

人工智能的气候影响悖论：文献述评与研究展望

陈占明¹，熊齐扬¹，胡姿彤^{1,2}，刘诗韵¹，马健洪^{1,*}

1. 中国人民大学 应用经济学院，北京 100872；

2. 中国地质大学（武汉） 经济管理学院，武汉 430078



摘要： 在全球气候治理的背景下，人工智能（Artificial Intelligence, AI）兼具高能耗基础设施与低碳转型关键工具的双重属性，构成了复杂的气候影响悖论。本文系统梳理并评述了国内外前沿文献，旨在厘清AI与能源消耗和碳排放之间的复杂关系。研究发现，算力基础设施的指数级扩张不仅加剧了运营能耗对区域电网的冲击，而且其硬件制造与电子废物处理环节的隐性碳排放正演变为不可忽视的环境负担。尽管AI在优化能源系统、驱动绿色技术创新及重塑供应链方面展现出显著的正外部性，但技术效率提升导致的物质、经济与社会行为层面的回弹效应，往往会削弱甚至抵消其减排收益，引发杰文斯悖论。因此，本文认为单纯依赖外生技术进步不足以规避AI的环境风险，必须构建涵盖全生命周期碳足迹核算、软硬件协同创新与市场化激励机制的复合治理框架。该框架旨在实现算力与绿电的深度融合，使AI真正成为实现可持续发展目标的关键推动力。

关键词： 人工智能；数据中心；碳足迹；回弹效应；低碳转型

中图分类号： X196 **文献标志码：** A **文章编号：** 2097-4981(2026)02-0315-16

DOI: 10.3724/j.issn.2097-4981.JECC-2025-0385

0 引言

人工智能（Artificial Intelligence, AI）作为通用目的技术（General Purpose Technology, GPT），正深刻重塑全球经济版图与气候治理格局。然而，技术进步具有鲜明的“双刃剑”属性^[1]，即AI既是高度依赖电力的高能耗基础设施，又是驱动资源优化和低碳转型的关键工具。这种内在冲突构成了复杂的气候影响悖论^[2-3]。一方面，受“缩放定律”驱动，算力需求的指数级增长导致直接能耗飙升，且其增速往往大于技术落地带来的能效优化；另一方面，AI赋能减排收益具有滞后性，且极易被效率提升诱发的回弹效应抵消。因此，如何平衡算力扩张与气

候目标，已成为关乎全球可持续发展的重要命题。

这一矛盾在中国和美国两个主要的AI市场中尤为突出。当前，全球AI发展高度集中。国际能源署（International Energy Agency, IEA）数据^[4]显示，2024年全球数据中心耗电量约为415 TWh，约占全球总用电量的1.5%。其中，美国（45%）位居榜首，其次分别为中国（25%）和欧洲（15%）。斯坦福大学发布的《2025年人工智能指数报告》^[5]指出，美国主导了全球的数字基础设施建设，2024年控制了约74%的高端智能算力，AI领域的私人投资达数千亿美元。然而，算力霸权也伴随着巨大的能源成本。2024年，美国数据中心能耗高达180~200 TWh^[4,6]，占全美总用电量的4%，并由此产生了约61 Mt的碳排放量。作

收稿日期： 2025-12-19；**接受日期：** 2026-02-03

基金项目： 中国人民大学“求是学术”育人育才项目（RUC25QSDL059）；中国人民大学科学研究基金（中央高校基本科研业务费专项资金、国家治理大数据和人工智能创新平台经费资助）项目（202530251）

作者简介： 陈占明（1984—），男，博士，教授，博士生导师，主要研究方向为能源经济与政策。E-mail: chenchanming@ruc.edu.cn

通讯作者： 马健洪（1996—），男，博士研究生，主要研究方向为生产网络一般均衡模型。E-mail: majianhong@ruc.edu.cn

引用格式： 陈占明，熊齐扬，胡姿彤，等. 人工智能的气候影响悖论：文献述评与研究展望[J]. 能源与气候变化, 2026, 2(2): 315–330.

CHEN Z M, XIONG Q Y, HU Z T, et al. The Climate Impact Paradox of Artificial Intelligence: Literature Review and Research Prospects [J]. Journal of Energy and Climate Change, 2026, 2(2): 315–330.

