

电算协同:数据中心电源配置、负荷建模与预测、电力调度*

王志新^{1,2} 邹怡寒¹ 王 简³ 陈 辉¹ 钱 进¹ 龚春阳¹

(1. 上海电力大学人工智能学部 上海 201306; 2. 上海交通大学电气工程学院 上海 200240;

3. 上海交通大学智慧能源创新学院 上海 200240)

摘 要:随着人工智能(artificial intelligence, AI)普及应用,数据中心(data center, DC)规模和算力负荷快速增长,其能耗管理成为当前能源电力领域关注的核心问题。DC 负荷与运行特性具有特殊性,能耗主要源自信息技术(information technology, IT)设备与冷却系统,负荷呈现高度的非线性与多源耦合特征,为 DC 能耗精准预测与电力调度带来挑战。系统阐述了 DC 电源配置、负荷建模与能耗预测方法,重点对比分析了统计模型、机器学习模型及深度学习模型在 IT 负荷与冷却能耗预测中的适用条件、特征构建方式与建模效果,总结了多源监测数据驱动下的关键特征选择与模型适配规律。同时,综述了基于负荷与能耗预测结果的 DC 参与电力系统调度的研究路径,从时间与空间两个维度分析了预测精度对需求响应、灵活调节能力挖掘及电-算协同运行的影响机制。最后,针对实际应用中面临的多源数据融合不足、模型可解释性有限及实时预测能力受限等问题,归纳了当前研究的主要挑战与发展趋势,并提出支撑 DC 数智化能耗管理与电算协同调度的关键研究方向与实施建议。

关键词:数据中心;负荷建模;能耗预测;电力调度;多源数据融合

中图分类号: TN919.5 **文献标识码:** A **国家标准学科分类代码:** 470

Electricity computing collaboration: Data center power configuration, load modeling and prediction, power dispatching

Wang Zhixin^{1,2} Zou Yihan¹ Wang Jian³ Chen Hui¹ Qian Jin¹ Gong Chunyang¹

(1. School of Artificial Intelligence, Shanghai University of Electric Power, Shanghai 201306, China;

2. School of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

3. Institute of Smart Energy Innovation, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: With the widespread deployment of artificial intelligence (AI) applications and the rapid expansion of digital infrastructure, data centers have experienced continuous growth in scale and computing demand, making energy consumption management a critical issue in the energy and power engineering domain. Distinct from conventional electricity consumers, data center loads exhibit unique operational characteristics, with total energy consumption dominated by information technology (IT) equipment and cooling systems. The strong non-linearity, multi-source coupling, and dynamic evolution of these components pose significant challenges to make accurate energy consumption forecasting and power system dispatch. The paper systematically reviews the research progress in data center power supply configurations, load modelling, and energy consumption forecasting. Particular emphasis is focused on a comparative analysis of statistical models, machine learning, and deep learning, which are applied to IT load and cooling energy prediction, applicable conditions, feature construction strategies, and modeling performance, etc. Based on extensive literature analysis, key feature selection principles and model adaptation patterns under multi-source monitoring data are summarized, highlight the role of workload characteristics, thermal environmental variables, and operational control parameters to improving prediction. Furthermore, data center participation mechanisms in power system operation, dispatch based on load and energy consumption forecasting are reviewed also. From both temporal and spatial perspectives, the influence of forecasting on demand response implementation, flexible regulation capability

收稿日期: 2025-12-08 Received Date: 2025-12-08

* 基金项目: 上海市科学技术委员会科技计划(24DZ3001500)、基于低碳数据中心“算力-电力”协同驱动的海上虚拟电厂优化运营策略及政策建议(25692109800)、现代电力系统仿真控制与绿色电能新技术教育部重点实验室(东北电力大学)开放课题(MPSS2025-01)项目资助

extraction, and coordinated operation between power system and computing resource is analyzed. The results indicate that high-resolution, reliable load forecasting is a prerequisite for unlocking the flexibility potential of data center and enhancing power-computing synergy in modern power system. Finally, considering practical deployment requirements, the paper identifies key challenges faced by current research, including insufficient multi-source data fusion, limited model interpretability, and constraints in real-time forecasting and decision-making. Cutting-edge and future research directions and implementation strategies are proposed to support intelligent energy management and coordinated power-computing operation of data centers, providing systematic references and methodological insights for researchers and practitioners in this field.

Keywords: data center; load modelling; energy consumption forecasting; power dispatching; multi-modal data fusion

0 引言

人工智能(artificial intelligence, AI)迅速发展和普及应用,算力需求量激增,数据中心(data center, DC)数量与规模呈指数级增长态势。DC 能耗主要来自 IT 设备和冷却系统,其中,IT 设备包括服务器、存储和网络设备,其消耗与 CPU/GPU 使用率、内存占用、磁盘 I/O 及网络流量密切相关^[1-3];冷却系统能耗则受服务器热负荷、机房布局、空气流动、环境温湿度及其耦合影响。

DC 规模与能耗激增具有如下特征:1)算力需求爆发式增长推动 DC 成为电力系统中占比持续攀升的核心负荷,其非线性、多源耦合能耗特性影响电网平稳运行^[4];2)赋予 DC 独特调控潜力,具备成为需求侧响应资源与灵活调控单元,具有时间-空间两个维度的灵活性^[5-6],确保 DC 能够在保障自身服务质量(QoS)与服务等级协议(SLA)前提下,深度参与电力系统需求侧响应或作为灵活性资源接受调度,实现冷却系统优化、数字化监控和余热回收应用。

AI 技术应用为负荷预测与电力调度提供了新的手段。传统统计学与解析建模方法能够在一定程度上描述 DC 的能耗规律,但往往难以应对其高度非线性、多维度耦合与动态变化的特性^[7]。机器学习和深度学习模型则展现出更强的建模能力,能够捕捉 CPU 利用率、网络流量与冷却负荷等多源指标之间的复杂时序关系,提升预测精度与鲁棒性,在多变量耦合预测和不确定性建模方面展现出显著优势,为 DC 能耗的动态建模与多时间尺度预测提供了有效支持。DC 的运行优化从经验驱动逐步走向数据驱动与智能决策。传统调度策略往往侧重单目标优化,如能耗最小化或任务完成时间最短化,而目前的研究大多在此基础上实现多目标的权衡与动态调节^[8]。在这种预测与调度之间的耦合机制下,使 DC 能够以更灵活的方式参与电力系统运行,成为电网调节的重要资源^[9-11]。在这种“精准预测-智能调度”的机制下,DC 得以以更灵活的方式深度融入电力系统运行,成为电网灵活调控的重要资源。

本文从负荷预测与电力调度两个关键环节入手,对 DC 运行数据进行深度特征挖掘,系统总结并对比近年来

用于 DC 能耗预测与电力调度的典型模型,分析其在不同应用场景下的适用性与实际效果差异。在此基础上,进一步从电-算协同视角探讨预测模型对调度策略设计与系统灵活性提升的支撑作用,明确 DC 的调节潜力与功能定位。最后,结合工程应用需求,展望 DC 能耗预测与电力调度技术在数据融合、模型可解释性与实时决策等方面面临的挑战与未来研究方向,以期为该领域的研究人员提供参考。

1 DC 电源配置

DC 电源配置涉及能量路由、电能质量管理、储能优化与柔性响应等多维度内容扩展。随着大规模 AI 计算、液冷技术与高密度机柜的普及,DC 电源系统呈现出更高的瞬时峰值、更复杂的能量耦合路径以及更强的动态特征,对配电网络、不间断电源(uninterruptible power system, UPS)架构及储能系统提出新的要求。

1.1 DC 电源配置基本要求

1) 主电源架构与冗余机制

DC 通常采用多级配电架构,从公用电网或多条市电接入,通过配电柜,再经过 UPS、不间断电源系统、静态转换开关(STS)、配电单元(PDU)和机柜级供电线路,最终供给服务器、存储与网络设备。这样的分层结构能够兼顾高功率密度、灵活扩容与可靠性^[12]。

为满足高可靠性需求,DC 普遍采用冗余供电设计。常见方案包括 $N+1$ 、 $2N$ 、 $2N+1$ 等配置。 $N+1$ 表示在满足全部负载的基础之上额外配备一个备份模块; $2N$ 表示双路独立系统,每一路均能承载全部负载。这样的冗余设计可以在某一路发生故障或检修时,保证另一侧电源继续不间断供电。而在“双母线+UPS 冗余”系统中,可显著提高 DC 的可用性与容错能力,对抗单点故障导致的服务中断风险。以某中大型 DC 为例,该配置使得即使进行了局部维护,也能保持全部 IT 设备不间断运行^[13]。因此,冗余和分级设计已成为现代 DC 电源系统的标准配置,是保障高可用、高可靠的重要基础。

2) 电源分配与机柜供电

DC 电源分配单元(power distribution unit, PDU)

