This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

Optimal Planning of Data Center Coupled with On-site Electricity–Heat–Cooling Conversion

Kang Wang, *Graduate Student Member, IEEE*, Hongxun Hui, *Senior Member, IEEE*, Chengjin Ye, *Member, IEEE*, and Yi Ding, *Member, IEEE*

***Abstract*—Amid escalating energy demands from data centers (DCs), waste heat recovery (WHR) emerges as a critical approach for enhancing energy sustainability and reducing carbon emissions. Existing WHR solutions for DCs remain limited in practice, constrained by limited heat-delivery distance and highly seasonal heat demand. To overcome these limitations, this paper develops an optimal planning framework for a hybrid cooling system, in which the adsorption chiller (ADS) operates in coordination with workload scheduling to enable on-site electricity-heat-cooling conversion. Firstly, an on-site electricity-heat-cooling operation model is developed to describe the coupling among server heat generation, waste heat recovery, and heat-driven refrigeration. Subsequently, a planning strategy for hybrid cooling systems is proposed to minimize the total annualized cost, including annualized investment cost and operation cost, via a discrete-difference-based Benders decomposition algorithm. Furthermore, a tailored reformulation method is designed to improve the tractability of the on-site electricity-heat-cooling operation model. Temperature-dependent ADS characteristics and related nonlinearities are modeled using two-dimensional SOS2 approximation to reformulate the problem into a mixed-integer linear problem for computational efficiency. Case studies demonstrate the effectiveness of the proposed planning strategy in realizing energy savings and carbon emission reductions.**

***Index Terms*—Data center, waste heat recovery, adsorption chillers, on-site electricity-heat-cooling conversion, planning strategy.**

I. Introduction

THE rapid expansion of artificial intelligence, cloud computing, and digital services has significantly increased the electricity consumption of data centers (DCs) [1]. Recent projections indicate that global DC electricity consumption could more than double by 2030, reaching approximately 945 TWh [2]. Meanwhile, in some regions, the accelerating growth of DC loads is outpacing local grid expansion, increasing supply pressure and limiting further DC deployment [3]. Under these conditions, improving DC energy efficiency is important for relieving local grid stress and enhancing operational sustainability. Since most of the electricity consumed by DCs is ultimately converted into heat, a large amount of low-grade waste heat is continuously generated during operation [4]. Recovering this waste heat offers a promising pathway to improve DC energy utilization.

This work is funded by the National Natural Science Foundation of China (52407075), the Science and Technology Development Fund, Macao SAR (0089/2026/RIB1, 0050/2025/AIJ, 001/2024/SKL, and 0002/2025/EQP), and the XPLORER PRIZE.

K. Wang and H. Hui (Corresponding author) are with the State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macao, 999078 China (email: yc47979@um.edu.mo, hongxunhui@um.edu.mo). C. Ye and Y. Ding are with the College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China (email: yechenjing@zju.edu.cn, yiding@zju.edu.cn).

Waste heat in DCs has characteristics distinct from conventional industrial sectors. Industrial waste heat from steel, cement, and chemical processes is often available at higher temperatures and can support a wider range of thermal applications. DC waste heat is typically lower in temperature and more closely tied to the thermal management of computing facilities. Therefore, waste heat recovery (WHR) pathways for DCs should be selected according to their specific heat-source characteristics. To harness the waste heat in DCs effectively, a wide array of waste heat recovery (WHR) technologies has been investigated. For example, reference [5] upgraded low-grade waste heat via heat pumps for district heating. A flexible workload management strategy was developed in [6] to enhance the on-site heat supply during cold waves. The compatibility of low-grade heat with emerging desalination technologies was discussed in [7]. Reference [8] assessed the feasibility of ORC deployment in DCs under varying ambient temperatures, electricity prices, and carbon costs. Despite these advances, the above waste heat utilization pathways exhibit limitations in universal applicability and overall efficiency. District heating exhibits highly seasonal demand [9] and is primarily viable in cold regions [10]. Desalination is limited by thermal losses over long-distance heat delivery [11]. Organic Rankine cycle systems generally have limited thermal-to-electric conversion efficiency for low-temperature heat sources [12]. These limitations indicate that DC-oriented WHR should place greater emphasis on heat-source quality and the spatiotemporal matching between heat supply and heat demand. Collectively, these studies do not yet offer a broadly applicable and high-efficiency solution for utilizing the large volumes of waste heat generated by DCs.

Based on the above characteristics, a more intrinsically compatible pathway is to use on-site waste heat to meet DC cooling loads. This is attractive because cooling accounts for 20–30% of total energy consumption [13]. Unlike traditional vapor compression refrigeration (VCR), thermally driven refrigeration technologies use heat to release and regenerate the refrigerant from the adsorbent/absorbent, thereby driving the cooling cycle and allowing waste heat to serve as the primary energy input [14]. By utilizing waste heat for cooling, DCs can realize on-site electricity–heat–cooling conversion, thereby broadening the applicability of WHR and improving energy utilization. Specifically, server electricity consumption is ultimately dissipated as heat, and the recovered heat is then converted into cooling to meet the DC cooling demand. Thermally driven cooling technologies are mainly categorized into absorption refrigeration and adsorption refrigeration. Absorption refrigeration suffers from low coefficient of performance (COP) and requires driving temperatures above 65 °C, which makes it incompatible with the low-grade waste heat in

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

TABLE I: Nomenclature of Main Symbols

Symbol	Description	Symbol	Description
A. Sets and indices			
$\mathcal{T}$	Set of scheduling periods	t	Indices of scheduling periods
$\mathcal{M}$	Set of operation scenarios	m	Index of operation scenario
$\mathcal{L}$	Set of DVFS frequency levels	l	Index of DVFS frequency level
$\mathcal{I}, \mathcal{J}$	Sets of breakpoints for ADS operation temperature	i, j	Indices of breakpoints for ADS operation temperature
$\mathcal{A}, \mathcal{B}$	Sets of breakpoints for generator heat input and COP	a, b	Indices of breakpoints for generator heat input and COP
$\mathcal{V}, \mathcal{R}$	Sets of breakpoints for air-flow and condenser-temperature	v, r	Indices of breakpoints for air-flow and condenser-temperature
$\mathcal{K}$	Set of Benders iterations	k	Index of Benders iteration
B. Parameters			
D	Number of days represented by each scenario	π_t^{elec}	Electricity price in period t
π_t^{carbon}	Carbon price in period t	γ_t	Carbon intensity of electricity in period t
c^{ADS}	Unit investment cost of one ADS unit	c^{VCR}	Unit investment cost of VCR capacity
r^{disc}	Discount rate for annualized investment-cost calculation	L	Service life of equipment
λ_t^{inter}	Arrival rate of interactive workloads in period t	$\lambda_t^{batch,0}$	Original arrival rate of batch workloads in period t
τ_t^{inter}	Response-time requirement of interactive workloads	τ_t^{batch}	Delay tolerance of batch workloads
f_l	Server operating frequency at DVFS level l	$\bar{\mu}_l$	Per-server processing speed at DVFS level l
P^{idle}	Idle power of one active server	κ^{DVFS}	Dynamic-power coefficient for DVFS operation
S^{max}	Maximum number of available servers	β	Proportion of generated heat removed by liquid cooling
r^{red}	Redundancy factor for capacity planning of cooling system	$\rho_w, c_{p,w}$	Density and specific heat capacity of water
ρ_{air}	Density of air	ε_{hex}	Effectiveness of the plate heat exchanger
ε_{ct}	Effectiveness of the cooling tower	COP^{LC}	COP of the liquid-cooling system
COP^{VCR}	COP of the VCR system	η_{fan}, K_a^{ct}	Efficiency of cooling-tower fan and coefficient of fan power
$P^{ADS,aux}$	Auxiliary power of one active ADS unit	h_{amb}	Enthalpy of ambient air
$\dot{V}^{liq}$	Volumetric flow rate of the liquid-cooling loop	$\dot{V}^{gen}$	Volumetric flow rate of the ADS generator-side loop
$\dot{V}^{cond}$	Volumetric flow rate of the ADS condenser-side loop	$\dot{V}^{evap}$	Volumetric flow rate of the ADS evaporator-side loop
M^{srv}, M^{gen}, M^{evap}, M^{ct}	Big-M constants for server and ADS operation constraints	$Q^{air,peak}$	Peak air-cooling demand used for capacity planning
ΔQ^{VCR}	Step size for VCR-capacity discrete difference	ϵ^{BD}	Convergence tolerance of the Benders decomposition
C. Variables			
θ^{plan}, θ^{sch}	Set of planning variables and scheduling variables	C^{inv}, C^{ope}	Annualized investment cost and operation cost
N_{plan}^{ADS},Q_{plan}^{VCR}	Planned number of ADS and planned VCR capacity	P_t^{SRV}	Power consumption of server racks in period t
P_t^{LC}	Power consumption of the liquid-cooling system in period t	P_t^{VCR}	Power consumption of the VCR system in period t
P_t^{ADS}	Auxiliary power consumption of ADS units in period t	P_t^{fan}	Power consumption of cooling-tower fan in period t
$\delta_{t_1\to t}^{batch}$	Batch workload shifted from period t_1 to period t	λ_t^{batch}	Batch workload processed in period t after scheduling
S_t^{inter}	Number of servers assigned to interactive workloads	S_t^{batch}	Number of servers assigned to batch workloads
S_t	Total number of active servers in period t	f_t	Selected server operating frequency in period t
μ_t	Selected per-server processing speed in period t	$z_{t,l}$	Binary variable for selecting DVFS level l
$N_{t,l}$	Number of active servers operating at DVFS level l	N_t^{ADS}	Number of active ADS units in period t
δ_t^{ADS}	Binary operating-state variable of the ADS system	Q_t^{liq}	Heat load removed by liquid cooling in period t
Q_t^{air}	Air-cooling demand served by ADS and VCR in period t	Q_t^{VCR}	Cooling output supplied by the VCR system in period t
q_t^{gen}	Generator heat input of one ADS unit in period t	q_t^{evap}	Evaporator cooling output of one ADS unit in period t
q_t^{cap}	Maximum evaporator cooling capacity of one ADS unit	Q_t^{ct}	Heat rejected by the cooling tower in period t
$T_t^{liq,in/out}$	Inlet/outlet temperature of the liquid-cooling loop	$T_t^{gen,in/out}$	Inlet/outlet temperature of the ADS generator-side loop
$T_t^{cond,in/out}$	Inlet/outlet temperature of the ADS condenser-side loop	$T_t^{evap,in/out}$	Inlet/outlet temperature of the ADS evaporator-side loop
$\dot{V}_t^{ct}$	Air volumetric flow rate through the cooling tower	Δh_t	Enthalpy difference between saturated outlet air and ambient air
COP_t^{ADS}	COP of the ADS in period t	$\lambda_{i,j,t}^{ADS}$	SOS2-2D interpolation weight for ADS performance
$\lambda_{a,b,t}^{hc}$	SOS2-2D interpolation weight for heat-to-cooling coupling	$\lambda_{v,t}^{ct}$, $\lambda_{v,r,t}^{hr}$	SOS2 interpolation weights for fan power and heat rejection
η	Auxiliary variable approximating the operation cost	μ_N^k, μ_Q^k	Cut coefficients for ADS-unit number and VCR capacity
UB^k, LB^k	Upper and lower estimation at iteration k	Gap^k	Relative convergence gap at iteration k