为AI市场的另一增长极，中国在数字中国建设与“人工智能+”行动驱动下，正经历算力、算法与数据要素的爆发式增长。2024年，中国数字经济核心产业的增加值约占国内生产总值的10%^[7]。2024年，中国数据中心的用电量约为166 TWh，占全社会用电量的1.68%，碳排放量达85.9 Mt^[8]。2025年上半年，中国智能算力规模达788百亿亿次/秒^[9]。更为严峻的是，相关研究预测中国AI碳足迹的碳达峰时间可能推迟至2038年，滞后于总体达峰目标^[10]。

面对气候挑战，全球已初步构建起AI治理框架。联合国教科文组织^[11]与欧盟率先强调了全生命周期评估AI能耗的必要性，并制定了数据中心可再生能源使用目标。美国依据《2020年能源法》强化了数据中心资源使用的监管和报告机制。德国通过《能源效率法》要求自2026年7月起，新建设施的电能利用效率（Power Usage Effectiveness, PUE）须低于1.2。中国立足于碳达峰碳中和（“双碳”）目标与数字经济发展的客观要求，构建了供需协同的政策体系。在供给侧，中国通过《数据中心绿色低碳发展专项行动计划》等政策文件，确立了新建枢纽节点绿电占比超过80%的要求，并积极探索“算电协同”的市场激励机制；在需求侧，中国在《关于推进“人工智能+”能源高质量发展的实施意见》和《中华人民共和国国民经济和社会发展第十五个五年规划纲要》中，分别提出AI赋能新型电力系统和千行百业战略路径。尽管政策体系不断完善，但在“适度超前建设新型基础设施”与“双碳”目标限制的背景下，数字基础设施的能耗扩张与碳减排约束之间仍存在不确定性，构成了亟待破解的气候影响悖论。

当前，大量研究探讨了AI气候影响，并在理论框架与实证分析方面取得了较为丰富的成果。为厘清AI与能源消耗和碳排放之间的复杂关系，本文通过系统梳理与评述国内外文献，深入探究AI气候影响的作用机制，重点回答3个基本问题：1）随着数字基础设施的快速发展，能源消耗与碳排放面临哪些挑战？2）AI助力中国构建新型能源体系，推动重点工业领域节能降碳有哪些赋能路径，节能收益在何种程度上被效率提升引发的回弹效应所抵消？3）如何最大限度地发挥AI在减排方面的净效益，从而为实现中国式现代化与“双碳”目标作出贡献？

本文的研究框架如图1所示：第1节评估AI发展带来的能源消耗和碳排放成本，第2节介绍AI在推动低碳转型方面的赋能机制，第3节分析AI减排效果面临的回弹效应，第4节探索发展可持续AI的可行路径，第5节总结研究结论并提出未来展望。

1 AI的能源消耗与碳排放成本

1.1 AI基础设施的运营能耗

1.1.1 算力规模扩张与电力消耗

数据中心作为AI发展的底层物理支柱，其能源消耗的历史演变呈现出明显的阶段性特征。第一波建设浪潮源于20世纪末互联网的普及，2000—2005年美国数据中心的电力消耗量增长约一倍^[12]。然而，在2007—2008年后，全球数据中心能耗增速放缓，甚至与持续增长的互联网服务需求出现了“脱钩”现象^[13]。这一变化主要归因于两大因素：一是规模效应，计算负载向能效更高的超大规模数据中心聚集；二是技术效应，服务器硬件能效得以改善，运营管理得到优化。二者共同降低了单位计算任务的能耗^[14]。2017年后，能效改善带来的平衡被打破，数据中心能耗迈入新一轮急剧增长期^[15]。其核心驱动力为云计算的成熟、流媒体的渗透及AI技术的商业化落地。2015—2024年，专门用于并行处理的加速服务器容量增速达到传统服务器的4倍。此类硬件处理特定任务的能效更高。然而，得益于其强大的算力，许多受限于计算瓶颈的大规模AI训练等新应用得以实现。这导致服务需求总体增长，从而抵消了硬件能效的提升^[4]。

图2为全球数据中心用电量的预测结果。由图可知，数据中心的用电量存在高度不均衡态势，重点集中于主要经济体。IEA数据显示，美国、中国及欧洲的数据中心用电量之和约占全球数据中心总用电量的85%^[16]。IEA预测显示，随着AI训练和推理需求的不断增长，全球数据中心电力需求有望在2026年较2022年水平实现翻倍^[17]；此外，至2030年，全球AI服务器的能源消耗预计将较2024年水平增长5倍^[18]。中美两国将主导全球约80%的能耗增量^[4]。值得注意的是，美国作为全球算力基础设施最为密集的国家，其境内数据中心的能源负载强度远超全球