或母线将变压器输出的电力分配到各个机柜,高密度 DC 采用母线+高压配电减少线损、增强扩容灵活性。

机柜通常配备两个独立电源回路,对应两个 PDU 或供电链路,服务器则采用双电源设计分别接入两个回路,消除单一路线或 PDU 单点故障风险,提升供电可靠性^[14]。IT 设备功率密度持续上升,PDU 和配电系统要求越来越高,包括每机柜多分支回路、功率监控与计量、热插拔开关/断路器、以及易于扩容/改造/负载重分配等能力,提出将母线供电与模块化 PDU 结合,以支持 30 kW 以上的供电需求,同时确保灵活性与效率^[15]。

1.2 绿色 DC 电源配置

DC 开始考虑整合电、气、热、冷等多种能源,实现不同能量之间的耦合互补与高效利用,系统能够将传统孤立的供能方式转变为一体运营模式,从而提升整体能源效率以实现能源结构多样化、供电可靠性和碳排放降低的双重目标。在多能耦合架构中,光伏、风电与生物质能等可再生能源通过升压变压器提升至 10 kV/35 kV,汇集至主供电母线构成向 DC 供电的主要能源路径之一。同时,在可再生能源出力富余时,部分电能被引导至制氢单元,用于电解制氢与氢气储存。所储氢在需要时通过燃气轮机等氢能发电装置回转为电能,并重新并入主供电母线,满足 DC 瞬时功率需求,实现短时能量平衡与峰谷削减功能。电源接入 DC 后,经多级降压变压器逐级转换至 0.4 kV 低压交流等级,在高压直流 (HVDC) 供电系统中转化为当前国内 DC 普遍采用的约 24 V 直流电压,最终输送至机柜内 IT 设备,形成从多能耦合与电力汇集、到中压输配与直流适配,再到末端 IT 负载供电全链路能源体系(图 1)。

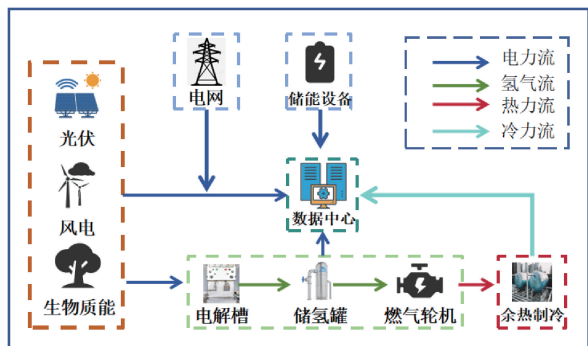


图 1 绿色 DC 多能耦合供能架构

Fig. 1 Multi-energy coupling energy architecture of green DC

文献[16]提出将光伏、风电、柴油发电机、储能电池与传统供电系统混合配置,通过多目标优化来设计 DC 供电系统,系统性比较不同混合配置,评估其在可再生渗透率与经济性上的特点。文献[17]分析了不同 IPS 架构

对功率损耗和服务器可用性的影响,随着服务器数量和电源供应单元(power supply unit, PSU)数量的增加,内部配电系统损耗将上升,若不当设计可能导致 PDU 过载、服务器断电或服务可用性下降。

可再生+储能+混合电源集成设计,使 DC 能够在电网供应不足、可再生能源出力波动或极端气候情况下,通过内部储能维持运行,降低对传统市电的需求,为未来“电-算协同”与电力系统灵活性奠定了基础。

2 DC 负荷建模

DC 负荷建模是对 DC 内部各种电力需求在时间与条件变化下的描述与预测,负荷模型为能耗预测与电力调度的基础,是理解 DC 能耗动态、调整资源配置、提高效率 and 降低能耗的关键环节^[18-21]。DC 负荷分为 IT 设备、制冷与其他负荷等 3 种类型。

DC 能源使用效率用于量化 DC 能耗,表征 DC 消耗的所有能源与 IT 负载使用的能源之比,与 DC 基础架构效率成反比^[22-24],如式(1)所示。

$$PUE = \frac{P^{DC}}{P^{IT}} \quad (1)$$

式中: P^{IT} 为 IT 设备的总能耗; P^{DC} 为 DC 的总能耗。

除了采用 PUE 值进行评估能耗,DC 能耗也可以从内部的用电设备获得。DC 内部的用电设备主要分为 IT 设备、冷却设备和电气设备^[25] 3 类。其中,IT 执行数据传输、存储和处理,冷却设备维持服务器温度,电气设备用于 DC 内部照明、供配电^[26]。

$$P^{DC} = P^{IT} + P^{CL} + P^L \quad (2)$$

式中: P^{CL} 表示冷却设备的总能耗; P^L 表示电气设备总能耗。

2.1 IT 负荷模型

IT 负荷作为核心能耗来源,通常占总能耗的 50% 左右。其主要包括服务器、存储设备与网络交换设备等的功率消耗。IT 负荷随业务请求量、资源利用率及任务分布动态变化,表现出显著的非线性与时变特征^[27-29]。当任务量上升或任务复杂度增加时,相关硬件资源利用率随之上升,从而显著拉动整体电力需求。当前,对 IT 设备能耗进行建模的主流方法主要包括加性模型、统计回归模型、分段线性化模型,回归能耗模型。IT 设备能耗建模模型对比如表 1 所示。

1) 加性模型

该模型假设 IT 设备总能耗可由其各内部部件能耗的代数和如式(3)所示。

$$P_{IT} = P_{CPU} + P_{mem} + P_{disk} + P_{net} + \Delta P \quad (3)$$

式中: P_{CPU} 、 P_{mem} 、 P_{disk} 、 P_{net} 分别表示处理器、内存、磁盘及网络接口的能耗; ΔP 为模型修正项。

表 1 IT 设备能耗模型对比

Table 1 Comparison of IT energy consumption models

模型	应用场景	模型特点	优点	局限性
加性模型	单台服务器或节点	处理器、等能耗,线性叠加	能耗拆分与能效评估	描述动态能耗与非线性难
统计回归	多设备能耗预测	多元拟合能耗与资源利用率	平均运行状态下能耗水平	依赖训练样本
分段线性	非线性能耗拟合	独立线性函数拟合	高负载区间性能退化与能效下降	结构复杂,泛化能力较弱
回归能耗	DC 建模与能效评估	CPU 利用率预测因子	应用于机房能耗预测与能效分析	难以捕捉短时间尺度能耗

2) 统计回归模型

统计回归模型假设 IT 设备能耗由静态能耗与动态能耗两部分如式(4)所示。

$$P_{IT} = P_{base} + \beta_1 U_{CPU} + \beta_2 U_{mem} + \beta_3 U_{disk} \quad (4)$$

式中: P_{base} 为静态能耗; U_{CPU} 、 U_{mem} 、 U_{disk} 分别表示 CPU、内存及磁盘利用率; β_i 为回归系数。

3) 分段线性化模型

由于 IT 设备在不同负载区间下的能耗响应存在非线性特征,线性模型难以准确描述其在高负载区间的性能退化如式(5)所示。

$$P_{IT} = \begin{cases} P_{idle} + k_1 U_{CPU}, & 0 \leq U_{CPU} \leq u_1 \\ P_{max} + k_2 U_{CPU}, & u_1 \leq U_{CPU} \leq 1 \end{cases} \quad (5)$$

式中: U_{CPU} 为 CPU 利用率; P_{idle} 与 P_{max} 分别为空闲与满载能耗,对应分界点 u_1 处的拟合系数 k_1 、 k_2 一般由实测

数据确定。

4) 回归能耗模型

文献[30]提出 DC 统计回归能耗模型,是业界最常用的 IT 设备功率线性拟合模型,如式(6)所示。

$$P_{server} = P_{idle} + (P_{max} - P_{idel}) \cdot u_{cpu} \quad (6)$$

2.2 制冷负荷模型

DC 总能耗中制冷系统约占 35%~45%,包括冷水机组、风机、水泵与冷却塔等设备。由于制冷系统运行受外部气象条件、IT 热负荷以及控制策略等多因素影响,其能耗模型具有显著的非线性和时变性^[31-33]。现有研究通常将制冷负荷模型划分为经验模型、线性回归模型、热力学机理模型和混合智能模型 4 类^[34-37]。对比各种 DC 制冷负荷模型结果如表 2 所示。

表 2 DC 制冷负荷模型对比

Table 2 Comparison of DC cooling load models

模型类型	应用场景	模型特点	优点	局限性
经验模型	早期能耗评估与能效标定	线性回归或多元回归建立制冷负荷与 IT 负载、环境温度差之间关系	结构简单、参数易标定,制冷与负载、温差线性关系	未考虑制冷机理及动态特征,环境变化不敏感
线性模型	负荷分区控制与能效分析	IT 负载区间制冷分段线性拟合,刻画不同运行区间能效变化	较好刻画不同负载区间制冷效率变化趋势	分段连续性弱,难以捕捉快速变化
热力模型	精细化仿真与冷却系统优化	考虑冷凝温度、蒸发温度及工质特性等因素	环境与设备热力学参数变化,具有物理解释性	模型复杂,参数获取困难,工质与结构依赖强
智能模型	大规模 DC 能耗预测与自适应控制	结合数据驱动与物理机理,采用深度学习对多维输入非线性映射	能同时捕捉多维非线性与时序动态关系,预测精度高	依赖训练数据,解释性较弱,模型泛化性