DCs [15]. Moreover, absorption refrigeration entails practical concerns such as corrosiveness and large space footprint [16]. In contrast, an adsorption chiller (ADS) operates effectively at driving temperatures as low as 50–60 °C, making it better suited for the waste heat in DCs [17]. Recent techno-economic analyses have validated ADS as a promising approach for energy savings and emission reductions in DCs [18]. In [19], a hybrid cooling architecture is proposed where waste heat from liquid-cooled racks drives an ADS, and the chilled water is then supplied to air-cooled racks for reducing the electricity load of VCR. Reference [20] developed a high-power-density ADS using an encapsulated composite adsorbent, enabling stable cooling driven by ultra-low-grade waste heat. In [21], a numerical model is developed to predict ADS performance. However, existing studies mainly focus on ADS component performance or techno-economic assessment. The role of ADS in system-level coordination that incorporates the on-site electricity–heat–cooling conversion has received less attention. This omission can lead to a mismatch between waste-heat supply and cooling demand, limiting energy savings and long-term economic performance.

To address these limitations, system-level selection and capacity planning of relevant energy equipment are necessary. This is because DC energy efficiency depends on how waste heat is utilized within the on-site electricity–heat–cooling conversion. Existing research on DC planning can be broadly grouped into single-energy approaches and multi-energy approaches. Single-energy approaches primarily focus on electrical infrastructure and associated flexibility resources. For example, reference [22] developed a two-stage planning framework for DCs that simultaneously optimizes equipment capacity and geographical placement within the power grid.

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

Reference [23] proposed a multi-objective planning method that jointly determines the DCs and energy storage systems in smart grids. Multi-energy approaches account for energy interactions and enable coordinated optimization across multiple energy carriers. Reference [24] proposed a planning method for regional integrated electricity–heat systems with DCs, incorporating a heat-recovery model to meet district heating demand. Reference [25] introduced a multi-stage coordinated planning framework for multiple energy hubs considering the spatiotemporal workload regulation of DCs. Nevertheless, electricity-centric planning overlooks electricity–cooling–heat coupling in DCs, thereby underexploiting the potential of multi-energy coordination and WHR. Although multi-energy planning studies offer a more comprehensive perspective, WHR is still modeled as an exogenous option with fixed efficiency and ADS-enabled waste-heat-driven cooling is rarely incorporated. Such simplification fails to capture the on-site electricity–heat–cooling conversion in DC operation, thereby limiting improvements in system-level energy efficiency.

Therefore, achieving efficient utilization of waste heat requires a DC planning model that captures the on-site electricity–heat–cooling conversion. A practical implementation is a hybrid ADS-VCR architecture, in which recovered waste heat drives the ADS to produce chilled water, and the VCR system supplies the remaining cooling demand. Developing such an integrated planning model is nontrivial due to the following challenges.

1) On-site electricity–heat–cooling conversion in DC involves several interdependent processes, including workload scheduling, thermally driven refrigeration, cooling supply and heat rejection. Such complex coupling makes system-level modeling nontrivial.
2) The nonlinear characteristics of ADS and cooling towers, together with the planning–operation coupling in capacity planning, hinder the tractable solution of the optimal planning problem.

To address these challenges, this paper proposes an optimal planning framework for hybrid ADS-VCR cooling systems in DC. The framework embeds the on-site electricity–heat–cooling conversion into a DC operation model and optimizes equipment capacities under varying workloads. The main contributions are summarized as follows.

1) An on-site electricity–heat–cooling conversion model is established for a data center equipped with ADS. This model comprehensively includes server heat generation, heat recovery, and heat-driven refrigeration, thereby enabling closed-loop energy utilization.
2) An optimal planning model for a hybrid cooling system is developed considering the characteristics of on-site electricity–heat–cooling conversion. A discrete-difference-based Benders decomposition algorithm is proposed to efficiently solve the planning problem, thereby facilitating waste-heat utilization.
3) A tailored reformulation method for the on-site electricity–heat–cooling conversion model is designed to ensure computational tractability. Two-dimensional Special Ordered Set of type 2 approximation is employed to reformulate the temperature-dependent ADS performance and related nonlinearities, converting the original problem into a mixed-integer linear problem.

II. On-site Electricity–Heat–Cooling Conversion Framework for DC with ADS

A. WHR-enabled DC Infrastructure

The DC features multi-energy coupling that enables on-site waste-heat recovery. In this framework, the ADS serves as the key conversion device, using recovered waste heat to desorb the refrigerant from the adsorbent and thereby produce chilled water for cooling. A more detailed description of the ADS operating process is provided later in Subsection II-B3. The supporting infrastructure comprises liquid-cooled high-density racks, air-cooled low-density racks, ADS units, water-to-water heat recuperators, and a wet cooling tower.

Liquid-cooled racks and air-cooled racks exhibit different characteristics in terms of heat density, waste-heat temperature, and recovery potential. Liquid-cooled racks house high-density computing equipment such as CPUs and GPUs, which typically accounts for the dominant share of information technology (IT) electricity consumption [26]. The liquid-cooling systems can capture concentrated heat at relatively high outlet temperatures, typically around 55-$70^{\circ}\mathrm{C}$ [27], making this waste heat suitable for driving low-temperature ADS units. Air-cooled racks accommodate lower-density computing loads such as storage servers, which generally contribute a comparatively smaller share [28]. The waste heat from air-cooled racks is more dispersed and usually remains at a lower temperature level, approximately 30–$40^{\circ}\mathrm{C}$ [27]. This lower temperature leads to weaker recovery potential and lower thermodynamic usefulness. Since waste heat from liquid-cooled racks has a higher temperature and better recovery potential than that from air-cooled racks, WHR is applied exclusively to liquid-cooled racks. Accordingly, liquid-cooled racks are modeled as the recoverable heat source, while air-cooled racks are modeled as the cooling load.

