC2}}, P_{j,t}^{\text{BD2}} \right]^T, \forall j \in D, \forall t \in T \quad (14\text{b})$$

$$\mathbf{z} = \left[ w_{j,t}^{\text{B2}}, w_{j,t}^{\text{S2}}, A_{j,t}^{\text{R2}}, A_{j,t}^{\text{BW2}}, A_{j,t}^{\text{IW2}} \right]^T, \forall j \in D, \forall t \in T. \quad (14c)$$

The set  $x$  includes the continuous and discrete decision variables in the first stage, i.e. buying and selling power and operational variables of demand-side resources. The set  $y$  includes continuous decision variables and  $z$  includes discrete decision variables in the second stage. It should be noted that the impact of a single workload may be insignificant in the decision-making process, but the active state of a single server can have a great effect on the feasibility of the system. To this end,  $L_{j,t}^{IW2}$  can be relaxed into continuous variables, thus alleviating the computing burden.

If the uncertain variables are determined, the proposed two-stage RO model can be transformed into a deterministic model, which is essentially a mixed integer linear programming (MILP) problem and can be solved by commercial solvers such as CPLEX and GUROBI.

### 3.3. Solution methodology

As shown in formula (13), the established two-stage RO model contains a subproblem as a MILP problem with the discrete decision variables  $\mathbf{z}$  in (14c). In this case, the standard C&CG method is no longer applicable due to the non-convexity of the subproblem. In light of this, the nested C&CG method is adopted for solving (13) [22]. Specifically, the original problem in (13) is decomposed into an MP and an SP. The MP is: where  $\mathbf{y}_r$  and  $\mathbf{z}_r$  are the additional decision variables.  $\mathbf{u}_r^*$  denotes the found worst-case realization of the uncertain variable  $\mathbf{u}_r$  in the  $r$ th iteration of the SP, which can be expressed as:

$$F^{\text{SP}} : \max_{\mathbf{u} \in \mathbf{U}} \min_{\mathbf{y}, \mathbf{z} \in \Omega(\hat{\mathbf{x}}, \mathbf{u})} \mathbf{b}^T \mathbf{y} \\ s.t. \quad \begin{cases} \mathbf{Fy} + \mathbf{Gz} = \mathbf{f} - \mathbf{Ju} - \mathbf{E}\hat{\mathbf{x}} \\ \mathbf{Hy} + \mathbf{Iz} \geq \mathbf{h} - \mathbf{Ku} - \mathbf{L}\hat{\mathbf{x}} \\ \mathbf{u} \in \mathbf{U}, \mathbf{y} \in \mathbf{Y}, \mathbf{z} \in \mathbf{Z} \end{cases}. \quad (15b)$$

Due to the integer variables  $\mathbf{z}$ , the SP is further decomposed [22]. The outer-level models of the SP can be converted to:

where  $\alpha_s$  and  $\beta_s$  are the dual variables for the first and second row of formula (15b), respectively. The fifth to the ninth row in (15c) is the complementary slackness constraint with the big-M method.  $y_s$ ,  $\alpha_s$  and  $\beta_s$  indicate the additional decision variables at the  $s$ th iteration.  $\hat{z}_s$  comes from the inner-level models, which can be expressed as:

$$F_{\text{inner}}^{\text{SP}} : \min_{\mathbf{z}, \mathbf{y} \in \Theta(\hat{\mathbf{x}}, u^*)} \mathbf{b}^T \mathbf{y} \\ \text{s.t.} \quad \begin{cases} \mathbf{F}\mathbf{y} + \mathbf{G}\mathbf{z} = \mathbf{f} - \mathbf{J}\mathbf{u}^* - \mathbf{E}\hat{\mathbf{x}} \\ \mathbf{H}\mathbf{y} + \mathbf{I}\mathbf{z} \geq \mathbf{h} - \mathbf{K}\mathbf{u}^* - \mathbf{L}\hat{\mathbf{x}} \\ \mathbf{y} \in \mathbf{Y}, \mathbf{z} \in \mathbf{Z} \end{cases}. \quad (15d)$$

To this end, the proposed two-stage RO strategy can be solved through [Algorithm 1](#) with the nested C&CG method as follows:

#### 4. Case study

In this section, a tested case involving a data center park with three DCBs is implemented to verify the effectiveness of the proposed two-stage RO strategy. This case is implemented in Matlab R2016a using the Yalmip platform on a desktop with an Intel Core i9 CPU clocked at 3.0 GHz and 64GB RAM. The optimization model is solved using Gurobi 9.5 solver.

**Algorithm 1**

### The nested C&CG method.

- ```

1: Initialize  $LB = -\infty$ ,  $UB = +\infty$ ,  $k = 1$ ,  $\varepsilon = 0.1$ , and set  $\mathbf{u}_1^*$  with random values;
2: while  $UB - LB \geq \varepsilon \cdot LB$  do
3:   Solve the MP (15a) with  $\mathbf{u}_k^*$  and get the optimal solution  $(\hat{\mathbf{x}}_k, \hat{\eta}_k, \hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_k, \hat{\mathbf{z}}_1, \dots, \hat{\mathbf{z}}_k)$ ;
4:   Update  $LB = \mathbf{a}^\top \hat{\mathbf{x}}_k + \hat{\eta}_k$ ;
5:   Initialize  $SLB = -\infty$ ,  $SUB = +\infty$ ,  $p = 1$ , and set  $\hat{\mathbf{z}}_{k+1}^{(1)} = \hat{\mathbf{z}}_k$ ;
6:   while  $SUB - SLB \geq \varepsilon \cdot SLB$  do
7:     Solve the outer-level problem (15c) of the SP with fixed  $\hat{\mathbf{x}}_k$  and  $(\hat{\mathbf{z}}_{k+1}^{(1)}, \dots, \hat{\mathbf{z}}_{k+1}^{(p)})$ , then get the optimal solution  $(\hat{\mathbf{u}}_{k+1}^{(p)}, \hat{\theta}_{k+1}^{(p)}, \hat{\mathbf{y}}_{k+1}^{(1)}, \dots, \hat{\mathbf{y}}_{k+1}^{(p)})$ ;
8:     Update  $SUB = \hat{\theta}_{k+1}^{(p)}$ ;
9:     Solve the inner-level problem (15d) of the SP with fixed  $\hat{\mathbf{x}}_k$  and  $\hat{\mathbf{u}}_{k+1}^{(p)}$ , then get the optimal solution  $(\hat{\mathbf{y}}_{k+1}^{(p+1)}, \hat{\mathbf{z}}_{k+1}^{(p+1)})$ ;
10:    Update  $SLB = \max\{SLB, \mathbf{b}^\top \hat{\mathbf{y}}_{k+1}^{(p+1)}\}$ ;
11:    Let  $p = p + 1$ ;
12:  end while
13:  return  $\mathbf{u}_{k+1}^* = \hat{\mathbf{u}}_{k+1}^{(p)}$ 
14:  Update  $UB = \min\{UB, \mathbf{a}^\top \hat{\mathbf{x}}_k + SUB\}$ ;
15:  Let  $k = k + 1$ ;
16: end while