1) 经验模型

经验模型往往需要大量的监测数据,通过拟合冷却系统输入输出关系来建立能耗估算公式^[38]如下:

$$P_{cool} = a_0 + a_1 P_{IT} + a_2 (T_{out} - T_{in}) \quad (7)$$

式中: a_0 、 a_1 、 a_2 为回归系数; T_{in} 、 T_{out} 分别为送风与环境温度。

2) 线性回归与分段线性化模型

线性模型假设冷却系统能耗随 IT 负荷变化呈比例增长关系^[39]为:

$$P_{cool} = \beta_0 + \beta_1 P_{IT} \quad (8)$$

式中: β_0 为冷却系统基准能耗; β_1 反映制冷能耗随 IT 负荷增长的灵敏度。

实际运行中,当环境温度较高或负载接近上限时,制冷效率下降导致该关系非线性,分段线性化模型^[40]公式如下:

$$P_{cool} = \begin{cases} k_1 P_{IT} + b_1, & 0 \leq P_{IT} < P_1 \\ k_2 P_{IT} + b_2, & P_1 \leq P_{IT} < P_2 \\ k_3 P_{IT} + b_3, & P_{IT} > P_2 \end{cases} \quad (9)$$

式中: P_1 、 P_2 为负载分界点; k_i 为不同负载区间的拟合系数。

3) 热力学机理模型

热力学模型从物理层面描述冷水机组、冷却塔与空气处理单元的传热与能量^[41]平衡过程。冷水机制冷功

率的基本表达式为:

$$P_{chiller} = \frac{Q_{load}}{COP} \quad (10)$$

式中: Q_{load} 为需移除的总热负荷,通常由 IT 设备能耗决定;COP 为制冷机性能系数,其值随环境温度变化而变化^[42]:

$$COP = c_1 - c_2(T_{out} - T_{evap}) \quad (11)$$

式中: T_{out} 为冷凝温度; T_{evap} 为蒸发温度; c_1 、 c_2 为实验系数。

制冷能耗与环境条件及 IT 负荷间的耦合关系反映外界气候条件对制冷能耗的影响^[43]:

$$P_{cool} = \frac{P_{IT}}{c_1 - c_2(T_{out} - T_{evap})} \quad (12)$$

3 DC 负荷能耗预测

随着 DC 规模快速扩张以及异构算力和高密度服务器的普及,DC 的整体能耗管理与电力调度面临日益复杂的挑战。在能源消耗构成中,IT 设备能能耗和冷却系统能耗是最主要的两类目标,分别反映了计算负载运行和环境控制的电力需求^[44-47]。准确预测这两类能耗不仅是评估 DC 总能耗行为、优化资源调度和节能控制的基础,也是支撑“电-算协同”管理的关键前提。

3.1 DC 的预测目标

IT 设备能耗主要来源于服务器、存储与网络设备,其能耗变化与 CPU 使用率、内存占用、磁盘 I/O、网络流量等资源指标密切相关。设第 i 台服务器在时间 t 的能耗为 P_{it}^T ,其整体 IT 设备能耗预测可表示为所有服务器能耗的总和^[48]:

$$P_t^T = \sum_{i=1}^N P_{it}^T = \sum_{i=1}^N f_{IT}(CPU_{i,t}, Mem_{i,t}, \dots) \quad (13)$$

式中: f_{IT} 表示将服务器多维资源指标映射为能耗的预测函数; N 为服务器总数。

DC 能耗预测,不同系统环境下对 IT 设备能耗与冷却系统能耗的准确预测是构建高效资源调度、评估总能耗行为及支撑电算协同的关键基础^[49]。对于风冷主导的 DC,总能耗如式(14)所示。

$$P_{total}(t) = P_{IT}(t) + P_{cooling}(t) + P_{overhead}(t) \quad (14)$$

式中: $P_{it}(t)$ 表示 t 时刻的 IT 设备能耗; $P_{cooling}(t)$ 表示风冷冷却系统的能耗;而 $P_{overhead}$ 涵盖涵盖照明、配电损耗、UPS 损耗等辅助部分。在实际运行中,冷却能耗在总能耗中通常占有相当比例,一些研究指出在典型工况下,冷却部分可能占据 30%~40%甚至更高的份额。

3.2 预测模型

面向 DC 能耗预测采用的方法主要分为传统和 AI 两类方法、展现出不同适配性^[50-52]。

1) 传统预测方法

包括基于统计学和时间序列分析的模型,通过对历史能耗数据的趋势、周期与波动特征进行建模,实现功率或能耗短期预测。模型结构清晰、参数可解释性强,适合于数据维度较低、波动规律相对稳定场景^[53-54]。

服务器负荷能耗约占 DC 总能耗的 45%~50%,其变化与服务器 CPU 利用率、内存占用率、磁盘 I/O、网络活动等资源指标密切相关。传统研究多利用统计回归与自回归模型对这些指标与能耗间的关系进行拟合。采用多变量线性回归模型,将 DC 能耗与气象参数及运行指标建立线性映射关系,从而实现能耗预测,部分研究人员通过捕捉功率序列的自相关性实现短期预测。这类模型实现简单、计算成本低,适合在 DC 运行稳定、负载规律清晰的情境下使用。但由于 IT 负荷具有显著的非线性与随机性特征,当任务类型、计算密度或虚拟化策略变化较快时,基于平稳趋势的传统模型易出现滞后或精度下降。此外,这类模型依赖大量人工特征选择与参数调优,缺乏自适应与泛化能力。

冷却系统能耗约占 DC 总能耗的 40%~45%,受环境温度、湿度、气流组织、冷却模式及 IT 热负荷等多重因素影响。传统冷却负荷预测方法基于多元线性回归、灰色预测及时间序列回归模型,优点在于参数透明、解释性强,能够反映冷却系统在稳定运行阶段的能耗规律;但无法充分捕捉冷却系统的非线性耦合与动态特征,难以应对高频任务切换、变负载冷却控制或气候条件突变等复杂场景。

2) AI 预测方法

随着大型 DC 中算力负荷的持续爆发、硬件密度不断提升以及 GPU 驱动任务的急剧增加,基于 AI 方法的 DC 能耗预测模型显得日益关键。传统的统计预测方法逐渐无法满足非线性、多源数据、高频变化背景下的需求。以机器学习和深度学习为主的 AI 预测模型因其在时序特征抽取、非线性关系拟合等方面的优势,成为当前研究的主流。

在服务器能耗预测方面,传统的统计或静态能耗模型多聚焦于服务器规格、历史使用率,但在高动态、多任务插入、算力-电力耦合背景下,其表现仍受限。提出了一种离线能耗预测方法,预测特定应用程序在目标服务器上不同负载水平下的能耗,通过它们在不同物理服务器上的能耗来评估该方法的性能。通过神经网络使用服务器硬件规格输入特征,对空载能耗与满载能耗进行预测。在特征设计方面,服务器能耗预测模型通常纳入多个维度。硬件层面:服务器型号、CPU/GPU 类型、使用率、功率状态等。任务负载特性:任务类型、负载突变频率、并发任务数、任务持续时间。电力耦合相关:输入电压、电源效率、服务器间散热影响。在算法模型构建方

面,已有研究开始采用 AI 技术,识别硬件与负载特征与能耗之间的非线性映射。文献[55]提出基于多变量线性回归、支持向量回归(SVR)与人工神经网络(ANN)的 3 类主机能耗预测模型;通过基准测试采集主机性能与能耗的关联数据,融合回归分析与机器学习方法完成模型构建,依托 CloudSim 平台仿真验证资源管理算法有效性。为云 DC 提供兼具高精度与强普适性的能耗建模方案,可为能效评估与优化提供核心支撑,所提资源管理算法可直接落地于实际 DC 场景,有效降低运营成本与碳排放量,助力绿色计算理念落地与行业可持续发展。

IT 负荷能耗预测主要利用机器学习与深度时序模型从多维监测指标中自动学习能耗特征。ESTA-IoT 平台研究天气状况与 DC 能耗之间的相关性,采用多元线性回归模型,通过将服务器负载指标与机房环境参数共同纳入建模,实现了对整体能耗趋势的预测,其结果表明在负载波动相对缓慢的情况中线性模型具有较高适用性。与此同时文献[55]以自回归差分移动平均模型(autoregressive integrated moving average model, ARIMA)为基础,利用功率序列的自相关结构建立时间序列预测模型,同时测试了朴素方法、线性回归、长短期记忆网络(long short-term memory, LSTM)和 Prophet 方法,提出了一种使用决策树和一组独特的条件属性来理解预测模型误差的方法,实现了对 IT 能耗的趋势跟踪。这表明在功率变化具有明显周期性或延迟性的场景下,时间序列建模可以获得较为稳定的预测效果。通过分析多规模 DC 能耗特征,构建融合 LSTM 和卷积神经网络(convolutional neural networks, CNN)的混合深度学习模型实现高精度动态能耗预测并提出自适应工作负载调度、智能制冷控制和服务器动态管理等优化策略。然后通过仿真与实际部署,验证模型预测精度和优化策略节