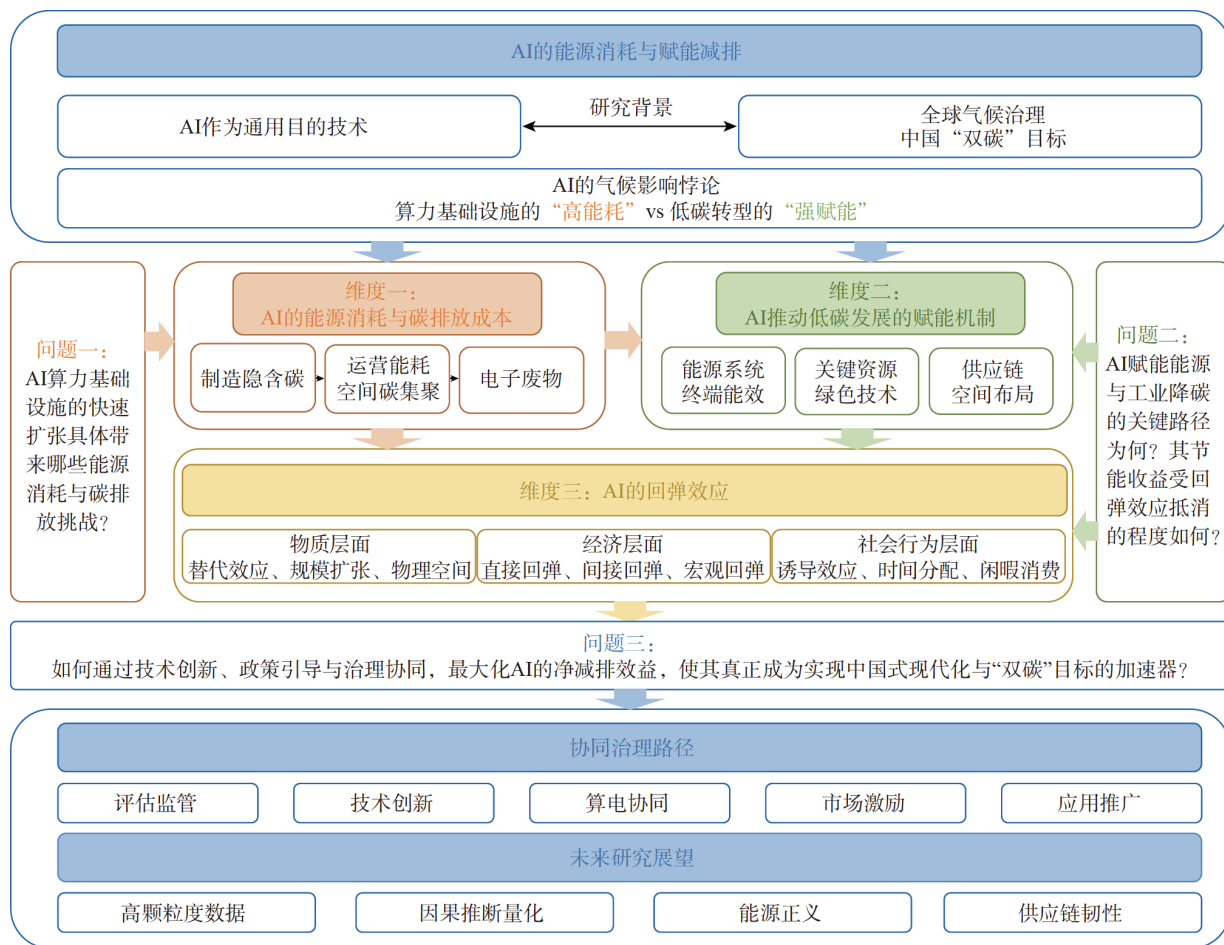


图1 研究框架

Fig.1 Research framework

平均水平。2024年，分摊至美国本土居民的平均数据中心电力负荷约为540 kWh，预计到2030年将突破1 200 kWh。相比之下，当前中国人均数据中心电力负荷仅为70 kWh，但增长势头强劲，预计到2030年将攀升至200 kWh。若缺乏有效的脱碳措施，大规模的能耗扩张将会给地球边界带来严峻考验。研究表明，当前全球数字内容消费形式约占用了1.5℃温控目标下人均碳预算的40%^[19]。若不采取更为积极的脱碳举措，日益增长的AI需求将进一步加重全球气候压力。

除总量增长带来的压力外，数据中心的空间集聚特性对局部电网构成了严重的非对称冲击。研究表明，由于网络延迟敏感性高、基础设施依赖性强及产业集群效应的影响，数据中心表现出空间集聚趋势^[20]。这种高度集聚会导致局部电力负荷激增，

且由于流量与能耗在时空上的错配，容易引发“碳效率陷阱”^[21]。当前超大规模数据中心的单体功耗已达数百兆瓦且呈上升趋势，这直接引发了区域性的电力短缺危机^[22]。2022年，爱尔兰数据中心的电力消耗量较2015年增长了3倍以上，占该国全社会总用电量的18%，且预计到2031年这一比例将攀升至28%。2020年，新加坡数据中心用电量约占全社会总用电量的7%，这使得新加坡政府暂时停止了新建项目，并开始实行更严格的能耗和用地准入标准^[23]。IEA分析报告进一步表明，超过50%的规划或在建大型数据中心选址靠近大城市，这一布局导致并网排队周期延长至7~10年^[18]。这种集中建设既加剧了区域电网拥堵，又加大了变压器、电缆等关键电力设备的供应链压力，从而形成了限制能源稳定供应的局部瓶颈^[18]。