```

#### 4.1. Simulation setup

The adopted data center park with three DCBs is illustrated in Fig. 1. The time horizon is set to 24 h with an hourly resolution. The predicted day-ahead and real-time buying and selling electricity prices are shown in Fig. 2 [29]. The PV curve is taken from [30]. The BWs curve at each DCB and IW curve at the front-end server are modified based on reference [31]. The degradation cost of BESS is 0.02 \$/kWh. Other parameters involved in this experiment are provided in Table 1 [32].

The budget  $\Gamma$  of polytopic uncertainty sets for PV, BWs, and IWs are set as 6, 12, and 12, respectively. In this experiment, the uncertainty sets of PV output power, BWs, and IWs are derived from “realistic” historical data using the Gaussian distribution. The number of the sampled data of each uncertain variable is set as 10,000. The confidence level  $\rho$  of each uncertainty set is set as 95%. The confidence level  $\kappa$  used to calculate the radius of the Wasserstein ball is set as 95% as well. Fig. 3 shows the predicted values (blue dashed line), the real values (red solid line), and the uncertainty sets (gray shaded area) of these uncertain variables.

#### 4.2. Case study results and analyses

##### 1) The first-stage strategy

Fig. 4 depicts the predicted values and the optimal values of BWs obtained from the proposed first-stage scheduling strategy (described in Section 2.3) for different DCBs. The following observations and analyses are made from Fig. 4 on a temporal scale.

- Some BWs are scheduled to be processed in earlier or later time slots in a day, e.g., 3:00 and 24:00. This is because the buying electricity prices in these time slots are relatively lower than in other time slots, resulting in a reduction of the electricity cost. In addition, DCBs do not sell electricity to the utility because the PV generations have no output power in these time slots.
- Some BWs are also scheduled to be processed to the time slots with relatively high electricity prices, e.g., 8:00–15:00. It is because more PV is generated in these time slots without any cost even though the electricity prices are not very high from 5:00 to 7:00. In the day-ahead market, PV would be consumed on-site rather than being sold to the main grid at low selling electricity prices unless it is surplus. Therefore, the processed BWs can be powered directly in these time slots by PV without being scheduled to other time slots.
- Among the three DCBs, DCB2 has relatively more BWs to process than DCB1 and DCB3, as shown in Fig. 3(b). Therefore, it stocks too many BWs before 8:00, and it has to process more BWs than predicted from 8:00 to 10:00. Whereas, DCB1 does not need to process BWs at 8:00 with less PV because it stocks fewer BWs before 10:00. DCB3 maintains fewer BWs in stock before 10:00. However, given the three-hour limit on the maximum delay time, it has to process any stocked BWs coming from 5:00–7:00 at 8:00.

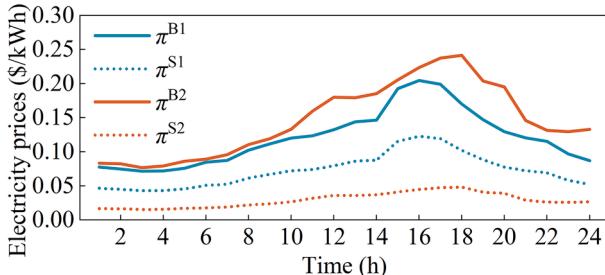


Fig. 2. The predicted day-ahead and real-time buying and selling electricity prices.

Table 1

Summary of parameters involved in the experiment.

| Parameter              | Numerical Value      | Parameter                | Numerical Value     |
|------------------------|----------------------|--------------------------|---------------------|
| $A_{\max}^{\text{DC}}$ | [4000, 4000, 4000]   | $E^{\text{RB}}$          | [400, 300, 500] kWh |
| $p_{\text{idle}}$      | [0.1, 0.1, 0.1] kW   | $SOC_{\min}$             | [0.1, 0.1, 0.1]     |
| $p_{\text{peak}}$      | [0.2, 0.2, 0.2] kW   | $SOC_{\max}$             | [0.9, 0.9, 0.9]     |
| $I^{\text{rate}}$      | [4, 4, 4] requests/s | $\eta^{\text{BC}}$       | [0.95, 0.95, 0.95]  |
| $C^{\text{PUE}}$       | [1.35, 1.4, 1.4]     | $\eta^{\text{BD}}$       | [0.95, 0.95, 0.95]  |
| $C^{\text{DT}}$        | [0.5, 0.5, 0.5] s    | $P_{\max}^{\text{B}}$    | [200, 150, 250] kW  |
| $TD$                   | [4, 5, 3] h          | $P_{\max}^{\text{B2}}$   | [100, 75, 125] kW   |
| $\eta^{\text{DC}}$     | [0.1, 0.1, 0.1]      | $P_{\max}^{\text{Grid}}$ | [1.5, 1.5, 1.5] MW  |

Note: the index  $j$  is neglected in this table.

- When the output power of PV is insufficient or even zero, during the time slot of 16:00–18:00, when the electricity prices are high, BWs are delayed to be processed afterward to reduce electricity costs.

Fig. 5 depicts the arriving IW at the front-end server and the optimal values of allocated IWs at three DCBs. The following observations and analyses are made from Fig. 5 on a spatial scale.

- The front-end server would prefer to allocate more IWs to DCB1 because the PUE index value of DCB1 is the smallest among the three DCBs. It means that DCB1 can process the same workloads with less power consumption.
- Different from general flexible loads, the capabilities of processing BWs and IWs are coupled and restricted by the total amount of servers. Hence, employing more servers for processing BWs would result in a reduced number of available servers for IWs. For example, DCB2 processes more stocked BWs at 9:00–10:00 and 24:00, when the front-end server almost only allocates IWs to DCB1 and DCB3.