The liquid-cooled racks are served by a liquid-cooling loop with intermediate rack-level water-to-water heat exchangers. The recovered waste heat is then directed to drive the ADS. The resulting chilled water is supplied to the computer room air handler that serves the air-cooled racks, thereby enabling closed-loop energy utilization. Meanwhile, the residual waste heat, together with the ADS's rejection heat, is ultimately rejected to the ambient environment through a wet cooling tower. The on-site electricity–heat–cooling conversion pathways of the DC are represented in Fig. 1. In summary, the ADS serves as the key heat-driven cooling component in the proposed framework. It uses recovered waste heat from liquid-cooled racks to produce chilled water for air-cooled racks, thereby reducing the cooling demand supplied by the VCR system. The VCR system supplies the remaining cooling demand. This structure enables on-site utilization of waste heat within the DC and forms the core electricity-heat-cooling conversion path considered in this study.

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

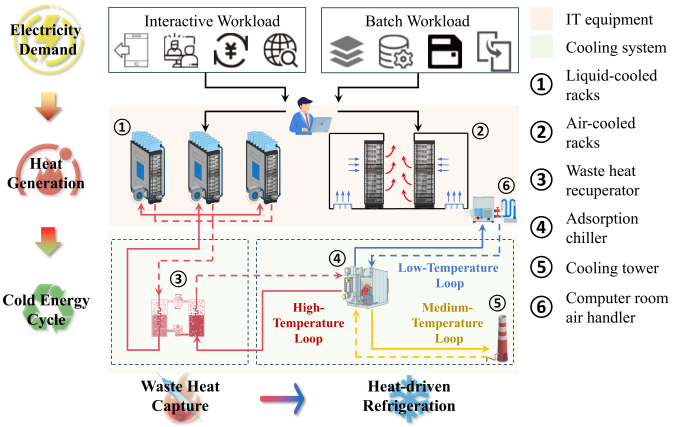


Fig. 1: On-site electricity–heat–cooling conversion pathways within DC.

B. On-site Electricity–Heat–Cooling Conversion Paths

1) Electricity Demand of Computing Workload: As indicated by the row corresponding to electricity demand in Fig. 1, computing workload is divided into interactive workload that requires timely responses and batch workload that can tolerate delays. The DC jointly optimizes workload scheduling and server operating frequency to reduce energy costs under time-varying electricity prices.

2) Heat Generation of Server Racks: As indicated by the row corresponding to heat generation in Fig. 1, the electricity consumption of server racks is ultimately converted into a heat load. To maintain rack temperature within allowable limits, the cooling system removes this heat. Specifically, high-density racks are cooled by the liquid-cooling loop, and low-density racks are cooled by in-row air-side units.

3) Waste-Heat-Driven Cooling Cycle: As indicated by the row corresponding to the cold energy cycle in Fig. 1, the waste heat recovered from liquid-cooled racks drives the ADS to produce chilled water, which is then used to cool low-density air-cooled racks. This closed-loop cycle enables on-site utilization of waste heat. Specifically, this cycle starts with waste-heat recovery from liquid-cooled racks and then proceeds through high-temperature loops, medium-temperature loops, and low-temperature loops to convert waste heat into cooling.

a) Waste Heat Capture: Liquid-cooled racks in the DC produce waste heat at higher temperatures than air-cooled systems. To harness this waste heat, a plate heat exchanger serving as the water-to-water waste-heat recuperator transfers heat from the primary liquid-cooling loop to the generator-side hot-water stream.

b) High-Temperature Loop in Waste-Heat-Driven Cooling: The high-temperature loop delivers the recovered hot-water stream to the generator of the ADS, namely the thermally driven desorption side of the cycle. The supplied heat elevates the temperature of the adsorbent material and promotes refrigerant desorption. This desorption step supplies refrigerant vapor for subsequent stages of the refrigeration cycle, thereby converting recovered waste heat into the driving input for cooling production.

c) Medium-Temperature Loop in Waste-Heat-Driven Cooling: The medium-temperature loop couples the condenser of the ADS units with a wet cooling tower to reject heat to the ambient environment. Cooling water circulates through the condenser to remove the heat of condensation from the refrigerant and then returns to the cooling tower, where fans and pumps facilitate airflow and water circulation to reject it to the ambient environment via evaporation.

d) Low-Temperature Loop in Waste-Heat-Driven Cooling: The low-temperature loop operates at the evaporator of ADS, producing chilled water that is supplied to the computer room air handler serving the air-cooled racks. Upon completion of the adsorption cycle, low-temperature chilled water is generated and distributed to the computer room air handler, offloading a substantial portion of the cooling demand from traditional VCR systems.

III. Problem Formulation for Optimal Planning of DC with On-Site Energy Conversion

The optimization problem is formulated to determine the optimal number of ADS units and the optimal capacity of VCR in a DC.

A. Objective Function

The objective is to minimize the total annualized cost, which includes the annualized investment cost and the annualized operation cost.

$$\min_{\theta^{plan}} C^{inv} + C^{ope} \tag{1}$$

$$C^{inv} = \left(c^{ADS} N^{ADS}_{plan} + c^{VCR} Q^{VCR}_{plan}\right) \times \frac{r^{disc}(1+r^{disc})^{L}}{(1+r^{disc})^{L}-1} \tag{2}$$

$$C^{ope} = \sum_{m\in\mathcal{M}} \min_{\theta^{sch}} \sum_{t\in\mathcal{T}} (\pi^{elec}_{t} + \pi^{carbon}_{t}\gamma_{t}) \times P^{total}_{m,t} \times D \tag{3}$$

$$P^{total}_{m,t} = P^{SRV}_{m,t} + P^{LC}_{m,t} + P^{VCR}_{m,t} + P^{ADS}_{m,t} + P^{fan}_{m,t} \tag{4}$$

where $P^{total}_{m,t}$ denotes the total electricity consumption in scenario m and period t, including the electricity consumption of server racks, the liquid-cooling system, the VCR system, ADS, and the cooling-tower fan.

The operation model is formulated separately for each representative daily scenario. For brevity, the scenario index m is omitted in the following equations.

B. Operation Model of Computing Workloads

For capacity planning and hourly operation scheduling, the IT workload model captures the relationships among workload scheduling, server operating frequency, server allocation, and the electricity consumption of server racks. Batch workloads are deferred within time tolerance to minimize electricity cost.

$$\lambda^{batch}_{t} = \sum_{t_1\in\mathcal{T}} \delta^{batch}_{t_1\to t} \tag{5}$$

$$\begin{cases} \lambda^{batch,0}_{t_1} \geq \delta^{batch}_{t_1\to t} \geq 0, & t_1 \leq t \leq t_1 + \tau^{batch}_{t_1} \\ \delta^{batch}_{t_1\to t} = 0, & \text{else} \end{cases} \tag{6}$$

$$\sum_{t\in\mathcal{T}} \delta^{batch}_{t_1\to t} = \lambda^{batch,0}_{t_1} \tag{7}$$

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

Workload conservation is enforced by requiring that each original batch workload is fully allocated across admissible execution times.

The average response time for interactive workloads is given as follows.

$$\frac{1}{\mu_t - \lambda_t^{inter}/S_t^{inter}} \le \tau_t^{inter} \tag{8}$$

DVFS with discrete frequency levels is modeled as follows.

$$\sum_{l\in\mathcal{L}} z_{t,l} = 1, \quad f_t = \sum_{l\in\mathcal{L}} \bar{f}_l z_{t,l}, \quad \mu_t = \sum_{l\in\mathcal{L}} \bar{\mu}_l z_{t,l} \tag{9}$$

The required numbers of active servers can be expressed as follows.

$$S_t^{inter} = \lambda_t^{inter} \sum_{l\in\mathcal{L}} \frac{z_{t,l}}{\bar{\mu}_l - 1/\tau_t^{inter}} \tag{10}$$

$$S_t^{batch} = \lambda_t^{batch} \sum_{l\in\mathcal{L}} \frac{z_{t,l}}{\bar{\mu}_l} \tag{11}$$

$$S_t = S_t^{inter} + S_t^{batch}, \quad 0 \le S_t \le S^{max} \tag{12}$$

The electricity consumption of server racks is modeled as follows.

$$P_t^{SRV} = \sum_{l\in\mathcal{L}} \left(P^{idle} + \kappa^{DVFS}(\bar{f}_l)^3\right) N_{t,l} \tag{13}$$

To preserve linearity and enable efficient solving, a big-M linearization is adopted to couple $N_{t,l}$ with S_t and $z_{t,l}$.

$$N_{t,l} \le M^{srv} z_{t,l} \tag{14}$$

$$N_{t,l} \ge S_t - M^{srv}(1 - z_{t,l}) \tag{15}$$

$$0 \le N_{t,l} \le S_t \tag{16}$$

where constant M^{srv} is set as S^{max}.

The electricity consumption of server racks P_t^{SRV} translates into an equivalent heat dissipation, which is divided into liquid-cooling and air-cooling portions.

$$Q_t^{liq} = \beta P_t^{SRV}, \quad Q_t^{air} = (1-\beta) P_t^{SRV} \tag{17}$$

The heat load drives the subsequent cooling system models, with Q_t^{liq} providing waste-heat input to the ADS and Q_t^{air} being met jointly by the ADS and VCR.