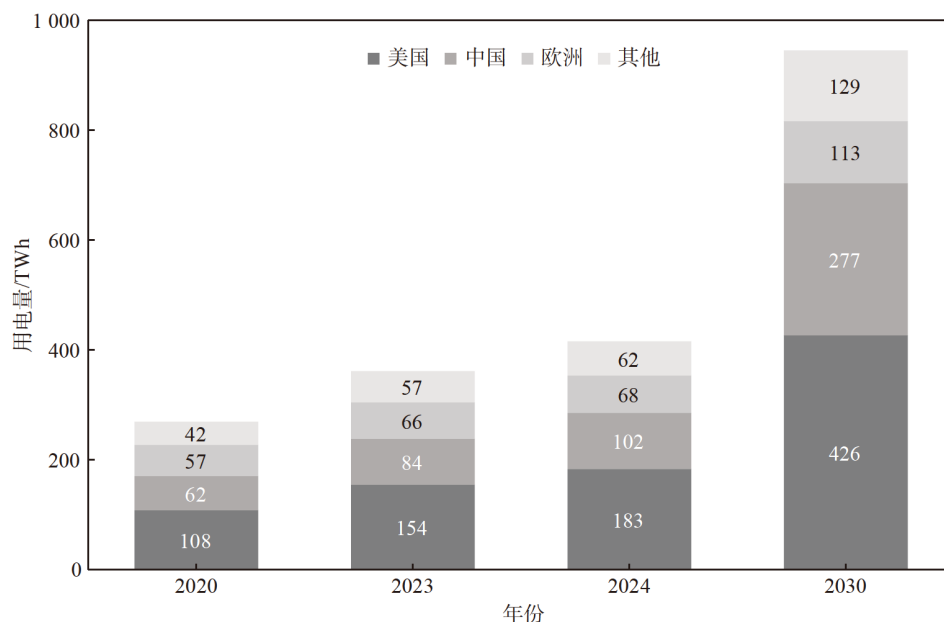


图2 全球数据中心用电量预测结果

Fig.2 Forecast results of global data center electricity consumption

注：数据来源为IEA。

能耗的大幅增长会直接引发温室气体排放量的显著上升。当前，数据中心所产生的碳排放量占全球碳排放总量的1.4%，约25%的数据处理负载和各类AI应用密切相关^[24]。随着大语言模型的爆发式兴起，该比例的增长速度不断加快。谷歌^[25]、百度^[26]、微软^[27]等互联网巨头的年度可持续发展报告显示，2020—2024年AI导致的电力消耗和碳排放量明显提高。然而，企业通过购买可再生能源证书进行碳抵消可能会掩盖真实的碳排放量；若采用基于位置的核算方法报告实时的电力消耗，部分科技企业的实际碳排放量会显著高于报告值^[28]。

1.1.2 AI模型训练与推理的能源需求

AI训练是通过处理海量数据来识别内在规律的过程，具有耗时长、耗能高的特征^[29]。近年来，前沿AI模型为追求卓越性能，其训练数据量呈指数级增长，而最大规模的AI训练任务所需的计算量每隔6~10个月就会翻倍^[30]，造成能耗和碳排放量的增长。此类高强度任务大多依靠高性能图形处理器（Graphics Processing Unit, GPU）集群完成。例如，GPT-4的训练总功率高达22 MW，相当于150个同时运行的高功率电动汽车充电站^[31]，日均能耗相当于发达经济体28 500个家庭的日均用电量^[6]。值得注意

的是，训练能耗存在规模效应，与模型参数量、数据规模及硬件效率紧密相关^[32]。此外，模型训练并非一次性完成，通常需进行预训练后的微调。虽然该阶段计算强度较低，但仍会带来额外能耗^[33-34]。表1为典型AI模型训练阶段的能耗与碳排放量测算数据。

AI推理是训练完成的模型对新输入的数据进行计算并输出预测结果的过程。尽管单次推理任务的能耗远低于训练阶段的总能耗，但凭借庞大的用户基数与高频次调用需求，AI推理的累积能耗正日益凸显^[38]。DE VRIES^[41]研究指出，在AI技术的加持下，互联网服务的能耗快速增长，一次常规搜索的能耗为0.3 Wh，ChatGPT的单次响应能耗约为2.96 Wh，而AI驱动搜索能耗高达8.9 Wh。数据显示，在AI全生命周期中，约60%~90%的能耗集中于推理阶段^[37,42-43]。尽管不同来源的估算口径存在差异，但AI推理阶段的累积能耗正逐渐超越训练阶段已成为行业共识^[24]。推理阶段的能耗强度主要受模型架构、任务类型、推理时长等因素影响。一个模型的参数越多，通常意味着其算力资源消耗越多。然而，采用混合专家等模型架构可在模型参数一定的情况下有效提升能效^[44]。任务类型的异质性也会导致能耗存在差异。例如，生成图像或视频的能耗强度远高