能效果。指出优化策略在正常负载下降低 PUE 值 1.32,平均节能 12.5%;极端情况下仍稳定运行。提出了基于经验模态分解和 Prophet 的服务器能耗预测模型。对于单一的服务器能耗序列,使用 Prophet 模型对服务器能耗进行预测,同时采用了经验模态分解对服务器能耗序列进行分解,然后再对分解的分量分别建立 Prophet 模型进行预测,最后累加各分量预测值,得到服务器能耗预测结果。

在 DC 能效优化的研究中,冷却系统能耗预测是整体能耗建模与能效评估的核心支撑环节,其预测性能的优劣直接决定了后续节能策略的制定精度与落地效果。特征工程作为冷却能耗预测模型构建的基础,环境变量因直接影响冷却系统的换热效率与运行负荷,特征设计方面,冷却能耗预测模型一般考虑多个维度:环境变量:室外温度、湿度、冷却水温度差、冷却塔入口温度等。冷却系统运行参数:压缩机频率、冷却水泵流量、风扇转速、热通道/冷通道配置、机房气流管理等。文献[56]回顾了冷却系统能耗建模方法,提出服务器热负载:服务器能耗提升会带来更大热负荷,从而影响冷却系统的能耗需求。在算法模型构建方面,冷却能耗预测以及优化研究也常常采用 AI 方法。文献[57]指出,DC 冷却系统优化,机器学习用来分析能耗模式、预测负载并动态管理设备能耗。性能评估与部署,冷却能耗预测模型需满足实时或近实时响应能力、良好适配环境变化、与服务器负载变化联动、并可执行控制策略。

AI 方法在 DC 能耗预测领域呈现出“由浅入深、由单一到融合”的趋势。早期浅层模型注重特征回归与模式识别,而深度模型则关注时间依赖与多源数据融合。当前研究主要方向正逐步从单纯的预测精度提升,转向可解释性、泛化能力与实时性的综合优化。表 3 所示为基于模型的 DC 负荷能耗预测应用。

表 3 基于模型的 DC 负荷能耗预测应用

Table 3 Application of load and energy consumption prediction based on models in data centers					
文献	模型及算法	适用场景	优点	局限性	误差
[55]	Prophet	强周期性时间序列预测	门槛较低,强可解释性,支持机理分析	非线性拟合,不适合强耦合	MAPE: A 9.9%, B 43.90%, C 18.17%, D 11.1%
	LSTM	时序依赖显著	捕捉长时间依赖与动态变化	训练代价高,需大量历史数据	MAPE: A39.69%, B37.44%, C 70.31%, D 41.53%
[56]	集成神经网络模型	DC 高精度预测场景	鲁棒性强	模型复杂度高,训练与预测成本高,算力需求大	MAPE:7.21%
	强化学习	DC 动态资源调度,冷却系统自适应优化场景	动态环境适应能力强	样本效率低,训练周期长	降低约 18.2%冷负荷

4 DC 电力调度策略

随着云计算与大规模算力服务的快速发展,DC 规模

持续扩大并呈现显著的时空分布性^[58]。DC 同时具备可调节、可迁移的潜在灵活资源属性,其与电网的协同已成为新型电力系统负荷侧灵活调控的关键路径^[59]。传统静态化、单目标导向的调度策略难以同时满足性能保证、

能效最优化与电网交互需求,因此近年来研究重心逐步向基于时间灵活性与基于空间灵活性多目标、耦合优化转移^[60-62]。从时间与空间维度,系统梳理 DC 参与电网侧调度方向,从“适用场景-优势-局限”的框架展开分析,明确各类方法的技术路径与应用边界。

4.1 基于时间灵活性的 DC 电力调度

基于时间灵活性的调度核心是通过“负荷时序重构”响应电网需求,其关键在于根据工作负载的时间约束差异设计差异化调度机制在 QoS 与 SLA 的前提下^[63],通过动态调整任务执行时序重构用电曲线,实现电网削峰填谷、需求响应或低碳时段负荷聚合。已有实证研究^[64-66]表明,该类调度可使电网负荷峰谷差削减^[67],同时提升 DC 电价套利与低碳收益,是 DC 参与电网互动的基础路径。本节按工作负载时间约束特性细分研究方向,系统分析各类方法的技术逻辑与电网价值。

1) 面向时间敏感型工作负载研究

时间敏感型工作负载通常包括在线服务、实时推理、搜索请求处理以及交互式网站服务,其特点为严格的 SLA 时延约束、对瞬时资源可用性极度敏感、需要快速调度与弹性伸缩能力。参与电网调度的核心问题是“时延保障”与“电网响应灵活性”的平衡。因此,针对时间敏感型任务的调度研究主要围绕:时延模型构建、资源动态分配、排队与拥塞避免、实时负载波动预测等方面展开。形成了算法优化、系统架构设计、策略对比 3 类技术路径,实现低延迟与能耗成本平衡^[68]。

针对时间敏感型负载的电网响应需求,文献[69]从算法层面提出多目标 QoS-DPSO 调度方法,将任务完成时间、可靠性与电网削峰需求纳入优化目标,通过离散粒子群优化在满足微秒级时延约束的前提下,实现负荷峰段柔性削减,为 DC 参与电网日间紧急需求响应提供了算法支撑。从系统架构切入,针对性解决“低延迟与高 CPU 利用率”的核心矛盾,提出 Shenango 调度系统通过细粒度核心重分配机制,在保障延迟敏感型应用延迟的同时,提升 CPU 利用率,间接降低 DC 对电网峰段电力的依赖度,减少峰谷负荷差。进一步拓展技术路径,对比强化学习、自适应批处理等调度策略在电网实时负载波动场景下的适配性,明确强化学习方法在动态跟踪电网调度信号中的优势,其自适应负载管理机制可使 DC 对电网调度信号的响应延迟缩短至秒级。时间敏感型工作负载的调度研究已形成“约束建模-方法优化-性能验证”的完整框架,实现了“时延保障”与“电网响应”的初步协同。但现有研究仍聚焦于中低速电网调度信号响应,对高动态电网调节需求、多资源维度协同调度的探索不足。

2) 面向时间延迟容忍型工作负载研究

时间延迟容忍型负载通常包括离线训练作业、批处理任务、大规模数据分析、备份任务以及部分异步服务。

这类任务的显著特征是在较大时间窗口中灵活安排;可分解、切片或迁移;对 SLA 的实时性要求较低,但对总体完成时间仍有约束;在调度时常与电价、碳排放和负荷峰谷联动。提出了一种多维资源联合调度框架。联合优化卸载选择、计算资源分配以及能耗分配,以最小化所有任务的加权延迟和能耗总和。其“可分批执行+在时间窗口内灵活调度”的机制为延迟容忍型任务提供了可直接用于削峰与低价时段迁移的基础。文献[70]提出 HIPPIE,一种混合并行训练框架,针对大规模训练任务提出了时间可变的资源分配策略,能够根据集群拥塞情况延迟或提前子任务执行,形成一种“任务灵活分配+等待队列调整”的机制。文献[71]提出了基于碳信号的调度框架,通过动态调整训练任务在不同时间片的执行比例,使训练在低碳时段集中进行。该工作直接面向 DC,且明确利用了任务的时间弹性,为时间延迟容忍负载参与电网低碳调度提供了可操作的方案。

4.2 基于空间灵活性的 DC 电力调度

空间灵活性指在地理上分布的 DC 之间迁移或重定向计算任务与电力负荷的能力。通过地理迁移,运营方可以将计算负荷从高成本/高碳/拥堵网络区域转移到电价低、可再生资源丰富或输电网络冗余的地区,从而实现电网层面的输电拥堵缓解、可再生并网平衡与成本最优。空间柔性的研究涉及电力系统的输配电约束、传输潮流影响以及跨区域市场机制的协调。

1) 面向配电网场景的空间灵活性优化

在配电网层面,空间灵活性主要依赖 DC 在城市或区域范围内的任务迁移能力,通过在本地集群之间重定向负载实现对局部电网的支撑,例如,减轻某一馈线的拥塞、利用某区域的分布式光伏、电价差异化调度等。因此研究重点包括局地电网约束、节点电价(DLMP)、分布式能源耦合与任务跨站点调度策略。文献[72]对极端故障下灵活配电网(FDN)中互联网 DC(IDC)的供电恢复问题,提出一种时空调节协调恢复方法。通过制定 IDC 工作负载的时空转移策略和基于灵活站(FS)的网络重构策略,建立混合整数二阶锥规划模型,实现源、网、储、荷侧的时空协同,以最小化 IDC 数据丢失损失。文献[73]提出一种考虑 DC 建筑(DCB)时空灵活性的配电网能源管理模型,整合多能源枢纽与储能,通过建立配电网与多能源枢纽的耦合模型,结合 DCB 的时空负荷转移特性及余热回收利用,采用鲁棒优化处理不确定性。表明了 DC 建筑的时空灵活性,改善了调度条件。文献[74]提出了基于分段线性化的 IT 设备数据-能耗模型,描述 DC 的能耗特性,考虑时空负载调节的潜力,构建了 DC 的运营灵活性分析模型,以适应 ADN 在复杂环境中的运营需求。案例研究表明,通过工作负荷的时空调节,IDC 的能效可以有效提升。IDC 的灵活调度还能