C. Operation Model of ADS with On-site Waste Heat Utilization

Waste heat from liquid-cooled racks is recovered to drive the ADS, which supplies chilled water to air-cooled racks and partially displaces the VCR system.

1) Waste Heat Capture from Liquid-Cooled Racks: The waste heat is recovered through the liquid-cooling loop and modeled as the available heat source for driving the ADS. The heat transferred to the ADS is further limited by the heat-exchanger effectiveness, the recoverable temperature difference, and the ADS operating constraints. The outlet temperature of the liquid-cooling loop is calculated as follows.

$$T_t^{liq,out} = T_t^{liq,in} + \frac{Q_t^{liq}}{\rho_w c_{p,w} \dot{V}^{liq}} \tag{18}$$

The electricity consumption of liquid-cooling system is modeled as follows.

$$P_t^{LC} = \frac{Q_t^{liq}}{\mathrm{COP}^{LC}} \tag{19}$$

2) High-Temperature Loop: This loop transfers waste heat from liquid-cooled racks to the generators of the ADS units via heat exchangers. The heat input raises the adsorbent temperature, promoting refrigerant desorption and converting waste heat into the thermal driving input for cooling production.

Waste heat is transferred to the generator via a plate heat exchanger.

$$T_t^{gen,in} = T_t^{gen,out} + \varepsilon_{hex}\left(T_t^{liq,out} - T_t^{gen,out}\right) \tag{20}$$

Assuming symmetric operation across the active ADSs, the total generator heat input is evenly distributed among them. The heat input to the generator is bounded by the available waste heat.

$$0 \le N_t^{ADS} q_t^{gen} \le Q_t^{liq} \tag{21}$$

This constraint ensures that the ADS generator heat input does not exceed the available liquid-cooling heat. This and analogous integer–continuous products can be automatically linearized within the YALMIP/Gurobi framework.

The heat transfer process of the generation-side loop is given as follows.

$$T_t^{gen,out} = T_t^{gen,in} - \frac{q_t^{gen}}{\rho_w c_{p,w} \dot{V}^{gen}} \tag{22}$$

3) Medium-Temperature Loop: This loop connects the condenser to a wet cooling tower for heat rejection.

Based on the energy balance principle across the ADS, the required total heat rejection can be determined.

$$Q_t^{ct} = N_t^{ADS}(q_t^{gen} + q_t^{evap}) \tag{23}$$

This equation represents that the heat rejected by the cooling tower is equal to the sum of the generator heat input and the evaporator heat load.

The outlet temperature of condenser is calculated as follows.

$$T_t^{cond,out} = T_t^{cond,in} + \frac{Q_t^{ct}}{\rho_w c_{p,w} \dot{V}^{cond}} \tag{24}$$

The heat rejection capacity of cooling tower is modeled as follows.

$$Q_t^{ct} = \rho_{air} \dot{V}_t^{ct} \varepsilon_{ct} \Delta h_t \tag{25}$$

$$\Delta h_t = h_{sat}(T_t^{cond,out}) - h_{amb} \tag{26}$$

where $h_{sat}(T_t^{cond,out})$ represents the enthalpy of saturated moist air at $T_t^{cond,out}$. The details about air enthalpy and the saturation humidity ratio approximation are provided in Appendix A.

Based on the air flow rate, cooling tower fan power is modeled as follows.

$$P_t^{fan} = \frac{(K_a^{ct} \dot{V}_t^{ct})^3}{\eta_{fan}} \tag{27}$$

4) Low-Temperature Loop: This loop at evaporator produces chilled water to the computer room air handler serving

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

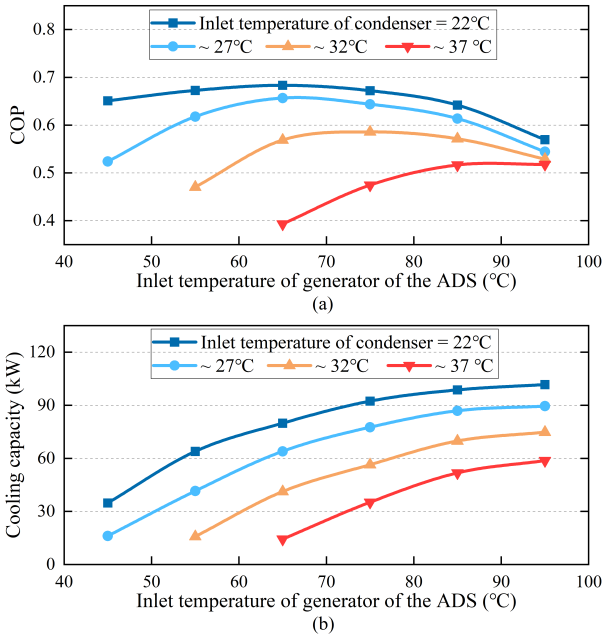


Fig. 2: Performance curves of an ADS under different operating temperatures: (a) COP; (b) cooling capacity.

the air-cooled racks, thereby offsetting a portion of the VCR cooling load and lowering electricity consumption.

For the proposed capacity planning problem, the ADS is represented by a quasi-steady operation model. The ADS cycle dynamics are assumed to be handled by the local equipment controller. The ADS performance is characterized using a calibrated Gordon-Ng model for thermal-driven chillers, which is expressed in a general form as follows.

$$\mathrm{COP}_t^{ADS} = f\left(T_t^{gen,in}, T_t^{cond,in}, T_t^{evap,in}, T_t^{evap,out}, q_t^{cap}\right) \tag{28}$$

The above semi-empirical relation characterizes the coupling between the COP and the cooling capacity under specified operating temperatures. Fig. 2 is derived from the manufacturer performance tables provided in [29]. The COP and cooling capacity curves are obtained by fixing the evaporator-side temperature.

Then, the cooling output delivered by the evaporator is modeled as follows.

$$q_t^{evap} = q_t^{gen} \cdot \mathrm{COP}_t^{ADS} \tag{29}$$

$$0 \le q_t^{evap} \le q_t^{cap} \tag{30}$$

The evaporator return water temperature is calculated as follows.

$$T_t^{evap,in} = T_t^{evap,out} + \frac{q_t^{evap}}{\rho_w c_{p,w} \dot{V}^{evap}} \tag{31}$$

The chilled water from the ADS is supplied to the computer room air handler, thereby reducing the VCR load.

$$q_t^{evap} N_t^{ADS} + Q_t^{VCR} = Q_t^{air} \tag{32}$$

The electricity consumption of VCR is calculated as follows.

$$P_t^{VCR} = \frac{Q_t^{VCR}}{\mathrm{COP}^{VCR}} \tag{33}$$

The electricity consumption of ADS auxiliaries is modeled as a fixed per-unit power.

$$P_t^{ADS} = P^{ADS,aux} \times N_t^{ADS} \tag{34}$$

The ADS auxiliary power is typically weakly dependent on the cooling output and is therefore approximated as a constant for planning-oriented optimization.

5) Operating Bounds: The operating bounds of the ADS and VCR are defined according to their respective planned capacities. For the ADS, the number of active units is constrained by the planned number of installed ADS units.

$$N_t^{ADS} \in \mathbb{Z}_{\ge 0}, \quad \delta_t^{ADS} \in \{0,1\} \tag{35}$$

$$\delta_t^{ADS} \le N_t^{ADS} \le N_{plan}^{ADS} \delta_t^{ADS} \tag{36}$$

$$0 \le q_t^{gen} \le M^{gen} \delta_t^{ADS} \tag{37}$$

$$0 \le q_t^{evap} \le M^{evap} \delta_t^{ADS} \tag{38}$$

$$0 \le Q_t^{ct} \le M^{ct} \delta_t^{ADS} \tag{39}$$

where the big-M constants M^{gen}, M^{evap} and M^{ct} are set to the corresponding physical upper bounds of each variable.

When $\delta_t^{ADS} = 0$, no ADS unit is active, the ADS generator heat input, evaporator cooling output, and ADS-related heat rejection are all forced to zero. When $\delta_t^{ADS} = 1$, the number of active ADS units is bounded by the installed ADS number.