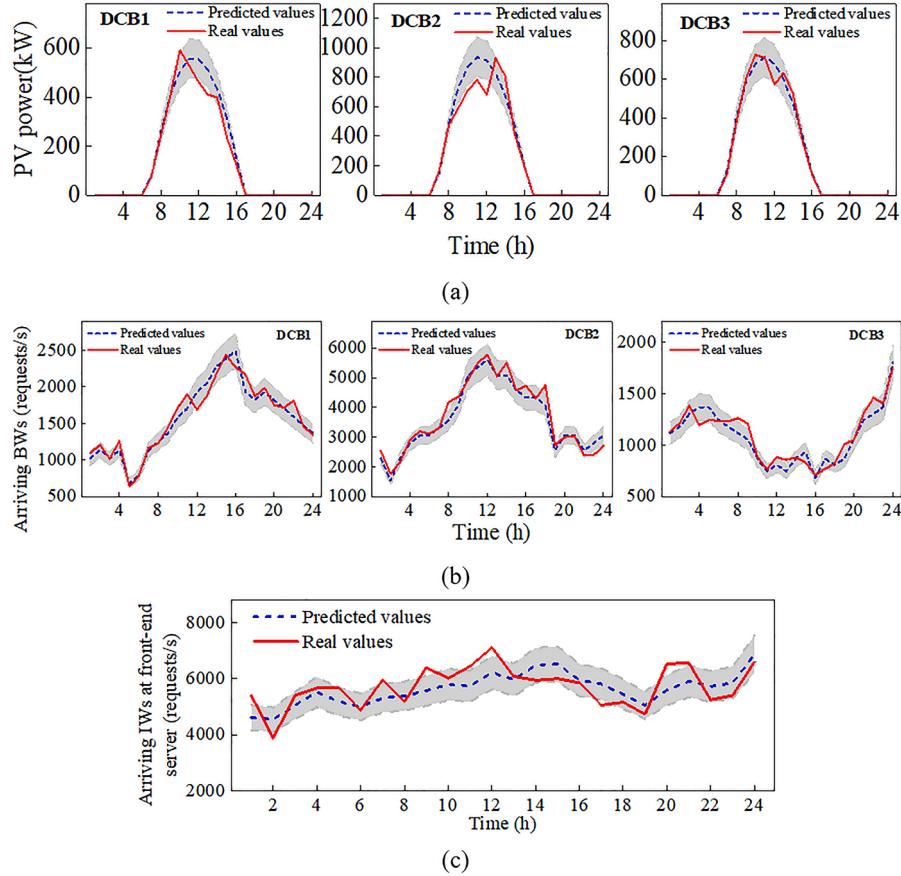
The first-stage power dispatching strategies of three DCBs are provided in Fig. 6. The positive values indicate the provided input power to DCB, while the negative values mean the consumed power.

As the selling electricity prices are lower than the buying electricity prices in the day-ahead market, PV takes the priority of being consumed locally. The following observations from Fig. 6 are made:

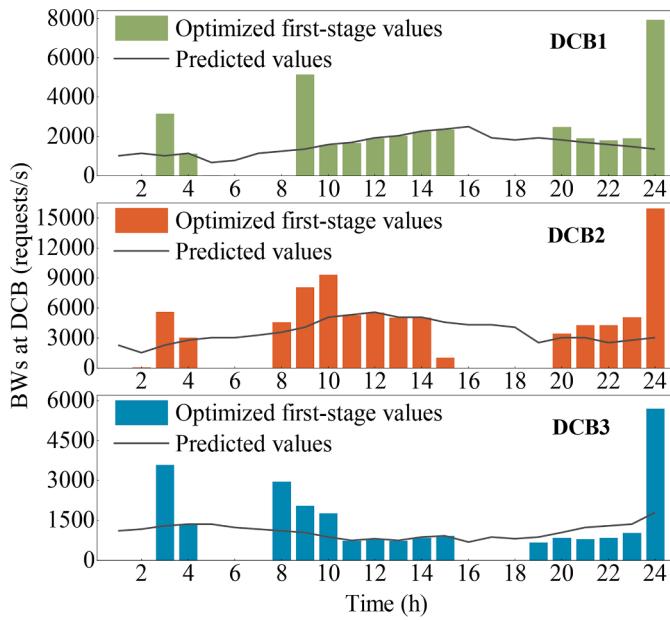
- If PV is insufficient, each DCB buys electricity from the main grid utility or uses the stored power in the BESS. The PUE index of DCB1 is the smallest among the three DCBs. In comparison, DCB1 has to buy more electricity from the main grid before 8:00 and after 16:00 because it has more IWs to process and less PV available. However, DCB2 and DCB3 only need to buy electricity when processing stocked BWs if PV is insufficient.
- The surplus electric power can be stored in the BESS when PV is sufficient. Especially, as DCB3 has fewer workloads to process, the surplus electric power from PV generation can even be sold to the main grid from 10:00 to 14:00.
- When the day-ahead electricity prices are relatively high, e.g. 16:00–17:00, BESS discharges power to reduce buying electricity from the main grid. Furthermore, to ensure sufficient capacity for charging power during the periods that PV is surplus, the BESS at each DCB also discharges power before 7:00, when the electricity prices are not very high.

##### 1) The second-stage strategy

The day-ahead strategy is comprised of the combined strategies from both two stages. The second-stage strategy is based on the set first-stage strategy, which cannot be changed due to the limitation of servers. To



**Fig. 3.** The predicted value, real value, and uncertainty sets of the uncertain variables. (a) PV. (b) BWs. (c) IWS.



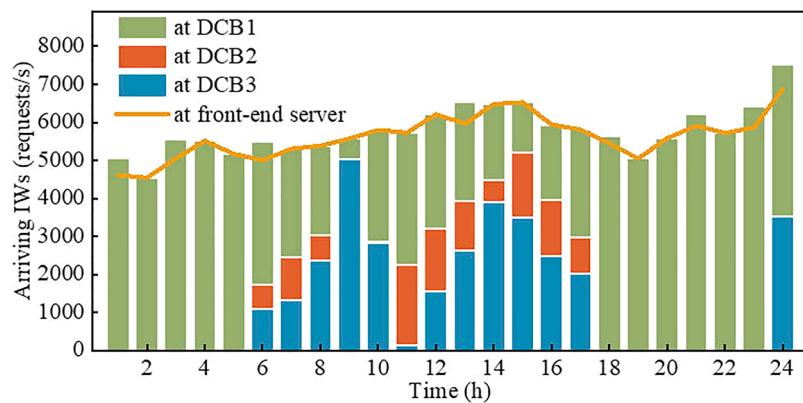
**Fig. 4.** The predicted values and optimal first-stage values of BWs processed at three DCBs.