减少 ADN 的电压违规和馈线负载不平衡,从而促进为 IDC 提供高质量的电源。

配电网场景的空间灵活性研究普遍强调“节点差异化电力条件+任务跨节点迁移机制”,无论是针对 DC 本身还是边缘节点,研究直接为 DC 在配电网的空间柔性调度提供理论基础。

2) 面向输电网场景的空间灵活性优化

空间灵活性涉及跨区域 DC 之间的任务迁移,通过响应跨区电价(LMP)、碳强度、输电拥塞情况以及可再生能源时空分布,实现大规模负荷的地理优化。研究核心大多关注:跨区传输约束、区域电价差异、可再生能源、碳驱动与任务迁移成本等问题。提出了跨区域 DC 之间的碳强度驱动调度方法,通过实时碳信号决定任务分布比例,以利用不同区域的低碳电源,体现了输电层级的空间柔性跨区域电力差异。通过不同的地理分布的云 DC,来优化利用廉价低碳能源驱动的计算能力,解决了工作负载不平衡的问题。为应对多目标调度的挑战,提出了一种基于多代理强化学习和 MASAC 方法的算法,通过与基于真实工作负载模式、能源价格和碳强度构建的云系统交互,学习最优的协作调度策略,提高了 GPU 利用率、降低能源成本和碳排放。建立了跨区域迁移对输电拥塞的影响模型,并提出在输电约束下优化 DC 任务迁移策略,实现跨区域需求响应,明确描述了“空间柔性+输电网耦合”工作。通过建模多个区域的风光出力、传输成本与任务迁移延迟,提出可再生驱动的跨区域调度机制。虽然部分区域并未包含大型 DC,但采用了分布式计算节点,整体方法可直接迁移至 DC 层级。输电网尺度的空间灵活性依赖跨区域资源差异,通过地理迁移实现“电价差异利用、可再生匹配”,形成 DC 参与新型电力系统的重要途径。

5 DC 应用与电算协同技术发展态势

5.1 DC 应用案例

1) 通信领域

中国电信中部云计算大 DC 采取机柜紧凑布置、及时关闭空闲机柜通风口等措施,年节电约 150×10^4 kWh;中国联通三江源国家大数据基地建立了基础设施能耗统计与巡检制度,通过定时检查空调、电力系统确保设备高效运行,PUE 为 1.16;中国电信兰州新区大数据产业园 1 机楼通过优化控制逻辑、延长自然冷却时长,年节电约 120×10^4 kWh;中国移动 DC 通过调节蓄冷罐回水温度与风机转速精细控制冷量;中国电信数字青海绿色大 DC 依靠气候优势采用间接蒸发冷却技术,全年自然冷却达 314 d;中国电信云计算重庆基地水土 DC 二期构建了多源复合制冷架构,动态调节温控设置;中国电信集团有限

公司云计算(青岛)基地 A2 数据机楼通过大数据建模优化设备运行参数,对 5 000 台设备进行精细化监控与休眠控制,PUE 为 1.25,提升 DC 能效水平。

2) 互联网领域

阿里巴巴浙江云计算 DC 仁和 C 楼、乌兰察布基地前旗 DCA 楼,采用系统性节能诊断体系和持续优化运行策略,PUE 为 1.17;润泽 DCA-18 通过冷冻水泵变频改造、电力室加装风管及水冷空调改造等,年节电约 $1\,300 \times 10^4$ kWh;润泽 DCA-7 采用冷板式液冷技术,PUE 为 1.10;抖音-中联绿色大数据产业基地 5 号楼采用浸没式液冷技术提高机柜运行效率;杭钢云计算 DC 利用 $34\,000\text{ m}^2$ 屋顶分布式光伏,年发电量超过 600×10^4 kWh;万国数据浦江数据中建立水环热泵系统为办公楼供热;浩云成都高性能云计算基地二号采用 DR4 配电架构提升负载率;五象云谷云计算中心通过分散布局电气设备缩短供电路径;中原 DC 项目 3 号 DC 与浩云北京 3 号 DC 建立回收体系和使用环保冷媒降低环境影响。

3) 公共服务领域

公共服务机构绿色 DC 通过集约化建设、绿色设计及智能运维等系统性措施实现能效提升。数字福建云计算中心整合 200 多个省级部门资源,促进业务协同与数据共享;济宁政务云 DC 改造闲置工业厂房缩短建设周期;雄安城市计算中心将机房主体建于地下,利用自然冷源与智能运维平台,实现冬季 PUE 低至 1.18;国家税务总局山西省税务局 DC 通过扩展机柜、增设冷通道实现精准送风;西安建筑科技大学 DC 采用微模块冷通道封闭设计,缩短送风距离;鹰潭高新区税务局 DC 通过一体化机柜与智能监控,PUE 较投用初期减小 0.32;南京市浦口区城市数字治理中心应用 AI 技术联动控制设备运行;国家税务总局山西省税务局 DC 利用数字孪生技术实现多系统实时监测;鄱阳县税务局中心机房落实绿色采购制度,建立废旧产品回收体系,全面降低环境影响,实现公共机构高效、低碳运行模式。

4) 能源、金融、智算领域

山东省城市商业银行合作联盟金融云 DC 结合 CFD 建模与氟泵自然冷却技术,PUE 同比降低 6.5%;上交所金桥 DCD3 机房楼通过双冷源耦合自然冷源补偿、动态飞轮 UPS 等技术,年节电约 220×10^4 kWh;中国农业银行稻香湖机房采用“风冷+水冷”复合冷源架构,实现负载智能切换,年节电 400×10^4 kWh;海通证券张江 DC 在配电间部署智能巡检系统,实时采集多维度环境数据,提升运维效率;中央结算公司上海卡园 DC 利用虚拟化与云化技术构建弹性架构,并通过 AI 同步优化制冷与供电节电超 15%;国网安徽省电力有限公司 DC 自主研发监测工具,常态化清理“僵尸系统”减少算力浪费,节电 8%以上;国网安徽省电力有限公司 DC 机房进一步运用数字

孪生技术仿真能耗趋势,优化负荷可调比例,探索接入电网虚拟电厂平台,支持新型电力系统建设;中国电信安徽智算中心 A1 楼通过构建 3000 P 大模型训练集群与 300 P 液冷算力,结合任务调度优化使设备上架率达 86.1%,并融合多项 ISO 标准建立综合运维体系,以高质量管理保障智算服务效能。

5.2 电算协同技术研究与发展态势

在数字基础设施绿色化转型与“双碳”目标深入推进背景下,DC 能耗管理精细化、智能化水平已成为影响运行经济性与碳排强度的关键。为应对能效提升、电网协同与多目标优化多重挑战,构建具备高精度预测、可解释分析、实时优化与跨场景适应能力的电算协同系统,成为未来 DC 技术体系演进核心方向,凸显如下发展态势。

1) 多源感知高精度联合能耗预测体系构建

DC 监测系统的精细化发展,负荷状态实现了全场景、全过程与高频率的可观测,为构建高精度能耗预测模型奠定了坚实的数据基础。未来将在于进一步整合 IT 负荷、冷却系统运行参数、环境条件以及调控策略,实现多源信息的统一建模。融合深度学习与热物理机理模型混合式预测框架将成为主流,更精准描绘 IT 负荷与冷却负荷间复杂耦合关系,并在多时间尺度上实现稳定预测。同时,引入不确定性量化与置信度评估将成为预测体系重要组成,预测结果在电网互动与调度决策中可信用和可控性大幅提升。

2) 可解释性能耗建模与耦合机理研究

预测模型虽具备高精度,但黑箱性强、难以解释仍是制约落地应用的瓶颈。未来将从单纯结果预测转向“可解释的机理分析”,即在模型结构中显式融入热物理机制、冷却系统能效规律及服务器性能模型,实现从经验拟合到机理驱动的演进。可解释性技术将用于识别影响能耗变化的关键因素,解释冷却系统在不同负荷强度与环境条件下的行为空间。通过提升模型透明度,显著增强其在调控策略制定、故障诊断及运维优化。