For the VCR, the cooling output is bounded by the planned VCR capacity.

$$0 \le Q_t^{VCR} \le Q_{plan}^{VCR} \tag{40}$$

Additionally, the feasible operating region of the ADS is defined by the following temperature and flow constraints.

$$T_{\min}^{liq,in} \le T_t^{liq,in} \le T_{\max}^{liq,in} \tag{41}$$

$$T_{\min}^{cond,in} \le T_t^{cond,in} \le T_{\max}^{cond,in} \tag{42}$$

$$T_{\min}^{evap,out} \le T_t^{evap,out} \le T_{\max}^{evap,out} \tag{43}$$

$$T_{\min}^{gen,in} \le T_t^{gen,in} \le T_{\max}^{gen,in} \tag{44}$$

$$T_{\min}^{gen,out} \le T_t^{gen,out} \le T_{\max}^{gen,out} \tag{45}$$

$$\dot{V}_{\min}^{ct} \le \dot{V}_t^{ct} \le \dot{V}_{\max}^{ct} \tag{46}$$

where $(\cdot)_{\min}$ and $(\cdot)_{\max}$ denote the lower and upper bounds of the corresponding variables defined above.

IV. Model Reformulation and Decomposition-based Solution Approach

To solve the proposed optimization problem efficiently, the original problem is first reformulated into a mixed-integer linear problem (MILP). This reformulation is conducted in three steps. First, the nonlinear and temperature-dependent characteristics of the ADS in terms of COP and cooling capacity are incorporated into the optimization problem via a two-dimensional special ordered set of type 2 (SOS2-2D) approximation. Second, the bilinear term involving generator heat input, COP, and evaporator cooling output is reformulated using SOS2-2D approximation. Third, the nonlinear fan power curve is piecewise-linearized using special ordered set of type

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

2 (SOS2), and the cooling tower heat rejection capacity is approximated via an SOS2-2D approximation. After the MILP is obtained, a discrete-difference-based Benders decomposition is applied to separate long-term capacity planning decisions from short-term operation scheduling, then iteratively solve the planning problem.

A. Reformulation of Temperature-dependent ADS Performance

The COP and the cooling capacity are nonlinear functions of the operating temperatures. In this study, the temperature-dependent ADS performance is embedded into the planning and scheduling model through an optimization-compatible reformulation. Using the operating points and associated parameters provided in the manufacturer's performance tables [29], the COP and the cooling capacity under different temperature conditions are computed based on the semi-empirical relations introduced above. Then an SOS2-2D approximation is employed to represent these performance characteristics in a tractable form. Since the outlet temperature and inlet temperature of evaporator are fixed, the key inputs reduce to the generator inlet temperature $T_t^{gen,in}$ and condenser inlet temperature $T_t^{cond,in}$. An SOS2-2D approximation is employed as follows.

For each period t, SOS2 weights are introduced in each dimension.

$$\lambda_{i,t}^{gen} \geq 0, \quad \sum_{i\in\mathcal{I}} \lambda_{i,t}^{gen} = 1, \quad (\lambda_{i,t}^{gen})_{i\in\mathcal{I}} \in \text{SOS2} \tag{47}$$

$$\lambda_{j,t}^{cond} \geq 0, \quad \sum_{j\in\mathcal{J}} \lambda_{j,t}^{cond} = 1, \quad (\lambda_{j,t}^{cond})_{j\in\mathcal{J}} \in \text{SOS2} \tag{48}$$

where $\mathcal{I}$ and $\mathcal{J}$ are index sets of the breakpoints for $T^{gen,in}$ and $T^{cond,in}$; $\lambda_{i,t}^{gen}$ and $\lambda_{j,t}^{cond}$ are the SOS2 weights.

Two-dimensional interpolation is implemented as follows.

$$\lambda_{i,j,t}^{ADS} \geq 0, \quad \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \lambda_{i,j,t}^{ADS} = 1 \tag{49}$$

$$\sum_{j\in\mathcal{J}} \lambda_{i,j,t}^{ADS} = \lambda_{i,t}^{gen}, \quad \sum_{i\in\mathcal{I}} \lambda_{i,j,t}^{ADS} = \lambda_{j,t}^{cond} \tag{50}$$

$$(\lambda_{i,j,t}^{ADS})_{i\in\mathcal{I}} \in \text{SOS2}, \quad (\lambda_{i,j,t}^{ADS})_{j\in\mathcal{J}} \in \text{SOS2} \tag{51}$$

where $\lambda_{i,j,t}^{ADS}$ is the interpolation weight at grid point (i,j). Constraints (50) link the two-dimensional weights to the one-dimensional SOS2 weights, while (51) enforces adjacency in both dimensions. The detailed row-wise and column-wise implementation of (51) is provided in Appendix B.

Using $\lambda_{i,j,t}^{ADS}$, the temperatures and performance quantities are represented as convex combinations of the grid values.

$$T_t^{gen,in} = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \bar{T}_i^{gen} \lambda_{i,j,t}^{ADS} \tag{52}$$

$$T_t^{cond,in} = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \bar{T}_j^{cond} \lambda_{i,j,t}^{ADS} \tag{53}$$

$$\text{COP}_t^{ADS} = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \overline{\text{COP}}_{i,j}^{ADS} \lambda_{i,j,t}^{ADS} \tag{54}$$

$$q_t^{cap} = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \overline{q}_{i,j}^{cap} \lambda_{i,j,t}^{ADS} \tag{55}$$

where $\bar{T}_i^{gen}$ and $\bar{T}_j^{cond}$ are the breakpoints of the corresponding temperature; $\overline{\text{COP}}_{i,j}^{ADS}$ and $\overline{q}_{i,j}^{cap}$ are precomputed COP and cooling capacity at grid point (i,j). This yields a convex and tractable approximation of the nonlinear dependencies.

B. Reformulation of Heat-to-Cooling Coupling

The relationship in (29) contains the product of the heat input to generator and the COP of ADS. Similar to the temperature-dependent ADS performance, this relationship is also approximated using SOS2-2D interpolation.

For each period t, SOS2 weights are introduced in the heat input and COP dimensions.

$$\lambda_{a,t}^{q} \geq 0, \quad \sum_{a\in\mathcal{A}} \lambda_{a,t}^{q} = 1, \quad (\lambda_{a,t}^{q})_{a\in\mathcal{A}} \in \text{SOS2} \tag{56}$$

$$\lambda_{b,t}^{cop} \geq 0, \quad \sum_{b\in\mathcal{B}} \lambda_{b,t}^{cop} = 1, \quad (\lambda_{b,t}^{cop})_{b\in\mathcal{B}} \in \text{SOS2} \tag{57}$$

where $\lambda_{a,t}^{q}$ and $\lambda_{b,t}^{cop}$ are the SOS2 weights in the generator heat input and COP dimensions, respectively.

The two-dimensional interpolation weights are constrained as follows.

$$\lambda_{a,b,t}^{hc} \geq 0, \quad \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} \lambda_{a,b,t}^{hc} = 1 \tag{58}$$

$$\sum_{b\in\mathcal{B}} \lambda_{a,b,t}^{hc} = \lambda_{a,t}^{q}, \quad \sum_{a\in\mathcal{A}} \lambda_{a,b,t}^{hc} = \lambda_{b,t}^{cop} \tag{59}$$

$$(\lambda_{a,b,t}^{hc})_{a\in\mathcal{A}} \in \text{SOS2}, \quad (\lambda_{a,b,t}^{hc})_{b\in\mathcal{B}} \in \text{SOS2} \tag{60}$$

where $\lambda_{a,b,t}^{hc}$ is the corresponding SOS2-2D interpolation weight.

Using $\lambda_{a,b,t}^{hc}$, the evaporator cooling output is represented as follows.

$$q_t^{evap} = \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} \overline{q}_a^{gen} \overline{\text{COP}}_b^{ADS} \lambda_{a,b,t}^{hc} \tag{61}$$

where $\overline{q}_a^{gen}$ and $\overline{\text{COP}}_b^{ADS}$ are the breakpoints of the generator heat input and COP, respectively. The weights $\lambda_{a,b,t}^{hc}$ also reconstruct q_t^{gen} and COP_t^{ADS} from their respective breakpoints.

C. Reformulation of Cooling Tower Model

The cooling tower fan power in (27) exhibits a nonconvex cubic form and is linearized via an SOS2-based piecewise approximation. Moreover, the cooling tower heat rejection model in (25) contains the product between air flow rate and temperature-dependent enthalpy difference, which is handled by an SOS2-2D approximation.

A piecewise linear approximation with SOS2 is utilized to linearize the nonconvex cooling tower fan power constraints.

$$\lambda_{v,t}^{ct} \geq 0, \quad \sum_{v\in\mathcal{V}} \lambda_{v,t}^{ct} = 1, \quad (\lambda_{v,t}^{ct})_{v\in\mathcal{V}} \in \text{SOS2} \tag{62}$$

$$\dot{V}_t^{ct} = \sum_{v\in\mathcal{V}} V_v^{break} \lambda_{v,t}^{ct}, \quad P_t^{fan} = \sum_{v\in\mathcal{V}} P_v^{fan} \lambda_{v,t}^{ct} \tag{63}$$

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

where $\lambda_{v,t}^{ct}$ denotes the SOS2 weight for the air volumetric flow rate at breakpoint v and time period t; V_v^{break} denotes the breakpoint for the air volumetric flow rate at index v; P_v^{fan} denotes the breakpoint for fan power at index v.