this end, despite the realization of arriving BWs in the second stage, the BWs in the set first stage cannot be rescheduled to other time slots. It means BWs do not have the capabilities of demand response in the second stage. However, the day-ahead two stages' optimization

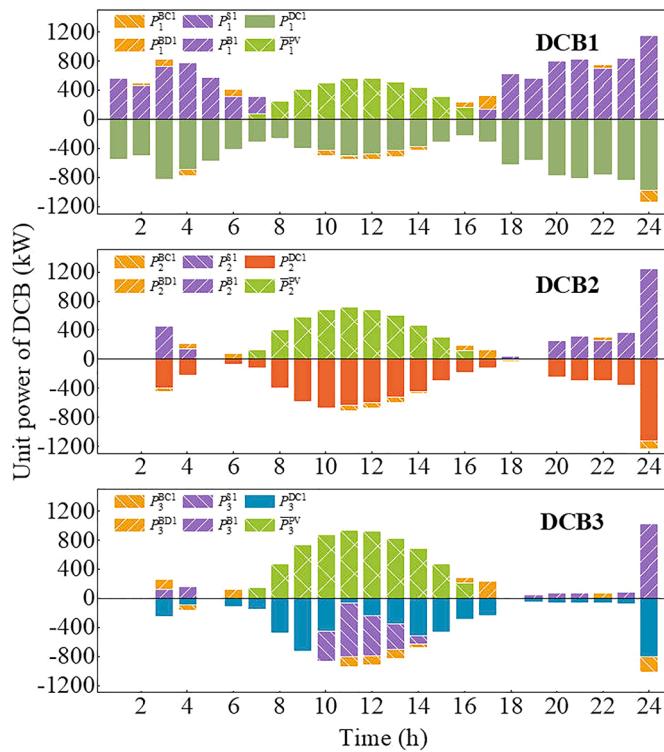
strategies are coupled, as modeled in [Section 2.3](#) and [Section 2.4](#). Considering the realization of uncertain scenarios (e.g. the worst-case scenario) in the second stage, the front-end server has some redundancy at specific time slots, where the total number of allocated IWS exceeds the arriving IWS, as illustrated in [Fig. 5](#).

Redundant servers are deployed to handle deviations in the quantities of uncertain workloads, such as BWs and IWS. [Fig. 7](#) provides the number of two stages' active servers at three DCBs. To adapt to each worst uncertain scenario generated in the inner iteration of [Algorithm 1](#), most of the redundant servers are active with no workloads to process in the first stage. However, the first-stage strategy is not completely satisfied with all the worst uncertain scenarios. Therefore, some remaining redundant servers are active in the second stage. Compared with DCB1 and DCB3 in [Fig. 7](#), DCB2 is equipped with more active redundant servers. It is because more uncertain BWs need to be processed in DCB2, as shown in [Fig. 3\(b\)](#). According to [Fig. 4](#) and [Fig. 5](#), DCB3 has the fewest BWs and IWS to process and therefore does not require a significant number of redundant servers.

Considering the power adjustment of the second stage with the worst-case scenario, the two-stage power dispatching strategy is shown in [Fig. 8](#). The BESS of each DCB accordingly adjusts the charging and discharging strategy to adapt to the updated power consumption of DCBs and PV generations. Compared with the strategy in [Fig. 6](#), the charging power is reduced and the discharging power is increased at some time slots such as 10:00–14:00 in [Fig. 8](#). Meanwhile, the discharging power is reduced before 7:00 because sufficient capacity is reserved in the BESS to charge electric power afterward. It reduces the charging and discharging cost of BESS. The additional insufficient electric power has to be bought in the second stage with real-time electricity prices to fit the uncertainties at some time slots, such as 8:00–15:00.



**Fig. 5.** The arriving IWs at the front-end server and optimal first-stage values of allocated IWs at three DCBs.



**Fig. 6.** The optimal first-stage power dispatching strategy for three DCBs.

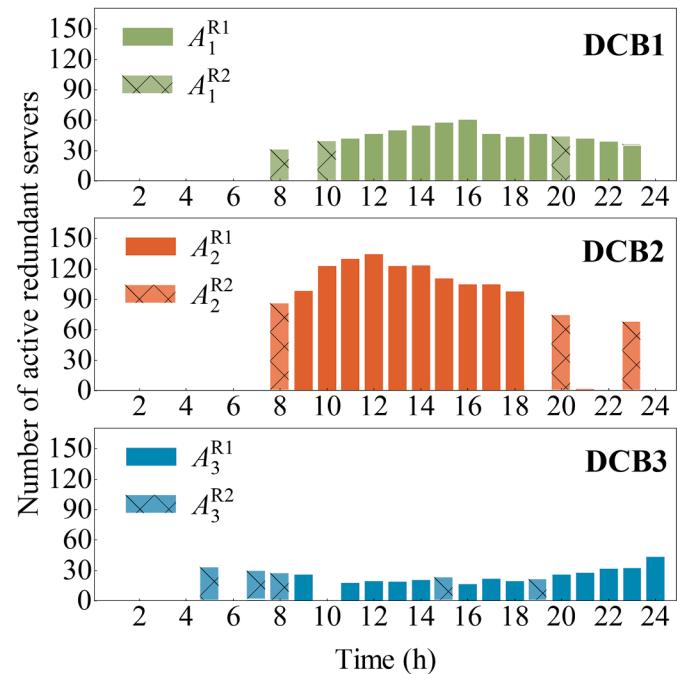
#### 4.3. Comparisons with different demand-side resources

This paper presents a spatial-temporal correlated scheduling model of workloads and electric power. However, due to the limitation of active servers within DCBs, the demand response capacities of workloads are subject to certain constraints. Existing literature has only separately considered the uncertainties associated with BWs [33] and IWs [16]. However, the simultaneous consideration of the uncertainties of both BWs and IWs would make the problem more complex. Therefore, four cases with different demand-side resources using the proposed two-stage RO method are compared as follows:

**Case 1: (none)** IWs in the front-end server are distributed equally to three DCBs. BWs are seen as fixed workloads to be processed at each specific time.

**Case 2: (only IWs)** IWs are demand-side resources that can be scheduled spatially from the front-end server to different DCBs. BWs are seen as fixed workloads to be processed at each specific time.