表1 典型AI模型训练阶段的能耗与碳排放量测算数据

Table 1 Estimated energy consumption and carbon emissions for typical AI model training

模型	发布年份	能耗/MWh	碳排放量/t	数据来源
BERT	2018	1.5	0.7	STRUBELL 等 ^[33]
GPT-2	2019	1.7*	0.7*	STRUBELL 等 ^[33]
GPT-3	2020	1 287.0	552.0	PATTERSON 等 ^[32]
LaMDA	2021	451.0	26.0	THOPPILAN 等 ^[35]
Gopher	2021	1 151.0*	380.0	RAE 等 ^[36]
GLaM	2021	456.0	40.0	PATTERSON 等 ^[37]
BLOOM	2022	433.0	24.7	LUCCIONI 等 ^[38]
PaLM	2022	3 436.0*	271.4	CHOWDHERY 等 ^[39]
GPT-4	2023	42 400.0*	—	IEA ^[4]
DeepSeek-V3	2024	2 341.0*	—	DeepSeek-AI 等 ^[40]

注：*表示基于硬件参数与训练时长推算的预估值。

于文本任务。同时，提高推理时长会使AI模型延长思考时间，提升输出质量，但会导致推理能耗成倍增加^[45]。

1.2 AI设备的全生命周期碳排放

1.2.1 制造阶段的隐含碳排放

虽然运营阶段的碳排放通常在AI设备全生命周期中占据主导地位，但制造阶段累积的隐含碳排放正逐渐成为不可忽视的环境负担^[46]。研究显示，一台典型的高性能服务器在制造阶段的能耗已超过10 MWh，相当于其5年运营总能耗的12.5%^[4]。就芯片而言，制造阶段能耗约占全生命周期总能耗的20%^[47-48]。随着制程工艺向微缩化发展，制造阶段的能耗强度快速增加。以前沿的3 nm芯片为例，单片制造约消耗2.3 MWh电力^[49]。其中，60%用于沉积、光刻和刻蚀等核心工序，剩余40%用于冷却系统等辅助工序^[49]。这表明能源强度的增加与芯片复杂度呈正相关。新一代芯片通过集成更复杂的金属层与晶体管结构来提升算力。虽然单位计算任务的运营能耗可能会因效率提高而降低，但制造阶段的隐含碳排放量却会进一步升高^[50]。由于制造环节主要依赖电力驱动，其碳足迹主要受生产所在地的能源结构影响。目前，全球超70%的先进制程产能集中于依赖煤炭等化石能源的东亚地区^[51]。这使得芯片制造阶段电力消耗导致的碳排放量居高不下。相比之下，制造过程中的直接燃烧排放贡献较弱，占比通常仅

为碳足迹的5%^[15]。

1.2.2 电子废物的回收挑战

生成式AI的爆发式增长导致对高性能GPU的需求快速上升，压缩了数字基础设施的使用年限。WANG等^[52]研究指出，随着模型计算需求每数月翻倍，服务器的更替周期缩短。在快速的AI发展情景下，硬件寿命可能从传统的3~5年大幅缩短至1.5年左右，这将导致大量的服务器、冷却系统及配套设施提前退役。此外，芯片出口限制等地缘政治风险增强了企业囤积先进芯片、提前退役现有机架的动机，进一步加剧了电子废物的问题。WANG等^[52]预测，若不采取循环经济措施，2020—2030年全球AI产业将累计产生1.2~5.0 Mt的电子废物。作为全球电子消费大国，中国到2030年的电子产品废弃总量预计将增至27.22 Mt，年均增长率达10.4%^[53]。

2 AI推动低碳发展的赋能机制

2.1 电力系统的运行优化与终端能效提升

在电力系统运营方面，AI在需求侧管理和电网运行优化方面展现出巨大的节能潜力。ROLNICK等^[54]指出，机器学习可以通过精准预测电力供需波动，加强电网对新能源的消纳能力。同时，AI提高了天气预报的准确性，助力电网运营商更精准地预估太阳能电池板和风力涡轮机的输出情况，有效平抑了可再生能源的间歇性波动^[55-56]。周剑等^[57]发现，