For the cooling tower heat-rejection constraint in (25), an SOS2-2D approximation is adopted to represent the coupling between the air volumetric flow rate and the temperature-dependent enthalpy difference. Following the same interpolation logic used in the preceding SOS2-2D reformulations, the cooling-tower breakpoint sets $\mathcal{V}$ and $\mathcal{R}$ are used for the air flow rate and condenser outlet temperature, respectively. Each interpolation weight $\lambda_{v,r,t}^{hr}$ corresponds to one grid point formed by the air-flow breakpoint indexed by v and the condenser-temperature breakpoint indexed by r. The enthalpy difference is first evaluated at each condenser-temperature breakpoint, and the heat-rejection capacity is then approximated by the weighted combination of these two-dimensional grid values. This treatment allows the nonlinear cooling tower heat-rejection relation to be embedded into the MILP in a tractable form.

D. Discrete-Difference-based Benders Decomposition Algorithm

After the above transformations, the operation problem can be reformulated as a MILP. For a given pair of planning decisions $(N_{plan}^{ADS}, Q_{plan}^{VCR})$, the optimal annual operation cost is obtained by solving the following problem.

$$\begin{aligned}&\phi\big(N_{plan}^{ADS}, Q_{plan}^{VCR}\big)\\&= \min_{\theta^{sch}} \sum_{m\in\mathcal{M}}\sum_{t\in\mathcal{T}} \left(\pi_t^{elec} + \pi_t^{carbon}\gamma_t\right) P_{m,t}^{total} D \qquad (64)\\&\text{s.t.} \quad (5)-(7),\ (9)-(24),\ (30)-(63).\end{aligned}$$

where $\phi(\cdot)$ is the value function of optimal annual operation cost.

The original planning problem couples capacity decisions with operation scheduling. The planning decisions include the number of ADS units and the VCR capacity. The operation scheduling coordinates workload scheduling, waste-heat recovery, ADS cooling output, VCR cooling output, and cooling-tower operation. Directly solving the integrated model is computationally demanding. Therefore, a discrete-difference-based Benders decomposition algorithm is applied. The master problem (MP) determines the installed ADS units and VCR capacity, while the subproblem (SP) evaluates the operation cost for a given planning solution by solving (64). The resulting operation-cost information is used to construct Benders cuts, which are added to the MP iteratively. The MP includes a cooling-capacity adequacy requirement. The peak air-cooling demand is obtained from the workload-only scheduling results, then this peak demand is scaled by the redundancy factor for capacity planning. Details of the above information are provided in Appendix C.

V. Case Study

A. Experimental Setups

In this case study, the parameters related to electricity price and carbon emissions are obtained from [30], [31]. The related operating parameters of ADS and VCR are obtained from [19]. The DVFS-related parameters are obtained from [24]. The parameters related to workload scheduling in DC are from [1]. Houston is selected as the case study location due to its representative large-scale DC deployment conditions. The corresponding hourly meteorological data, including ambient temperature, relative humidity, and other relevant parameters are retrieved from [32]. For each month, a representative day is extracted from hourly meteorological data using k-means clustering, resulting in 12 representative scenarios for planning. The relevant parameters used in the case study are organized and made available at [33]. To examine the performance of the proposed method, the following benchmarks are established for comparison:

Scenario 1 (S1): Conventional cooling configuration relying solely on VCR, with no ADS deployed.

Scenario 2 (S2): Cooling operation under a fixed computing workload profile, where workload scheduling is disabled.

Scenario 3 (Proposed, S3): Integrated optimization with VCR and ADS, where the number of ADS units is optimally determined and workload scheduling is enabled.

B. Fitting Performance of the ADS Operation Model

To obtain an accurate surrogate that can be embedded directly into the optimization model, SOS2-2D formulation is applied to approximate the COP and cooling capacity. Fig. 3 presents the fitting performance of the proposed SOS2-2D model using 20 segments along each dimension. As shown, both COP and cooling capacity are fitted with high accuracy. Fig. 4 compares the fitting error, the embedding-induced error, and the solution time across different methods. Among the approaches that can be directly embedded into optimization models, SOS2-2D achieves superior accuracy without a significant increase in computational time, making it more suitable for practical applications. The multilayer perceptron (MLP) can achieve high fitting accuracy. However, although the discretization technique enables a theoretically exact reformulation for embedding MLP into an optimization problem, numerical issues may still lead to deviations that accumulate across layers. In contrast, linear regression is easy to embed but cannot capture nonlinear relationships.

Table II summarizes the approximation accuracy of the proposed methods for fan power, heat rejection, and the heat-to-cooling relationship under different parameter settings. For the heat-to-cooling relationship, the reported errors denote the maximum deviation. As the number of SOS2 segments increases, the approximation accuracy improves accordingly. This study adopts 30 segments for fan power, heat rejection and heat-to-cooling relationship.

C. Economic Analysis of Hybrid Cooling System

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

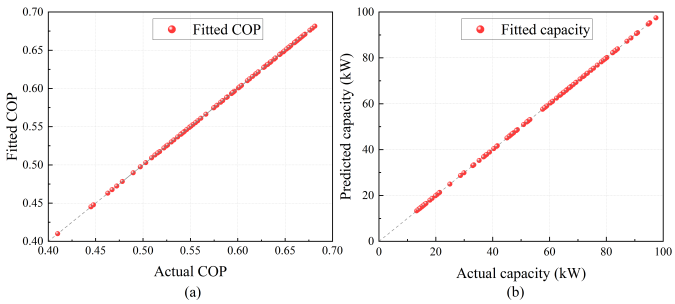


Fig. 3: Fitted curves of the SOS2-2D model for ADS: (a) COP; (b) cooling capacity.

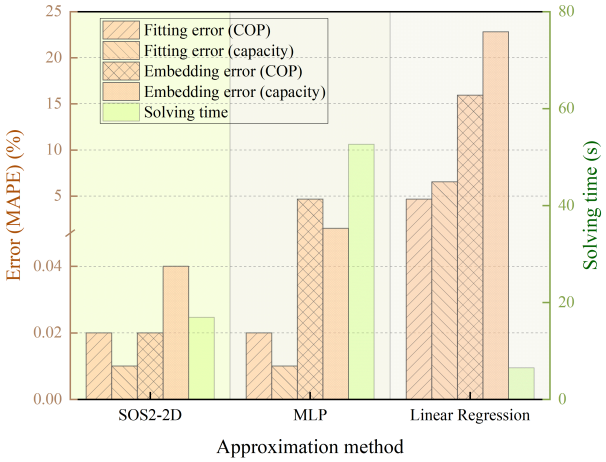


Fig. 4: Comparison of fitting error, embedding-induced error, and solution time across different methods.

Under the considered techno-economic parameters, the minimum total annualized cost is achieved with four ADS units. Fig. 5 compares the cooling-related annualized costs under different scenarios, including the annualized investment cost and the annual operation cost. It can be observed that adopting waste-heat-driven ADS increases the annualized investment cost, while it significantly reduces the annual operation cost by 46.16%. In particular, the related cooling electricity cost and the carbon emission cost decrease by 46.41% and 44.89%, respectively. As a result, the cooling-related total annualized cost is reduced by 19.01%, demonstrating the economic benefit of ADS from the perspective of the cooling subsystem. It should be noted that the electricity cost associated with server racks and liquid cooling system is excluded from Fig. 5 to highlight the economic impact of ADS on DC planning, since ADS planning does not alter the electricity consumption of

TABLE II: Relative errors of approximation components under different parameter settings.

Approximation component	Parameter setting	Relative error (%)
Cooling tower fan power model	Segments: 10 / 20 / 30	13.42 / 2.85 / 0.95
Cooling tower heat rejection model	Segments: 10 / 20 / 30	14.50 / 3.05 / 0.99
Heat-to-cooling relationship	Segments: 10 / 20 / 30	0.39 / 0.07 / 0.03

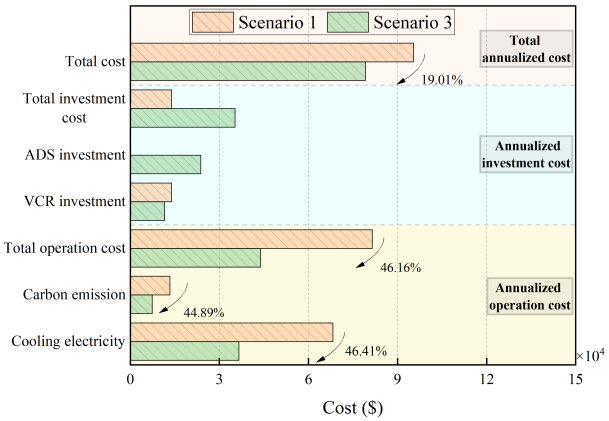


Fig. 5: Comparison of cooling-related costs under different scenarios.

these components.