**Case 3: (only BWs)** BWs are demand-side resources that can be



**Fig. 7.** The number of redundant servers in two stages at three DCBs.

scheduled temporally across time scales. IWs in the front-end server are distributed equally to three DCBs.

**Case 4: (IWs and BWs)** BWs and IWs can be scheduled cooperatively across spatial and temporal scales (the case studies in Section 4.2).

The real values of uncertain variables are used to generate the ex-post results. The economic comparison of ISC under four cases is depicted in Fig. 9. The total electricity costs considering demand response capacities of IWs and BWs in Case 4 decrease by 13.7% compared to Case 1, where no demand-side resources are taken into account. It demonstrates the great advantages of the proposed method in reducing the energy cost of ISC. There are no significant differences between Case 1 and Case 2, or between Case 3 and Case 4, whereas there are obvious differences between Case 1 and Case 3, or Case 2 and Case 4. The results indicate that the demand response capacity of BWs is greater than that of IWs. It is because there is not a significant difference in the PUE metric across three different DCBs. Nonetheless, the PV generation outputs promote DCBs to schedule the processing of BWs across time scales to absorb these renewable energies. In general, the spatially transferrable IWs can be scheduled across different DCBs actively to coordinate with temporally transferrable BWs, thereby taking full

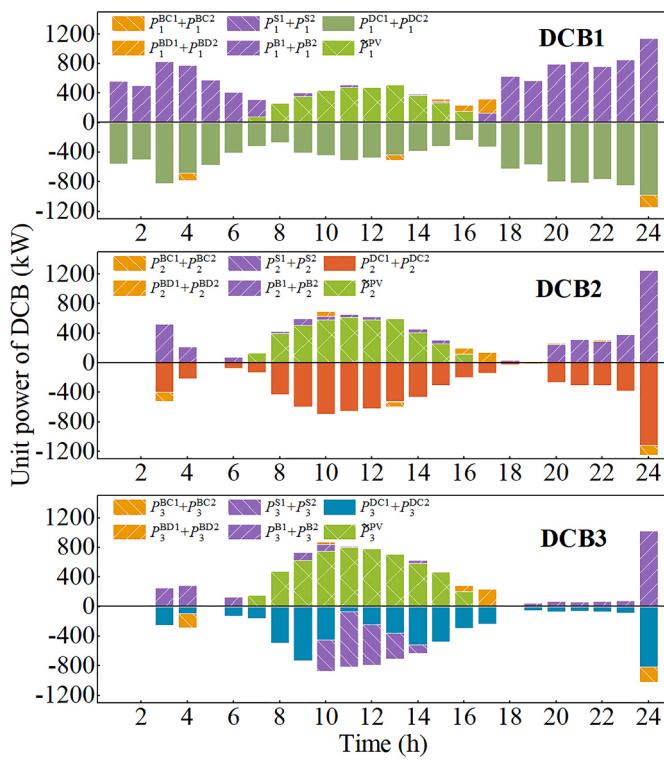


Fig. 8. The two-stage power dispatching strategy for three DCBs.

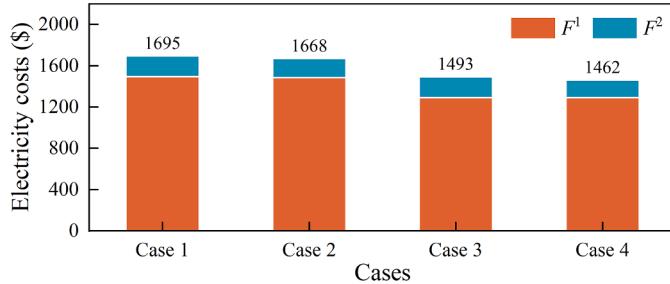


Fig. 9. The economic comparison of ISC with different demand-side resources in four cases.

advantage of active servers and reducing the total energy cost.

#### 4.4. Comparisons under different confidence levels

The risk confidence level  $\rho$  plays an important role in the conservativeness of the data-driven uncertainty sets. In this part, we conduct a sensitivity analysis of the risk confidence level in deciding the uncertainty sets, as illustrated in Fig. 10. We test four different confidence levels (i.e. 85%, 90%, 95%, and 99%) for the proposed data-driven method. Fig. 10(a) and Fig. 10(b) depict ambiguous boundaries of uncertainty sets of PV generation and arriving BWs for each DCB. Fig. 10(c) depicts the arriving IWs at the front-end server. Fig. 10(d) shows the total electricity cost of ISC under different confidence levels.

As shown in Fig. 10(a)–(c), the ambiguous boundary of uncertainty sets  $\Delta u$  expands as the risk confidence level  $\rho$  increases due to the decrease of the tolerance of risk. Meanwhile, Fig. 10(d) shows that the two-stage operation cost of ISC increases as well when the confidence level gets large. It is because the larger uncertainty sets worsen the found ‘worst-case scenario’ when the uncertain variables reach the boundaries, leading to a relatively higher conservativeness. These outcomes indicate that an appropriate confidence level is crucial to strike a balance

between the economic efficiency and conservativeness of the scheduling strategy.

#### 4.5. Comparisons of different optimization strategies

To access the effectiveness of the proposed two-stage RO method, we compare its performance with two other strategies presented in the literature: namely, the DO strategy [13] and the SO strategy [16]. Additionally, a general RO strategy [18] is provided for comparison with the provided data-driven RO strategy.

**Strategy 1 (DO):** The uncertainties of PV, BWs, and IWs are ignored in the scheduling strategy. In other words, the deterministically predicted values of PV, BWs, and IWs are adopted.

**Strategy 2 (SO):** This strategy addresses the uncertainties associated with PV, BWs, and IWs using the stochastic programming method. To generate a sufficient number of samples, the uncertain parameters are assumed to follow a Gaussian distribution, from which 10,000 samples are generated. These samples are then clustered into 10 typical scenarios using the k-means method.

**Strategy 3 (General RO):** The provided two-stage RO strategy considers uncertainties arising from PV, BWs, and IWs, which are modeled using polyhedron uncertainty sets, with the boundaries of these sets determined subjectively. Specifically, the upper and lower boundaries of PV and workloads (BW or IW) used in this part are all set at a 15% deviation from the predicted values.