基于深度强化学习的算法能够有效涵盖碳成本约束，降低微电网系统的综合运行成本。面对日益复杂的电力系统形态，XIE等^[58]提出了“技术-市场-政策”三元架构，论证了AI算法在处理高维决策与不确定性问题方面的优势，可有效加速电网的碳中和转型进程。

在终端用能侧，AI塑造的数据驱动决策体系正推动能源管理从“粗放管理”向“智能调控”转变。屠西伟和史丹^[59]认为，数字技术可实现对能源消耗全过程的实时监测，降低冗余环节造成的能源损耗。这一优化能力在建筑领域的应用效果尤为显著。DING等^[60]针对美国中型办公楼开发的一套评估方法显示，利用AI对设备控制进行优化可使到2050年商业建筑的能耗和碳排放分别降低约40%和90%，大幅降低了实现净零建筑的成本溢价。

此外，AI技术也被全面应用于数字基础设施自身的能效优化。针对5G网络中流量与能耗的时空错配，LI等^[21]提出了“DeepEnergy”深度强化学习方案，通过动态调控基站工作状态，能够使中国超71%的省级行政区避免“碳效率陷阱”，在全国层面的理论减排潜力超20 Mt。在数据中心运维方面，杨震等^[61]通过实验证实了神经网络模型在优化PUE方面的显著成效。此外，LAL和YOU^[62]提出了跨系统融合模式，建议将AI数据中心运营与加密货币挖矿进行协同，预计到2030年可减少700 Mt的碳排放量，为AI的绿色发展提供了新思路。

2.2 关键资源的供应保障与绿色技术创新

在能源供给侧，AI通过强化可预测性与调度灵活性，打破了能源产业对化石能源的路径依赖^[63]。ZHANG等^[64]利用多源地理空间数据及机器学习模型，量化了中国城市屋顶光伏高达3 000~4 000 Mt的减排潜力，为分布式能源布局提供了数据支撑。在上游关键矿产领域，VESPIGNANI和SMYTH^[65]指出，AI技术通过削减采矿项目的“后端风险溢价”，增强了锂、钴等资源的供应链韧性，从而有效降低了能源转型成本。

在绿色创新领域，AI正在重塑研发范式。例如，AI可以帮助研究人员从海量候选库中筛选出高性能的电池材料^[66-67]。在实证层面，WANG等^[68]、鲁锦

涛等^[69]分别从宏观经济、区域创新扩散视角证实了AI技术对低碳技术进步的正向驱动机制。VERENDEL^[70]基于专利引用的分析结果进一步揭示，AI已深度嵌入交通、能源和工业领域的绿色发明，这表明其已成为推动全球低碳技术创新与扩散的核心驱动力。

同时，AI能有效提升环境气候的监测能力。例如，EYRING等^[71]证实了机器学习在构建高保真地球系统模型和预测极端气候风险领域的价值。目前，AI结合遥感技术已被广泛应用于识别森林砍伐^[72]、监测甲烷泄漏^[73]及评估珊瑚礁健康状况^[74]等实际场景。

2.3 供应链管理与空间布局优化

在供应链管理层面，AI可以通过缓解信息不对称和优化资源配置来提高全球价值链的运行效率。钞小静和沈路^[75]的研究表明，焦点企业的AI技术进步对下游具备显著的绿色溢出效应，能有效规避“长鞭效应”导致的资源错配。HONG和XIAO^[76]进一步指出，AI与区块链技术的结合能降低供应链中的信息壁垒，在碳核算及循环经济方面发挥关键作用。

在空间布局层面，数字基础设施的科学规划能够从根源上减少AI的碳足迹。XIE等^[77]以中国“东数西算”工程为例，发现通过跨区域的算力调度，将能耗负载引导至可再生能源丰富的西部地区，在2020—2050年将实现2 125 Mt的碳减排量。

3 AI的回弹效应

在量化了AI的直接环境成本和其作为减排工具的潜在环境收益后，评估其净气候影响仍存在巨大挑战。若仅关注直接能耗与减排收益的线性叠加，往往会忽视效率提升引发的系统性行为和结构变化，即回弹效应。本节将深入剖析这一核心悖论，分析其在物质、经济及社会行为层面的多维传导机制。