To further interpret the practical heat-recovery level, a system-level heat-utilization indicator is calculated across all representative days. This indicator is defined as the ratio of the total ADS generator heat input to the total liquid-cooled server heat. The resulting value is 38.22%, which means that 38.22% of the liquid-cooled server heat enters the ADS generator side as thermal driving input. The remaining heat stays outside the ADS generator operation. This value indicates that the model does not assume full utilization of waste heat. Heat-exchanger effectiveness, temperature degradation, and ADS operating conditions jointly limit the amount of heat converted into useful cooling.

Fig. 6 further illustrates the planning tradeoff under different numbers of ADS units. Fig. 6(a) shows the capacity shares of VCR and ADS in the hybrid cooling system. As the number of ADS units increases, the planned VCR capacity gradually decreases, while the ADS share increases accordingly. This reflects the substitution effect of waste-heat-driven cooling on conventional VCR-based cooling. Fig. 6(b) presents the total annualized cost and the capital payback period under different numbers of ADS units. The cost profile provides a direct numerical check of the planning solution obtained by the proposed decomposition method. The minimum total annualized cost is achieved when four ADS units are installed, which is consistent with the solution identified by the proposed method. This result confirms that the selected configuration has the lowest cost among the evaluated ADS configurations. Meanwhile, the capital payback period increases with the number of ADS units, indicating that excessive ADS installation requires a longer time to recover the additional capital investment.

D. Performance Analysis of Integrated Computing and Cooling Scheduling

Fig. 7(a) illustrates the workload profiles under scenarios with and without temporal scheduling of batch workloads. Under temporal scheduling, the volume of computing workloads decreases during periods of high electricity prices and increases during periods of low electricity prices. Fig. 7(b) depicts the corresponding server operating frequency subject to the latency requirements for interactive workload. It can

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

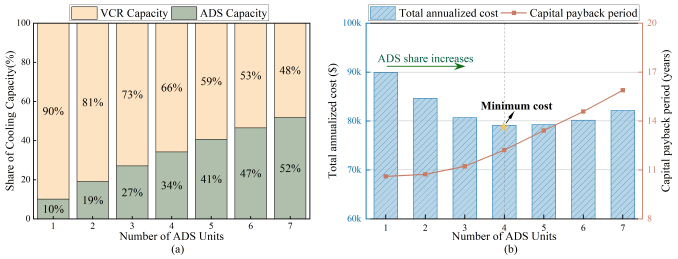


Fig. 6: Planning results under different numbers of ADS units: (a) capacity shares of VCR and ADS; (b) total annualized cost and capital payback period.

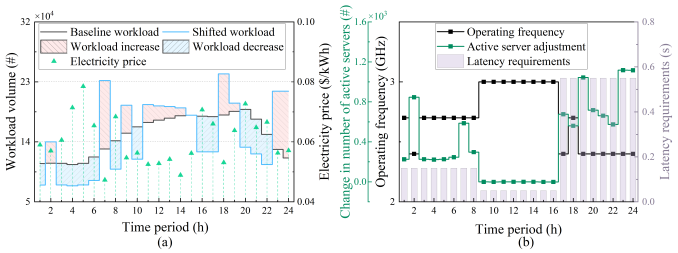


Fig. 7: Workload profiles with and without temporal scheduling: (a) batch workloads; (b) interactive workloads.

be observed that, during periods of more relaxed latency requirements, DVFS reduces the server operating frequency and increases the number of active servers to minimize electricity consumption. In contrast, during periods of higher latency requirements, the server operating frequency is maintained at elevated levels to ensure timely task responses.

The various electricity consumption and operating costs are presented in Table III. All results in this table are obtained for one representative day scenario. It can be observed that the scheduling approach that accounts for temporal scheduling of workloads reduces electricity consumption by 16.45%. This is because, when latency requirements are less stringent, DVFS can reduce the electricity consumption of server racks by operating servers at lower frequencies. Since electricity consumption of workloads is ultimately dissipated as heat, the same rationale applies to the cooling system.

E. Sensitivity Analysis of Ambient Temperature and Humidity

The operating performance of ADS is sensitive to ambient temperature and humidity. Fig. 8 illustrates the monthly cooling-related electricity consumption and the corresponding electricity savings. The cooling-related electricity consumption first increases and then decreases over the year. The electricity savings are generally more pronounced in spring and autumn, while the savings become relatively smaller in summer.

To decouple the impacts of temperature and humidity, Fig. 9 presents the electricity consumption for cooling and the electricity savings rate under varying ambient conditions. For clarity, three representative temperature levels and humidity levels are defined as Low-T (10°C), Med-T (23°C), High-T (35°C), Low-W (5 g/kg), Med-W (15 g/kg) and High-W (25 g/kg). In Fig. 9(a), the electricity consumption increases monotonically with humidity. However, the electricity savings rate is non-monotonic with respect to humidity for all temperature levels. It remains relatively stable from Low-T to Med-T but drops noticeably at High-T. This behavior indicates that the relative benefit of ADS diminishes in hot and humid environments, because a higher wet-bulb temperature raises the condensing temperature and weakens heat rejection, thereby lowering the effective COP and cooling capacity. Meanwhile, the associated auxiliary electricity consumption increases, such as the power demand of cooling tower fans, which further offsets the electricity savings. Fig. 9(b) shows a pattern similar to Fig. 9(a). The total electricity consumption increases with temperature across all humidity levels. For Low-W, the electricity savings rate generally increases with temperature. For Med-W, the electricity savings rate remains generally stable. In contrast, under High-W, the electricity-saving rate decreases sharply as temperature increases.

TABLE III: Electricity consumption and electricity cost comparison under different operation scenarios.

Scenario	S2	S3
Electricity consumption of server racks (MWh)	41.21	34.43 (-16.45%)
Electricity cost of server racks (kUSD)	2.56	2.01 (-21.83%)
Electricity consumption of cooling (MWh)	2.89	2.31 (-20.12%)
Electricity cost of cooling (kUSD)	0.17	0.13 (-23.69%)
Total electricity consumption (MWh)	44.10	36.74 (-16.69%)
Total electricity cost (kUSD)	2.74	2.14 (-21.95%)

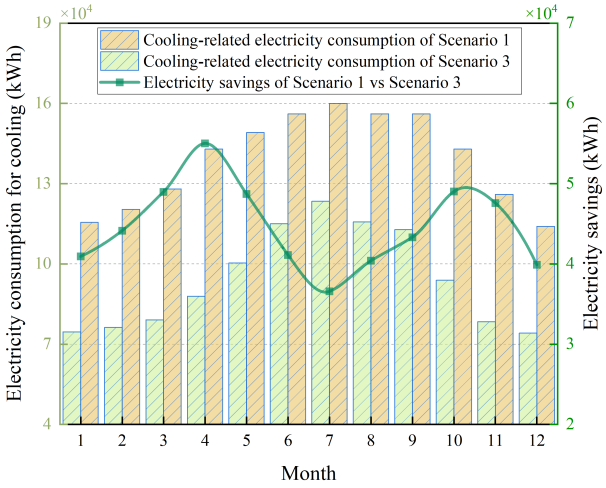


Fig. 8: Monthly electricity consumption for cooling and energy savings achieved by hybrid cooling systems.

VI. CONCLUSION

This paper proposes an optimal planning framework for the hybrid ADS-VCR cooling system in DC to utilize low-grade waste heat for on-site cooling. The proposed SOS2-2D approximation yields a tractable MILP with validated modeling accuracy. Compared with the VCR-only baseline, integrating waste-heat-driven ADS increases capital investment but reduces operation cost. Under the optimal sizing decisions, the total annualized cost is minimized, while the payback period increases with the installed number of ADS units. Moreover, joint scheduling of computing workloads and the cooling system further improves the economic performance of the

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Smart Grid. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSG.2026.3750936

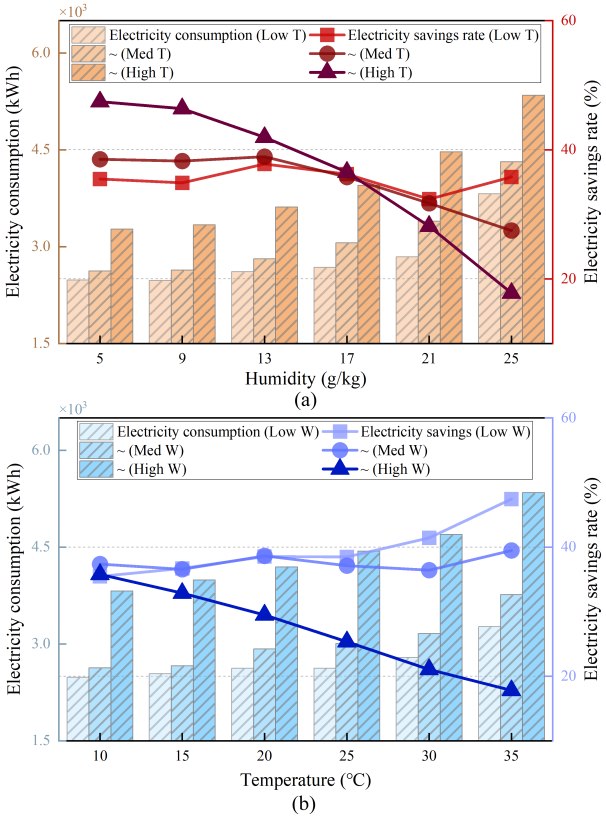


Fig. 9: Electricity consumption for cooling and energy-saving rate under various ambient conditions: (a) Varying humidity with temperature fixed at three levels; (b) Varying temperature with humidity fixed at three levels.

hybrid system. Sensitivity analysis shows that the electricity savings rate is highly dependent on ambient conditions.