3) 实时调控高效、多目标协同优化调度

未来将更深度嵌入电力系统调度,优化目标不再局限于能效或成本,而是能耗、运行成本、碳排放、电网互动及服务质量实现多目标、动态权衡。调度算法需具备以下特征:实时性强,能在秒级或分钟级完成调度计算,适应高频电网调度需求;多目标协同,利用分层优化、进化算法或滚动优化等框架处理多目标间的动态冲突;设备与功率联动,实现 IT 任务调度、冷却系统控制、储能系统充放电等多环节协同;电网互动能力增强,可根据价格、碳强度及网源荷状态进行任务迁移与负荷移峰。这种高效调度体系将推动 DC 从“被动负荷”转向“可调节资源”,提升电网可再生能源消纳能力,在源-网-荷-储协同中发挥更大作用。

4) 跨场景适应性电算协同技术

DC 部署环境、气候特征、冷却方式以及 IT 任务结构在不同站点间差异显著,使得模型在跨场景应用时存在明显的泛化性能下降。未来必须解决模型迁移能力不足的问题,通过迁移学习、领域自适应、少样本学习等方法减少模型对特定场景数据的依赖,使预测与调度系统能够在“跨机房类型、跨气候区域、跨任务结构”的条件下保持稳定性能。与此同时,构建标准化数据接口与可扩展模型框架,使系统在新增机柜、扩容集群或更换冷却架构后仍具备快速适配能力。最终目标是形成可在大规模、多区域算力基础设施中统一部署的协同运行系统,实现算法与工程实践的稳定协同。

6 结 论

现有研究围绕 DC 算力负荷与能耗行为,从负荷结构特征、功耗映射建模及时序预测方法等方面展开了较为系统的分析。相关文献普遍指出,DC 负荷以 IT 设备与冷却系统为主要构成,其功耗演化受算力任务强度、资源利用率及运行环境等多因素共同影响,呈现出明显的非线性和多时间尺度特征。尽管数据驱动模型在短期功耗预测精度方面取得一定进展,但受限于真实功耗数据获取困难、功耗标签构建方式差异较大,以及 IT 负荷与冷却能耗耦合关系刻画不充分等,现有模型在可解释性、跨场景适应性及工程适用性方面存在不足,难以直接支撑电力系统层面规划与调度分析。

面向算力负荷持续增长及电力系统低碳转型的背景,未来研究有必要在现有工作的基础上进一步深化。一方面,应结合监测数据、建模假设与物理约束,完善功耗映射与数据构建方法,提高功耗建模结果的可信度与一致性;另一方面,有必要从系统层面加强 IT 负荷与冷却系统的协同建模与联合预测,以更准确地反映 DC 整体能耗的动态演化特性。在此基础上,将 DC 负荷建模与预测结果进一步融入源-网-荷-储协同运行、电力系统调度及相关市场机制分析中,系统评估其作为可调节负荷资源的运行行为与调节潜力。

参考文献

- [1] 孙宏斌, 潘昭光, 郭庆来. 多能流能量管理研究: 挑战与展望[J]. 电力系统自动化, 2016, 40(15): 1-8. SUN H B, PAN ZH G, GUO Q L. Energy management for multi-energy flow: Challenges and prospects [J]. Automation of Electric Power Systems, 2016, 40(15): 1-8.
- [2] 梁晨, 曾博, 雷乐意, 等. 算力-电力联合市场下数据中心与配电网集成规划: 一种多目标区间-随机优化方法[J]. 电力系统保护与控制, 2025, 53(16):

- 120-135.
- LIANG CH, ZENG B, LEI L Y, et al. Integrated planning of data centers and distribution networks under the computing-electricity joint market: A multi-objective interval-stochastic optimization approach [J]. *Power System Protection and Control*, 2025, 53 (16): 120-135.
- [3] 曹雨洁, 丁肇豪, 王鹏, 等. 能源互联网背景下数据中心与电力系统协同优化 (二): 机遇与挑战[J]. *中国电机工程学报*, 2022, 42(10): 3512-3527.
- CAO Y J, DING ZH H, WANG P, et al. Coordinated operation for data center and power system in the context of energy internet (II): Opportunities and challenges [J]. *Proceedings of the Chinese Society of Electrical Engineering*, 2022, 42(10): 3512-3527.
- [4] 丁肇豪, 曹雨洁, 张素芳, 等. 能源互联网背景下数据中心与电力系统协同优化: 数据中心能耗模型[J]. *中国电机工程学报*, 2022, 42(9): 3161-3177.
- DINF ZH H, CAO Y J, ZHANG S F, et al. Coordinated operations of highly renewable power systems and distributed data centers [J]. *Proceedings of the Chinese Society of Electrical Engineering*, 2022, 42 (9): 3161-3177.
- [5] 李明睿, 朱介北, 俞露杰, 等. 实现新能源最大利用的多数据中心自适应负荷调度算法[J]. *中国电机工程学报*, 2025, 45(6): 2071-2084.
- LI M R, ZHU J B, YU L J, et al. Adaptive spatio-temporal tasks scheduling algorithm in regional interconnected data centers [J]. *Proceedings of the Chinese Society of Electrical Engineering*, 2025, 45(6): 2071-2084.
- [6] 张程, 江伟龙, 罗玉锦, 等. 考虑风光不确定性的多数据中心微网与共享储能协同优化运行[J]. *电力系统保护与控制*, 2025, 53(4): 72-84.
- ZHANG CH, JIANG W L, LUO Y J, et al. Collaborative optimal operation of multiple data center microgrids and shared energy storage considering wind power and photovoltaic uncertainty [J]. *Power System Protection and Control*, 2025, 53(4): 72-84.
- [7] ZOU Y H, CHEN H, WANG J, et al. Data center computing load forecasting based on feature modeling and the informer model [C]. 2025 The 10th International Conference on Power and Renewable Energy, 2025, 2: 745-750.
- [8] 徐婷, 张海峰, 杨昊, 等. 数据中心微电网能量管理随机优化模型研究[J]. *电子测量与仪器学报*, 2019, 33(8): 101-107.
- XU T, ZHANG H F, YANG H, et al. Research on stochastic optimization model of energy management for data center microgrids [J]. *Journal of Electronic Measurement and Instrumentation*, 2019, 33 (8): 101-107.
- [9] 史久根, 贾坤荣, 邾伟, 等. 基于软件定义的数据中心网络虚拟机迁移规划[J]. *电子测量与仪器学报*, 2018, 32(2): 119-127.
- SHI J G, JIA K Y, ZHU W, et al. Multi-controller deployment algorithm based on load balance in software defined network [J]. *Journal of Electronic Measurement and Instrumentation*, 2018, 32(2): 119-127.
- [10] YOO S J B. Prospects and challenges of photonic switching in data centers and computing systems [J]. *Journal of Lightwave Technology*, 2021, 40 (8): 2214-2243.
- [11] SAXENA D, SINGH A K, BUYYA R. OP-MLB: An online VM prediction-based multi-objective load balancing framework for resource management at cloud data center [J]. *IEEE Transactions on Cloud Computing*, 2021, 10(4): 2804-2816.
- [12] 王晴, 刘友波, 黄杨, 等. 考虑数据中心需求响应的城市电网阻塞管理[J]. *电网技术*, 2020, 44(8): 3129-3138.
- WANG Q, LIU Y B, HUANG Y, et al. Congestion management of urban power grid considering demand response of data center [J]. *Power System Technology*, 2020, 44(8): 3129-3138.
- [13] 王奖, 张勇军, 李立涅, 等. 数据中心园区能源互联网的关键技术与发展模式[J]. *中国工程科学*, 2020, 22(4): 65-73.
- WANG J, ZHANG Y J, LI L CH, et al. Key technologies and development modes of park energy internet in data centers [J]. *Strategic Study of CAE*, 2020, 22(4): 65-73.
- [14] HE Z G, DING T, LIU Y, et al. Analysis of a district heating system using waste heat in a distributed cooling data center [J]. *Applied Thermal Engineering*, 2018(141): 1131-1140.
- [15] LIU L J, SUN H B, LI C, et al. Exploring customizable heterogeneous power distribution and management for datacenter [J]. *IEEE Transactions on Parallel and Distributed Systems*, 2018, 29(12): 2798-2813.
- [16] RAVADANEH S N, OSKUEE M R J, KARIMI M. Multi-objective planning model for simultaneous reconfiguration of power distribution network and allocation of renewable energy resources and capacitors