3.1 理论框架与学术争议：AI时代的杰文斯悖论

回弹效应，亦称杰文斯悖论，指技术进步在提升效率、降低单位服务成本的同时，可能刺激更大规模的使用需求，最终导致资源消耗总量不减反

增^[78]。自19世纪经济学家杰文斯^[79]首度揭示这一现象后,其已在能源、交通等领域得到验证^[80-83]。针对AI领域的气候影响前景,学术界存在明显分歧。作为技术乐观派,PATTERSON等^[37]认为,硬件与算法的持续创新将大幅提升AI能效,使碳足迹趋于平稳甚至下降。然而,更多学者揭示了“数字经济悖论”的存在。WANG等^[68]指出,尽管数字技术能通过低碳创新遏制碳排放,但数字产业化主导地位的增强往往会抵消这一减排收益,导致碳排放总量不降反增。史丹等^[84]进一步提出了数智化的“能效悖论”:尽管在微观层面数智化通常能改善能源效率,但在宏观层面却难以观察到同等幅度的效率提升。其根源在于数智化本身并非被直接用于提高能源效率,而是通过深度嵌入经济活动,重塑生产消费方式,从而以更间接、更复杂的路径影响能源投入和产出效率。因此,数智化带来的技术进步并不一定会导致能源消费的减少。

3.2 AI回弹效应的类别

AI在物质层面的回弹效应本质是数字替代的减量和规模扩张的增量相互作用的结果。一方面,替代效应体现为数字化服务对传统实体产品或流程的取代,包括数字导航代替纸质地图、生成式内容取代部分传统出版物^[23]。研究表明,相较于人类创作,AI生成文本和图像的碳排放量低3~4个数量级^[85]。尽管此类替代在理论上存在“去物质化”的环保潜力,但其气候净影响取决于数字基础设施能耗与实体产品碳足迹的对比结果。另一方面,规模效应表现为当前AI模型的发展对参数与算力的无尽追求。受“缩放定律”的影响,模型能效的边际改善往往被模型参数的指数级膨胀所吞噬^[86]。CHEN等^[10]发现,2017—2024年中国智能算力规模的年均增速超过100%。宋德勇等^[34]进一步分析指出,智能化转型通过规模扩张和要素替代使企业碳排放量呈现先增后减的倒U形特征。

此外,AI物质回弹效应在地理上还表现为“空间回弹”,即在终端设备日益小型化的同时,用于AI训练与推理的数据中心却在地理空间上高度集聚。典型案例为美国弗吉尼亚州的数据中心走廊^[87]。这种空间集聚虽提升了聚集区内部的数据传输效率,

却对区域电网的稳定性带来了巨大压力^[88]。

在经济层面,AI带来的效率提升经由价格信号传导,诱发了多层次的回弹效应。直接经济回弹效应是杰文斯悖论的典型表现。GPU算力成本的下降刺激了算力需求和硬件出货量的上涨,导致能耗总量攀升。MASANET等^[13]指出,过去数据中心能效的提升曾一度抵消了需求增长所带来的能耗压力。但面对生成式AI的冲击,单纯的技术优化已触及瓶颈^[89]。间接经济回弹效应则源于收入效应,即AI带来的成本下降会提高消费者的实际收入,转而消费其他具有高能耗属性的产品或服务^[90]。技术驱动型淘汰也是产生间接经济回弹效应的诱因。由于智能家居与终端设备逐渐依赖云端AI能力,消费者为获取新功能而提高硬件更新频率,进而增加了设备制造与电子废物处理环节的资源消耗^[91]。宏观经济回弹效应进一步体现了AI作为GPT对经济结构的整体影响。科技巨头大规模签署可再生能源采购协议虽客观拉动了绿电投资,但也深刻改变了能源市场的供需格局^[92]。

在社会行为层面,AI通过重塑个体认知与组织行为模式,产生了更为深远的回弹效应。首先是诱导效应对消费行为的异化。石峰等^[93]提到,大数据驱动的推荐系统改变了营销模式及企业与消费者之间的互动逻辑,从而为平台创造了巨额商业价值^[94]。然而,诱导效应也助推了非必要消费的增长和“快时尚”文化的蔓延,进而增加了产品全生命周期的环境足迹^[95-96]。

其次是时间回弹效应针对被AI释放的时间资源的再分配问题。虽然智能导航和家务机器人为用户节省了通勤与劳动时间,但研究显示,这些被节省的时间往往被重新分配至网络购物、长途旅行等高碳休闲活动^[97]。考虑到全球闲暇消费导致的碳排放量占碳排放总量的18%,且高收入群体与低收入群体间存在巨大的碳足迹鸿沟^[98],若技术效率提升带来的时间红利转化为高能耗的休闲行为,将严重削弱甚至抵消AI的减排收益。

相较于传统自动化技术,AI还通过降低使用门槛和“注意力捕获”机制改变了用户的使用行为。在生产侧,生成式AI极大降低了内容创作的边际成本,但反而诱发了用户高频的探索性生成操作。这

种低门槛使用户停留在数字界面的时间不降反增，直接推高了模型推理阶段的累积能耗^[99]。在消费侧，深度学习的推荐算法将用户碎片化时间变现为短视频流等活动，导致时间被高能耗的数字消费所填补^[19]。尽管智能导航等单一应用可降低3.4%的碳排放量^[100]，但AI导致的社会行为改变会造成更高的总资源消耗^[101]。因此，TOMLINSON等^[85]指出，对AI气候影响的判断需要全面考虑职业替代、法律风险及社会回弹效应。