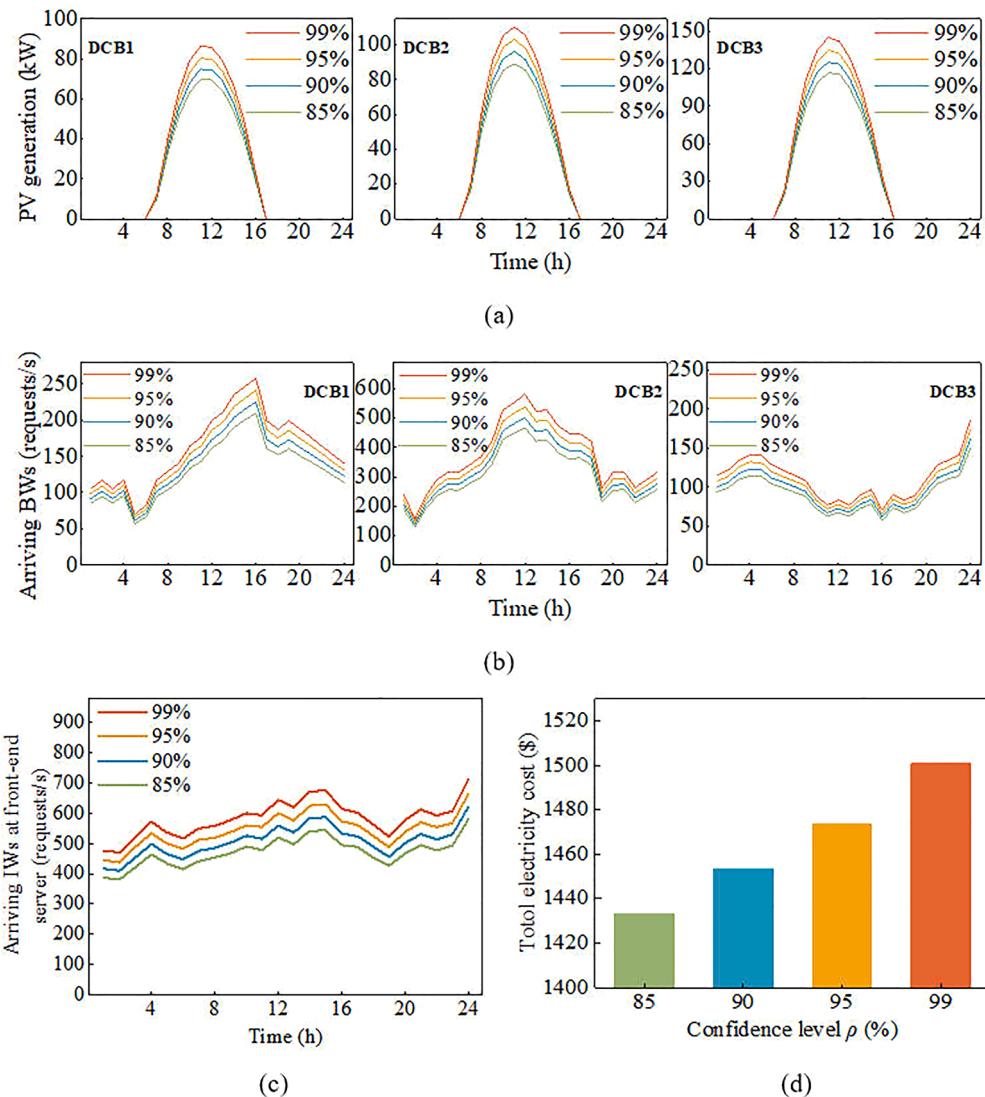
**Strategy 4 (Data-driven RO):** The provided two-stage RO strategy considers uncertainties arising from PV, BWs, and IWs, which are modeled using polyhedron uncertainty sets, with the data-driven boundaries described in Section 3.

The economic comparison of these four strategies with three scenarios (as described below) is presented in Table 2.  $F^1$ ,  $F^2$ , and  $F$  represent the first-stage, second-stage, and total electricity costs, respectively. The post-performance of different strategies is examined based on the day-ahead scheduling decisions obtained in these three scenarios.

- Scenario I provides ex-post results based on the predicted values of all uncertain variables.
- Scenario II provides the ex-post results in the presence of real values of uncertain variables depicted in Fig. 3.
- Scenario III shows ex-post results under the worst-case values of uncertain variables obtained in the proposed two-stage RO strategy with data-driven uncertainty sets.

As shown in Table 2, the total cost  $F$  of DO in Scenario I is the lowest among the four strategies, while the cost of general RO is the highest. This is because the DO strategy is optimistic by assuming all uncertain variables as known and equal to the predicted values. While the RO strategies, including general RO and data-driven RO, are conservative considering many worst-case scenarios in each iteration. As the boundaries of the uncertainty set used in the data-driven RO strategy are derived from historical data, which is a more objective approach, the cost associated with this method is lower compared to the general RO strategy. The SO strategy also takes uncertain scenarios into account, but it does not consider the worst-case scenarios. As a result, the total cost of the SO strategy in Scenario I falls between the cost of the DO strategy and RO strategies.

Once the worst-case realization of uncertainties occurs, the second-stage electricity cost of the DO with adjustable strategy, as shown in Scenario III, is significantly large. In comparison, the electricity cost of the data-driven RO strategy is the lowest because the worst-case scenario is considered in the day-ahead decision-making process. SO strategy is a compromise strategy in Scenario III and the performance is close to the DO strategy. This is because the worst-case scenarios are not involved in the typical scenarios of SO. The total cost associated with the general RO strategy is higher than that of the proposed data-driven RO



**Fig. 10.** The sensitivity analysis of the risk confidence level of the proposed data-driven method.

(a) the ambiguous boundary of PV's uncertainty sets. (b) the ambiguous boundary of BWs' uncertainty sets. (c) the ambiguous boundary of IWs' uncertainty sets. (d) the total electricity cost of ISC.

**Table 2**  
Economic comparisons of three strategies (\$).

| Strategy       | $F^1$ | Scenario I |      | Scenario II |      | Scenario III |      |
|----------------|-------|------------|------|-------------|------|--------------|------|
|                |       | $F^2$      | $F$  | $F^2$       | $F$  | $F^2$        | $F$  |
| DO             | 1159  | 0          | 1159 | 336         | 1495 | 403          | 1562 |
| SO             | 1176  | 1          | 1177 | 293         | 1469 | 368          | 1544 |
| General RO     | 1342  | -10        | 1332 | 148         | 1490 | 186          | 1528 |
| Data_driven RO | 1291  | 0          | 1291 | 171         | 1462 | 184          | 1475 |

strategy, indicating that the data-driven RO strategy is less conservative.