- with considering uncertainties [J]. Journal of Central South University, 2017, 24(8): 1837-1849.
- [17] ZHU W T, HAN M Y, Milanović J V, et al. Methodology for reliability assessment of smart grid considering risk of failure of communication architecture [J]. IEEE Transactions on Smart Grid, 2020, 11(5): 4358-4365.
- [18] 杨挺, 姜含, 侯昱丞, 等. 基于计算负荷时-空双维迁移的互联多数据中心碳中和调控方法研究[J]. 中国电机工程学报, 2022, 42(1): 164-177.
- YANG T, JIANG H, HOU Y CH, et al. Study on carbon neutrality regulation method of interconnected multi-datacenter based on spatio-temporal dual-dimensional computing load migration[J]. Proceedings of the Chinese Society of Electrical Engineering, 2022, 42(1): 164-177.
- [19] 陈冬林, 马一帆, 邹安琪. 基于可再生能源消纳的数据中心负荷调度研究综述[J]. 计算机工程与科学, 2023, 45(5): 790-801.
- CHEN D L, MA Y F, ZOU A Q. A survey of data center load scheduling based on renewable energy consumption [J]. Computer Engineering & Science, 2023, 45(5): 790-801.
- [20] KATAL A, DAHIYA S, CHOUDHURY T. Energy efficiency in cloud computing data centers: A survey on software technologies [J]. Cluster Computing, 2023, 26(3): 1845-1875.
- [21] SHAFIQ D A, JHANJHI N Z, ABDULLAH A. Load balancing techniques in cloud computing environment: A review [J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(7): 3910-3933.
- [22] 冯斌, 胡轶婕, 黄刚, 等. 基于深度强化学习的新型电力系统调度优化方法综述[J]. 电力系统自动化, 2023, 47(17): 187-199.
- FEN B, HU Y J, HUANG G, et al. Review on optimization methods for new power system dispatch based on deep reinforcement learning [J]. Automation of Electric Power Systems, 2023, 47(17): 187-199.
- [23] 苏娟, 董彦君, 赵晶, 等. “东数西算”背景下多数据中心联合消纳可再生能源途径研究综述[J]. 高电压技术, 2024, 50(01): 55-64.
- SU J, DONG Y J, ZHAO J, et al. Research of the approach to consuming renewable energy across multiple data centers in the context of east data, west calculation[J]. High Voltage Engineering, 2024, 50(1): 55-64.
- [24] 李远征, 龙信鑫, 周纯杰, 等. 高比例新能源电网-分布式数据中心集群协同优化运行研究[J]. 中国科学:技术科学, 2024, 54(1): 119-135.
- LI Y ZH, LONG X X, ZHOU CH J, et al. Coordinated operations of highly renewable power systems and distributed data centers [J]. Scientia Sinica Technologica, 2024, 54(1): 119-135.
- [25] 王昊, 李晖, 宋端正, 等. 面向云-雾计算系统中的遗传算法任务调度研究[J]. 电子测量与仪器学报, 2023, 37(8): 40-51.
- WANG H, LI H, SONG R ZH, et al. Research on genetic algorithm task scheduling in cloud-fog oriented computing systems [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 40-51.
- [26] VASQUES T L, MOURA P, DE ALMEIDA A. A review on energy efficiency and demand response with focus on small and medium data centers [J]. Energy Efficiency, 2019, 12(5): 1399-1428.
- [27] 耿圣杰, 贾燕冰, 江珂滕. 电网数据中心服务器容量及其综合供能系统联合规划策略研究[J]. 电网技术, 2022, 46(9): 3281-3292.
- GENG SH J, JIA Y B, JIANG K T. Joint planning strategy of grid data center server capacity and its integrated energy supply system [J]. Journal Power System Technology, 2022, 46(9): 3281-3292.
- [28] RADOVANOVIC A, CHEN B K, TALUKDAR S, et al. Power modeling for effective datacenter planning and compute management [J]. IEEE Transactions on Smart Grid, 2021, 13(2): 1611-1621.
- [29] 孙湛冬, 焦娇, 李伟, 等. 基于改进蚁群算法的电力云数据中心任务调度策略研究[J]. 电力系统保护与控制, 2022, 50(2): 95-101.
- SUN ZH D, JIAO J, LI W, et al. A task scheduling strategy for a power cloud data center based on an improved ant colony algorithm [J]. Automation of Electric Power Systems, 2022, 50(2): 95-101.
- [30] GAO J C, WANG H Y, SHEN H Y. Task failure prediction in cloud data centers using deep learning [J]. IEEE Transactions on Services Computing, 2020, 15(3): 1411-1422.
- [31] 朱介北, 杨顺坡, 俞露杰, 等. 数据中心制冷系统即时消纳可再生能源的建模和变温控制[J]. 电力系统自动化, 2022, 46(20): 13-22.
- ZHU J B, YANG SH P, YU L J, et al. Modeling and variable temperature control for instant consumption of renewable energy in data center cooling systems [J]. Automation of Electric Power Systems, 2022, 46(20): 13-22.
- [32] CHEN X T, ZHOU L Y, ZHANG C B, et al. Research status and future development of cooling technologies for

- green and energy-efficient data centers [J]. Strategic Study of Chinese Academy of Engineering, 2022, 24(4): 94-104.
- [33] WAN J X, ZHOU J, GUI X. Intelligent rack-level cooling management in data centers with active ventilation tiles: A deep reinforcement learning approach[J]. IEEE Intelligent Systems, 2021, 36(6): 42-52.
- [34] 路煜, 张路, 侯小凤, 等. 数据中心功耗削峰电池的可用性分析[J]. 计算机工程与科学, 2021, 43(1): 9-16.
- LU Y, ZHANG L, HOU X F, et al. Availability analysis of datacenter peak power shaving battery[J]. Computer Engineering & Science, 2021, 43(1): 9-16.
- [35] LIN W W, WU G X, WANG X Y, et al. An artificial neural network approach to power consumption model construction for servers in cloud data centers[J]. IEEE Transactions on Sustainable Computing, 2019, 5(3): 329-340.
- [36] LIN R H, DENG Y H. Allocating workload to minimize the power consumption of data centers[J]. Frontiers of Computer Science, 2017, 11(1): 105-118.
- [37] ZHOU X, URATA R, LIU H. Beyond 1 Tb/s intra-data center interconnect technology: IM-DD OR coherent? [J]. Journal of Lightwave Technology, 2019, 38(2): 475-484.
- [38] OUHAME S, HADI Y, ULLAH A. An efficient forecasting approach for resource utilization in cloud data center using CNN-LSTM model[J]. Neural Computing and Applications, 2021, 33(16): 10043-10055.
- [39] DAYARATHNA M, WEN Y J, FAN R. Data center energy consumption modeling: A survey [J]. IEEE Communications Surveys & Tutorials, 2015, 18(1): 732-794.
- [40] NADJABI C, LOUAHLIA H, LEMASSON S. A review of thermal management and innovative cooling strategies for data center[J]. Sustainable Computing: Informatics and Systems, 2018, 19(5): 14-28.
- [41] LI Z, KANDLIKAR S G. Current status and future trends in data-center cooling technologies [J]. Heat Transfer Engineering, 2015, 36(6): 523-538.
- [42] RAN Y Y, HU H, WEN Y G, et al. Optimizing energy efficiency for data center via parameterized deep reinforcement learning [J]. IEEE Transactions on Services Computing, 2022, 16(2): 1310-1323.
- [43] WANG M, CHANG B, WANG S G, et al. Availability- and traffic-aware placement of parallelized SFC in data center networks[J]. IEEE Transactions on Network and Service Management, 2021, 18(1): 182-194.
- [44] 丁肇豪, 曹雨洁, 张素芳, 等. 能源互联网背景下数据中心与电力系统协同优化 (一): 数据中心能耗模型 [J]. 中国电机工程学报, 2022, 42(9): 3161-3177.
- DIN ZH H, CAO Y J, ZHANG S F, et al. Coordinated operation for data center and power system in the context of energy internet (I): Energy demand management model of data center [J]. Proceedings of the Chinese Society of Electrical Engineering, 2022, 42(9): 3161-3177.
- [45] LIU Y N, WEI X X, XIAO J Y, et al. Energy consumption and emission mitigation prediction based on data center traffic and PUE for global data centers[J]. Global Energy Interconnection, 2020, 3(3): 272-282.
- [46] XIQBAL W, BERRAL J L, ERRADI A, et al. Adaptive prediction models for data center resources utilization estimation [J]. IEEE Transactions on Network and Service Management, 2019, 16(4): 1681-1693.
- [47] DEEPIKA T, PRAKASH P. Power consumption prediction in cloud data center using machine learning [J]. International Journal of Electrical Computer and Engineering, 2020, 10(2): 1524-1532.
- [48] CAO Z W, ZHOU X, HU H, et al. Toward a systematic survey for carbon neutral data centers [J]. IEEE Communications Surveys & Tutorials, 2022, 24(2): 895-936.
- [49] HIEU N T, DI FRANCESCO M, YLÄ-JÄÄSKI A. Virtual machine consolidation with multiple usage prediction for energy-efficient cloud data centers [J]. IEEE Transactions on Services Computing, 2017, 13(1): 186-199.
- [50] FARAHNAKIAN F, PAHIKKALA T, LILJEBERG P, et al. Energy-aware VM consolidation in cloud data centers using utilization prediction model[J]. IEEE Transactions on Cloud Computing, 2016, 7(2): 524-536.
- [51] CHENG Q X, BAHADORI M, GLICK M, et al. Recent advances in optical technologies for data centers: A review[J]. Optica, 2018, 5(11): 1354-1370.
- [52] KANG Y P, HAUSWALD J, GAO C, et al. Neurosurgeon: Collaborative intelligence between the cloud and mobile edge[J]. ACM SIGARCH Computer Architecture News, 2017, 45(1): 615-629.
- [53] SHI Y Y, XU B L, WANG D, et al. Using battery storage for peak shaving and frequency regulation: Joint optimization for superlinear gains[J]. IEEE Transactions on Power Systems, 2017, 33(3): 2882-2894.