REFERENCES

[1] M. Song, X. Yan, X. Wang, C. Gao, W. Zhang, S. Xia, Y. Zhao, and D. Abbess, "Optimal scheduling strategy of networked data centers in multiple transactive energy markets," *CSEE Journal of Power and Energy Systems*, 2024.

[2] Y. Sun, Z. Ding, P. Dehghanian, and F. Teng, "Learning-enabled adaptive power capping scheme for cloud data centers," *IEEE Transactions on Smart Grid*, 2025.

[3] Y. Sun, Z. Ding, Y. Yan, Z. Wang, P. Dehghanian, and W.-J. Lee, "Privacy-preserving energy sharing among cloud service providers via collaborative job scheduling," *IEEE Transactions on Smart Grid*, 2024.

[4] X. Yin, C. Ye, Y. Ding, and Y. Song, "Exploiting internet data centers as energy prosumers in integrated electricity-heat system," *IEEE Transactions on Smart Grid*, vol. 14, no. 1, pp. 167–182, 2022.

[5] L. Socci, A. Rocchetti, A. Verzino, *et al.*, "Enhancing third-generation district heating networks with data centre waste heat recovery: analysis of a case study in italy," *Energy*, vol. 313, p. 134013, 2024.

[6] X. Yin, C. Ye, Y. Ding, *et al.*, "Combined heat and power dispatch against cold waves utilizing responsive internet data centers," *IEEE Transactions on Sustainable Energy*, vol. 15, no. 2, pp. 819–834, 2023.

[7] K. Elsaid, E. T. Sayed, B. A. A. Yousef, *et al.*, "Recent progress on the utilization of waste heat for desalination: A review," *Energy conversion and management*, vol. 221, p. 113105, 2020.

[8] X. Zhou, Z. Xin, W. Tang, *et al.*, "Comparative study for waste heat recovery in immersion cooling data centers with district heating and organic rankine cycle (orc)," *Applied Thermal Engineering*, vol. 242, p. 122479, 2024.

[9] X. Yuan, J. Liu, S. Sun, X. Lin, X. Fan, W. Zhao, and R. Kosonen, "Data center waste heat for district heating networks: A review," *Renewable and Sustainable Energy Reviews*, vol. 219, p. 115863, 2025.

[10] Z. Cao, X. Zhou, H. Hu, Z. Wang, and Y. Wen, "Toward a systematic survey for carbon neutral data centers," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 895–936, 2022.

[11] P. Wang, S. Kowalski, Z. Gao, J. Sun, C.-M. Yang, D. Grant, P. Boudreaux, S. Huff, and K. Nawaz, "District heating utilizing waste heat of a data center: High-temperature heat pumps," *Energy and Buildings*, vol. 315, p. 114327, 2024.

[12] Y. Hao, H. Zhou, T. Tian, W. Zhang, X. Zhou, Q. Shen, T. Wu, and J. Li, "Data centers waste heat recovery technologies: Review and evaluation," *Applied Energy*, vol. 384, p. 125489, 2025.

[13] C. Ye, X. Yin, and Y. Ding, "Hierarchical control of internet data centers for providing fast electricity-heat regulation services," *IEEE Transactions on Smart Grid*, 2025.

[14] M. G. Gado, T. F. Megahed, S. Ookawara, S. Nada, and I. I. El-Sharkawy, "Potential application of cascade adsorption-vapor compression refrigeration system powered by photovoltaic/thermal collectors," *Applied Thermal Engineering*, vol. 207, p. 118075, 2022.

[15] Z. Cui, S. Du, and R. Wang, "Switchable multistage absorption chiller for ultra-low-grade waste heat driven cooling in data centers," *Applied Energy*, vol. 398, p. 126407, 2025.

[16] D. S. Ayou, M. F. V. Wardhana, and A. Coronas, "Performance analysis of a reversible water/libr absorption heat pump connected to district heating network in warm and cold climates," *Energy*, vol. 268, p. 126679, 2023.

[17] S. Du, Z. Cui, R. Wang, H. Wang, and Q. Pan, "Development and experimental study of a compact silica gel-water adsorption chiller for waste heat driven cooling in data centers," *Energy Conversion and Management*, vol. 300, p. 117985, 2024.

[18] Q. Pan, J. Peng, and R. Wang, "Application analysis of adsorption refrigeration system for solar and data center waste heat utilization," *Energy Conversion and Management*, vol. 228, p. 113564, 2021.

[19] R. Gupta and I. K. Puri, "Waste heat recovery in a data center with an adsorption chiller: Technical and economic analysis," *Energy Conversion and Management*, vol. 245, p. 114576, 2021.

[20] Z. Cui, S. Du, T. Zhao, *et al.*, "High-power-density adsorption chiller driven by data center waste heat using encapsulated composite as adsorbent," *Energy*, vol. 311, p. 133391, 2024.

[21] A. Velte-Schäfer, C. Teicht, M. Stahlhut, *et al.*, "Utilizing waste heat from data centers with adsorptive heat transformation–heat exchanger design and choice of adsorbent," *Energy Conversion and Management*, vol. 310, p. 118500, 2024.

[22] B. Zeng, Y. Zhou, X. Xu, *et al.*, "Bi-level planning approach for incorporating the demand-side flexibility of cloud data centers under electricity-carbon markets," *Applied Energy*, vol. 357, p. 122406, 2024.

[23] C. Guo, F. Luo, Z. Cai, *et al.*, "Integrated planning of internet data centers and battery energy storage systems in smart grids," *Applied Energy*, vol. 281, p. 116093, 2021.

[24] W. Li, T. Qian, Y. Zhang, *et al.*, "Distributionally robust chance-constrained planning for regional integrated electricity–heat systems with data centers considering wind power uncertainty," *Applied energy*, vol. 336, p. 120787, 2023.

[25] S. Zhang, J. Lyu, W. Jin, *et al.*, "Coordinated planning of multiple energy hubs considering the spatiotemporal load regulation of data centers," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 4193–4207, 2023.

[26] R. Kong, H. Zhang, M. Tang, H. Zou, C. Tian, and T. Ding, "Enhancing data center cooling efficiency and ability: a comprehensive review of direct liquid cooling technologies," *Energy*, vol. 308, p. 132846, 2024.

[27] M. Pärssinen, M. Wahlroos, J. Manner, and S. Syri, "Waste heat from data centers: An investment analysis," *Sustainable Cities and Society*, vol. 44, pp. 428–444, 2019.

[28] "2024 united states data center energy usage report," 2024. Report prepared with support from U.S. DOE.

[29] InvenSor GmbH, "Ltc 90 e plus product datasheet." [Online]. Available: https://www.aaasaveenergy.com/products/001/pdf/InvenSor_LTC90_e_plus_datasheet.pdf, 2019.

[30] Y. Bian, L. Xie, J. Ye, and L. Ma, "A new shared energy storage business model for data center clusters considering energy storage degradation," *Renewable Energy*, vol. 225, p. 120283, 2024.

[31] U.S. Environmental Protection Agency, "egrid summary tables 2023." [Online]. Available: https://www.epa.gov/system/files/documents/2025-06/summary_tables_rev2.pdf, 2025.

[32] NASA Prediction Of Worldwide Energy Resources (POWER) Project, "Data access viewer (dav)," n.d. NASA Langley Research Center; POWER project web tool.

[33] K. Wang, "Dataset for Optimal Planning of Data Center." IEEE DataPort, 2026. doi: 10.21227/vsk7-rt36. [Online]. Available: https://doi.org/10.21227/vsk7-rt36.

Authorized licensed use limited to: Zhejiang University. Downloaded on November 08,2026 at 03:31:26 UTC from IEEE Xplore. Restrictions apply.
© 2026 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.