In the presence of the real values of uncertain variables in Scenario II, the second-stage costs  $F^2$  in real-time and the total costs  $F$  of SO and two RO strategies are also lower than that of DO because uncertain factors are taken into account to optimize the day-ahead decisions. It makes the day-ahead decisions more robust to the difference between the real values and the predicted values. With the given real values, the total cost of RO is lower than SO. However, if the real values are closer to Scenario I with predicted values, the SO method becomes more cost-effective compared to the RO method. When the real values are closer to Scenario III with worst-case values, the RO method proves to be superior to the SO method in the economy.

#### 4.6. Impacts of robust budget parameters

The budget parameter  $\Gamma$  is used to adjust the conservativeness of the proposed two-stage RO strategy. Table 3 presents five sets of budget parameters for comparison with two scenarios to evaluate the efficiency of the budget parameters. Scenario II uses the real values of uncertain variables, as described in Section 4.5. Scenario IV shows ex-post results under the worst-case values of uncertain variables obtained in the proposed two-stage RO strategy with different  $\Gamma$ .

With the increase of  $\Gamma$ , the total operating costs based on the worst-case values in Scenario IV increase accordingly. This evidence indicates

**Table 3**  
Economic comparisons of different  $\Gamma$  (\$).

| $\Gamma$ | $F^1$                                         | Scenario II |     | Scenario IV |     |
|----------|-----------------------------------------------|-------------|-----|-------------|-----|
|          |                                               | $F^2$       | $F$ | $F^2$       | $F$ |
| 1        | $\Gamma^{PV}=0, \Gamma^{BW}=\Gamma^{FIW}=0$   | 1159        | 336 | 1495        | 0   |
| 2        | $\Gamma^{PV}=4, \Gamma^{BW}=\Gamma^{FIW}=6$   | 1250        | 207 | 1457        | 131 |
| 3        | $\Gamma^{PV}=6, \Gamma^{BW}=\Gamma^{FIW}=12$  | 1291        | 171 | 1462        | 184 |
| 4        | $\Gamma^{PV}=8, \Gamma^{BW}=\Gamma^{FIW}=18$  | 1319        | 163 | 1482        | 233 |
| 5        | $\Gamma^{PV}=10, \Gamma^{BW}=\Gamma^{FIW}=24$ | 1325        | 171 | 1496        | 250 |

that a larger  $\Gamma$  reflects a more conservative approach, resulting in increased PV generation reaching the lower bound and more workloads reaching the upper bound. To ensure processing more workloads (i.e. BWs and IWs) with limited PV generation in the worst-case scenario, the ISC would procure additional electricity from the day-ahead electricity market as a mitigation measure to counteract the effects of rising uncertainty.

Regarding the ex-post performance of the RO strategy with various budget parameters, in the presence of the real values of uncertain parameters, the RO strategy with  $\Gamma=0$  exhibits the highest second-stage cost  $F^2$  in real time. Nevertheless, the total cost  $F$  is not consistently higher when using different budget parameters. For instance, the total cost with real data based on the fifth set of  $\Gamma$  is more than that of DO with the first set, indicating an excessive level of conservativeness. Hence, the budget parameters should be set appropriately to balance the economy and the conservativeness of the proposed strategy.

## 5. Conclusion

This paper proposes a two-stage RO strategy for DCBs managed by an ISC. The proposed strategy takes into account the coordinated scheduling of workloads and electricity in the data center park in multiple sources of uncertainties. The spatial-temporal demand-side resources such as BESS, BWs, and IWs are leveraged to minimize the overall cost of electricity consumption. A data-driven method is proposed to determine the ambiguous boundaries of uncertainty sets based on the 1-Wasserstein metric. Eventually, the proposed two-stage RO model with a mixed-integer inner subproblem is solved using the nested C&CG algorithm.

Case studies verified the proposed method, and the following conclusions are drawn:

- 1) The inclusion of BWs and IWs as spatially and temporally transferable demand-side resources in DCBs holds significant potential, as they can respond to fluctuations in market electricity prices.
- 2) The regulation of uncertainty budget parameters based on data-driven uncertainty sets for the two-stage RO problem can be a vital tool for operators to balance conservativeness and the economy.
- 3) The proposed RO strategy offers superior performance when compared to DO and SO strategies, as it can better accommodate real scenarios close to the worst cases without requiring the probability distribution of uncertain parameters. The data-driven method also reduces the conservatism of the problem than general RO.
- 4) The increase in the risk level of obtaining the boundaries of uncertainty sets with the data-driven method makes the RO strategy more conservative.

In reality, DCBs in a data center park can be managed by more than one ISC. Future research will focus on exploring the potential benefits of cooperation between different ISCs in terms of energy management and transactions within the distribution network of the park.

## CRediT authorship contribution statement

**Zhihao Yang:** Conceptualization, Methodology, Software, Writing – original draft. **Anupam Trivedi:** Writing – review & editing, Visualization. **Haoming Liu:** Investigation, Validation. **Ming Ni:** Writing – review & editing, Supervision. **Dipti Srinivasan:** Supervision.

## 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

Data will be made available on request.

## Acknowledgements

This work was supported in part by the 2021 Graduate Research and Innovation Projects of Jiangsu Province, China under grant KYCX21\_0473 and the China Scholarship Council (CSC) program under project number 202106710110.