- [54] KALYAMPUDI P S L, KRISHNA P V, Kuppani S, et al. A work load prediction strategy for power optimization on cloud based data centre using deep machine learning[J]. *Evolutionary Intelligence*, 2021, 14(2): 519-527.
- [55] 王永真, 宋阔, 韩恺, 等. 算力-电力-热力协同视角下数据中心高密机柜全栈联动研究[J]. *中国科学: 技术科学*, 2025, 55(2): 213-232.
WANG Y ZH, SONG K, HAN K, et al. Full-stack linkage of high-density cabinets in data centers considering computing-electric-thermal power synergy [J]. *Scientia Sinica Technologica*, 2025, 55(2): 213-232.
- [56] HENZEL J, WRÓBEL Ł, FICE M, et al. Energy consumption forecasting for the digital-twin model of the building[J]. *Energies*, 2022, 15(12): 4318.
- [57] ZHANG Q, MENG Z, HONG X, et al. A survey on data center cooling systems: Technology, power consumption modeling and control strategy optimization[J]. *Journal of Systems Architecture*, 2021, 119: 102253.
- [58] CHANG Q K, HUANG Y F, LIU K Y, et al. Optimization control strategies and evaluation metrics of cooling systems in data centers: A review [J]. *Sustainability*, 2024, 16(16): 7222.
- [59] 刘成运, 孟超, 景锐, 等. A 级数据中心综合能源系统多目标优化设计和调度[J]. *电力系统自动化*, 2019, 43(14): 136-142.
LIU CH Y, MENG CH, JING R, et al. Multi-objective optimization design and scheduling of integrated energy system in a-level data center[J]. *Automation of Electric Power Systems*, 2019, 43(14): 136-142.
- [60] YIN X H, YE C J, DING Y, et al. Combined heat and power dispatch against cold waves utilizing responsive internet data centers [J]. *IEEE Transactions on Sustainable Energy*, 2023, 15(2): 819-834.
- [61] LIU Z, ZHANG C, DONG M X, et al. Markov-decision-process-assisted consumer scheduling in a networked smart grid[J]. *IEEE Access*, 2016, 5: 2448-2458.
- [62] HAN O Z, DING T, MU C G, et al. Multi-time scale optimal dispatch for the wind power integrated system with demand response of data centers based on neural network-based model predictive control [J]. *IEEE Transactions on Industry Applications*, 2023, 59(6): 7238-7249.
- [63] LIU Z H, LIU I, LOW S, et al. Pricing data center demand response[J]. *ACM SIGMETRICS Performance Evaluation Review*, 2014, 42(1): 111-123.
- [64] LIMA P A, NETO A S B, MACIEL P. Data centers' services restoration based on the decision-making of distributed agents [J]. *Telecommunication Systems*, 2020, 74(3): 367-378.
- [65] BAJRACHARYA A, KHAN M R A, MICHAEL S, et al. Forecasting data center load using hidden Markov model [C]. 2018 North American Power Symposium (NAPS). IEEE, 2018: 1-5.
- [66] LI Q, GAO D W, ZHANG H G, et al. Consensus-based distributed economic dispatch control method in power systems[J]. *IEEE Transactions on Smart Grid*, 2017, 10(1): 941-954.
- [67] LIN L, GUAN X, PENG Y, et al. Deep reinforcement learning for economic dispatch of virtual power plant in internet of energy[J]. *IEEE Internet of Things Journal*, 2020, 7(7): 6288-6301.
- [68] PIRAGHAJ S F, DASTJERDI A V, CALHEIROS R N, et al. ContainerCloudSim: An environment for modeling and simulation of containers in cloud data centers[J]. *Software: Practice and Experience*, 2017, 47(4): 505-521.
- [69] JING W P, ZHAO C Y, MIAO Q C, et al. QoS-DPSO: QoS-aware task scheduling for cloud computing system[J]. *Journal of Network and Systems Management*, 2021, 29: 5.
- [70] LI D S, LI S W, LAI ZH Q, et al. A memory-efficient hybrid parallel framework for deep neural network training[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2023, 35(4): 577-591.
- [71] BIAN J M, WANG L, REN S L, et al. CAFE: Carbon-aware federated learning in geographically distributed data centers[C]. *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems*, 2024: 347-360.
- [72] JIAN J, ZHAO J L, JI H, et al. Supply restoration of data centers in flexible distribution networks with spatial-temporal regulation [J]. *IEEE Transactions on Smart Grid*, 2023, 15(1): 340-354.
- [73] GHOBADI K, KHAN A A. Using spatial-temporal flexibility of data center building in energy management of distribution grid coupled with multi-energy hubs and energy storage [J]. *Energy Reports*, 2025, 14: 4477-4491.
- [74] ZHANG Y Y, LIANG C, WANG H, et al. A multi-objective interval optimization approach to expansion planning of active distribution system with distributed internet data centers and renewable energy resources[J]. *IET Generation, Transmission & Distribution*, 2024, 18(18): 2999-3016.

作者简介



王志新, 1985、1988、1994 年先后于浙江大学获得学士、硕士和博士学位, 现为上海交通大学研究员/博士生导师, 上海电力大学研究生导师, 主要研究方向风力发电及变流器技术和综合能源系统等。

E-mail: wangzxin@sjtu.edu.cn

Wang Zhixin received his B. Sc. degree, M. Sc. degree, and Ph. D. degree from Zhejiang University in 1985, 1988, and 1994, respectively. He is currently a researcher and Ph. D. supervisor at Shanghai Jiao Tong University, as well as a M. Sc. supervisor at Shanghai University of Electric Power. His main research interests include wind power generation and converter technology and integrated energy systems.



邹怡寒, 2023 年于苏州大学应用技术学院获得学士学位, 现为上海电力大学硕士研究生, 主要研究方向为数据中心、电网协调互动的数据中心功耗预测及其优化调度。

E-mail: 2389075184@qq.com

Zou Yihan received a B. Sc. degree from the Suzhou University of Applied Technology in 2023. He is currently a M. Sc. candidate at Shanghai University of Electric Power. His main research interests include data centers, specifically on predicting data center power consumption and optimizing scheduling in coordination with the power grid.



王简, 2018 年于华北电力大学获得学士学位, 2020 年于上海交通大学获得硕士学位, 2023 年于米兰理工大学获得博士学位, 现为上海交通大学国家电投智慧能源创新学院助理研究员, 主要研究方向为新型电力系统优化控制和电力电子控制。

E-mail: wangjianee@sjtu.edu.cn

Wang Jian received her B. Sc. degree from North China Electric Power University in 2018, M. Sc. degree from Shanghai Jiao Tong University in 2020, and Ph. D. degree from Politecnico di Milano in 2023. She is currently an assistant researcher at the State Power Investment Corporation Smart Energy Innovation Institute of Shanghai Jiao Tong University. Her main research

interests include optimization and control of new power systems and control of power electronic systems.



陈辉, 2006 年于江苏大学获得学士学位, 2009 年于上海大学获得硕士学位, 2014 年于上海大学获得博士学位, 现为上海电力大学教授, 主要研究方向为机器视觉与模式识别、机器人导航与地图构建 SLAM 和电力设备状态检测。

E-mail: chenhui@shiep.edu.cn

Chen Hui received her B. Sc. degree from Jiangsu University in 2006, M. Sc. degree from Shanghai University in 2009, and Ph. D. degree from Shanghai University in 2014. She is currently a professor at Shanghai University of Electric Power. Her main research interests include machine vision and pattern recognition, robot navigation and SLAM for map construction and power equipment condition monitoring.



钱进, 2023 年于南京师范大学中北学院获得学士学位, 现为上海电力大学硕士研究生, 主要研究方向数据中心能源配置规划。

E-mail: 2277348812@qq.com

Qian Jin received a B. Sc. degree from Zhongbei College of Nanjing Normal University in 2023. He is currently a M. Sc. candidate at Shanghai University of Electric Power. His main research interests include focusing on energy configuration planning for data center.



龚春阳(通信作者), 2013 年于上海电力大学获得学士学位, 2014 年于英国斯特拉斯克莱德大学获得硕士学位, 现为上海电力大学助理研究员, 主要研究方向为新能源发电规划与控制及互联多微网系统。

E-mail: sh19900307@163.com

Gong Chunyang (Corresponding author) received his B. Sc. degree from Shanghai University of Electric Power in 2013 and M. Sc. degree from the University of Strathclyde, UK, in 2014. He is currently an assistant researcher at Shanghai University of Electric Power. His main research interests include new energy power generation planning and control and interconnected multi-microgrid systems.