## References

- [1] Z. Liu, H. Yu, R. Liu, M. Wang, C. Li, Configuration optimization model for data-center-park-integrated energy systems under economic, reliability, and environmental considerations, *Energies* 13 (2020) 448.
- [2] E. Masanet, A. Shehabi, N. Lei, S. Smith, J. Koomey, Recalibrating global data center energy-use estimates, *Science* 367 (2020) 984–986.
- [3] C. Nadjahi, H. Louahla, S. Lemasson, A review of thermal management and innovative cooling strategies for data center, *Sustain. Comput. Inform. Syst.* 19 (2018) 14–28.
- [4] Q. Tian, Q. Guo, S. Nojavan, X. Sun, Robust optimal energy management of data center equipped with multi-energy conversion technologies, *J. Clean. Prod.* 329 (2021), 129616.
- [5] Z. Ding, L. Xie, Y. Lu, P. Wang, S. Xia, Emission-aware stochastic resource planning scheme for data center microgrid considering batch workload scheduling and risk management, *IEEE Trans. Ind. Appl.* 54 (2018) 5599–5608.
- [6] M. Chen, C. Gao, M. Song, S. Chen, D. Li, Q. Liu, Internet data centers participating in demand response: a comprehensive review, *Renew. Sustain. Energy Rev.* 117 (2020), 109466.
- [7] M. Chen, C. Gao, M. Shahidehpour, Z. Li, S. Chen, D. Li, Internet data center load modeling for demand response considering the coupling of multiple regulation methods, *IEEE Trans. Smart Grid*, 12 (2021) 2060–2076.
- [8] Z. Yang, M. Ni, H. Liu, Pricing strategy of multi-energy provider considering integrated demand response, *IEEE Access* 8 (2020) 149041–149051.
- [9] X. Ge, L. Shi, Y. Fu, S.M. Muyeen, Z. Zhang, H. He, Data-driven spatial-temporal prediction of electric vehicle load profile considering charging behavior, *Electr. Power Syst. Res.* 187 (2020), 106469.
- [10] B. Tan, H. Chen, X. Zheng, J. Huang, Two-stage robust optimization dispatch for multiple microgrids with electric vehicle loads based on a novel data-driven uncertainty set, *Int. J. Electr. Power Energy Syst.* 134 (2022), 107359.
- [11] J. Luo, L. Rao, X. Liu, Spatio-temporal load balancing for energy cost optimization in distributed internet data centers, *IEEE Trans. Cloud Comput.* 3 (2015) 387–397.
- [12] H. Ji, S. Chen, H. Yu, P. Li, J. Yan, J. Song, C. Wang, Robust operation for minimizing power consumption of data centers with flexible substation integration, *Energy* 248 (2022), 123599.
- [13] C. Jiang, C. Li Tseng, Y. Wang, Z. Lan, F. Wen, F. Chen, L. Liang, Optimal pricing strategy for data center considering demand response and renewable energy source accommodation, *J. Mod. Power Syst. Clean Energy*, 11 (2023) 345–354.
- [14] W.-Y. Chiu, W.-K. Hsieh, C.-M. Chen, Y.-C. Chuang, Multiobjective demand response for internet data centers, *IEEE Trans. Emerg. Top. Comput. Intell.* 6 (2022) 365–376.
- [15] H. Wang, J. Huang, X. Lin, H. Mohsenian-Rad, Proactive demand response for data centers: a win-win solution, *IEEE Trans. Smart Grid*, 7 (2016) 1584–1596.
- [16] T. Niu, B. Hu, K. Xie, C. Pan, H. Jin, C. Li, Spatial coordination between data centers and power system considering uncertainties of both source and load sides, *Int. J. Electr. Power Energy Syst.* 124 (2021), 106358.
- [17] Z. Ding, Y. Cao, L. Xie, Y. Lu, P. Wang, Integrated stochastic energy management for data center microgrid considering waste heat recovery, *IEEE Trans. Ind. Appl.* 55 (2019) 2198–2207.
- [18] T. Chen, Y. Zhang, X. Wang, G.B. Giannakis, Robust workload and energy management for sustainable data centers, *IEEE J. Sel. Areas Commun.*, 34 (2016) 651–664.
- [19] C. Duan, W. Fang, L. Jiang, L. Yao, J. Liu, Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric, *IEEE Trans. Power Syst.* 33 (2018) 4924–4936.
- [20] Z.-P. Yuan, P. Li, Z.-L. Li, J. Xia, Data-driven risk-adjusted robust energy management for microgrids integrating demand response aggregator and renewable energies, *IEEE Trans. Smart Grid*, 14 (2023) 365–377.
- [21] B. Zeng, L. Zhao, Solving two-stage robust optimization problems using a column-and-constraint generation method, *Oper. Res. Lett.* 41 (2013) 457–461.
- [22] L. Zhao, B. Zeng, An exact algorithm for two-stage robust optimization with mixed integer recourse problems, (2012).
- [23] D. Wang, J. Qiu, L. Reedman, K. Meng, L.L. Lai, Two-stage energy management for networked microgrids with high renewable penetration, *Appl. Energy*, 226 (2018) 39–48.
- [24] L. Ju, J. Wu, H. Lin, Q. Tan, G. Li, Z. Tan, J. Li, Robust purchase and sale transactions optimization strategy for electricity retailers with energy storage system considering two-stage demand response, *Appl. Energy*, 271 (2020), 115155.

- [25] X. Chen, Y. Liu, B. Li, Adjustable robust optimization in enabling optimal day-ahead economic dispatch of CCHP-MG considering uncertainties of wind-solar power and electric vehicle, *J. Ind. Manag. Optim.* 17 (2021) 1639.
- [26] X. Lu, P. Zhang, K. Li, F. Wang, Z. Li, Z. Zhen, T. Wang, Data center aggregators' optimal bidding and benefit allocation strategy considering the spatiotemporal transfer characteristics, *IEEE Trans. Ind. Appl.* 57 (2021) 4486–4499.
- [27] J. Wu, Y. Liu, X. Chen, C. Wang, W. Li, Data-driven adjustable robust day-ahead economic dispatch strategy considering uncertainties of wind power generation and electric vehicles, *Int. J. Electr. Power Energy Syst.* 138 (2022), 107898.
- [28] Y. Sun, R. Qiu, M. Sun, Optimizing decisions for a dual-channel retailer with service level requirements and demand uncertainties: a Wasserstein metric-based distributionally robust optimization approach, *Comput. Oper. Res.* 138 (2022), 105589.
- [29] New York independent system operator. 2022-08-29. Available: <https://www.nyiso.com/energy-market-operational-data>.
- [30] XIHE Energy. 2000-01-20. Available: <https://www.xihe-energy.com/>.
- [31] L. Gu, D. Zeng, A. Barnawi, S. Guo, I. Stojmenovic, Optimal task placement with QoS constraints in geo-distributed data centers using DVFS, *IEEE Trans. Comput.* 64 (2015) 2049–2059.
- [32] Z. Yang. Two-stage robust optimization strategy for spatially-temporally correlated data centers with data-driven uncertainty sets. 2022-12-29. Available: <https://dx.doi.org/10.21227/bcnz-s159>.
- [33] Y. Lian, Y. Li, Y. Zhao, C. Yu, T. Zhao, L. Wu, Robust multi-objective optimization for islanded data center microgrid operations, *Appl. Energy.* 330 (2023), 120344.