4 迈向绿色可持续AI的协同治理路径

前文分析表明，AI虽然具备推动低碳转型的赋能潜力，但伴随其全生命周期的高碳排放及效率提升导致的系统性回弹效应成为实现净环境效益的主要障碍。要破解这一气候影响悖论，依赖单一维度的技术效率提升已无法应对算力需求的指数级增长。在向可持续AI迈进过程中，供给侧与需求侧均面临深层的结构性难题。在供给侧，头部科技企业对模型参数与硬件细节的“黑箱化”给能耗评估带来了极大的困难；而且由于模型架构与硬件效率之间的非线性相互作用，单一调控难以取得理想效果^[43]。在需求侧，充分释放AI的节能潜力需克服高质量数据集匮乏、数字基础设施薄弱及隐私监管不确定性等多重障碍^[16]。为应对这些难点，本节提出构建“技术优化”和“市场约束”双轮驱动的协同治理框架。该框架包含全生命周期评估、软硬件协同创新、源荷互动调度及市场经济激励等多层级途径，既能抑制能源回弹效应，又能同时放大AI助力经济低碳转型的赋能效应，推动AI走向绿色可持续的发展道路。

4.1 评估与监管框架的构建

构建透明、统一的评估体系与监管框架是打破“数据黑箱”、推动AI绿色发展的首要前提。在评估指标与标准化方面，虽然CodeCarbon、MLCO₂等碳核算工具已投入应用^[102-103]，但由于统一评估边界与计算方法的缺乏，不同工具对同一AI模型的评估结果通常存在差异^[104]。因此，标准组织亟需建立一套纳入运营碳排放和隐含碳排放的全生命周期标准化

核算框架^[47]。此外，KAACK等^[24]强调，AI环境足迹的评估体系还应纳入回弹效应对电力系统产生的综合影响。

在数据透明度建设方面，建立开放的数据生态是关键。目前，主要的数据中心运营商常以商业机密为由拒绝披露能耗细节，数据的不透明性严重阻碍了行业的绿色转型^[5]。对此，政府应出台相关政策，鼓励企业披露AI模型的能耗数据，帮助用户在选择AI服务时做出基于实际需求的理性决策^[13]。虽然联合国教科文组织的《人工智能伦理建议书》和欧盟的《人工智能法案》已开始关注AI开发过程的透明度，但目前仍缺乏强制性的能耗报告规定^[11,105]。未来政策应平衡发展与监管，既要通过强制披露或激励措施增加信息透明度，又要避免采取过多的限制手段，阻碍AI的性能提升和发展潜力。要实现这一目标，AI开发者应披露模型架构和训练硬件类型等关键信息；数据中心运营商需报告本地电网的实时碳强度和能源结构；研究机构与标准组织则应制定统一的评估边界并提供开放共享的数据集^[106]。

4.2 软硬件协同的能效提升

摩尔定律逐步放缓使得物理极限显现，依靠半导体工艺微缩获得的性能红利已难以为继。未来的能效增长重点在于计算堆栈“顶层”的创新，即软件算法和硬件架构的深度协同改良^[107]。

从软件算法层面看，核心在于降低计算复杂度以实现能效突破^[29]。大语言模型高强度训练时，流水线并行^[108]、张量并行^[109]、内存优化^[110]及低精度计算^[111]等技术明显缩减了训练周期并降低了能耗。在推理阶段，混合专家等稀疏化架构只会激活部分网络参数，以在保持模型容量的同时大幅削减单次推理的计算量^[112]。剪枝、量化与知识蒸馏等模型压缩技术则切实减少了部署成本^[113]。动态计算图和自适应推理技术使模型能够根据输入的复杂度灵活调配资源，避免了简单任务的“过度计算”^[114]。与此同时，软件平台的工程化优化有着巨大潜力。Facebook经由预计算和缓存机制在翻译任务中实现了6.7倍的能效提升^[115]，并且通过简化网络结构与超参数调优迭代直接降低了能耗^[116]。

在硬件架构层面，开发高能效的专用AI芯片是

突破算力瓶颈的关键。专用芯片通过摒弃通用处理器中的冗余功能模块,将晶体管资源集中到并行计算单元,在深度学习等特定领域实现了2~3个数量级的能效飞跃^[107]。相较于中央处理器(Central Processing Unit, CPU), GPU大幅提升了效率;而在特定任务中,张量处理器能效高达当前CPU或GPU的30~80倍^[117]。英伟达公司的Blackwell等新一代硬件架构,不仅能够支持万亿参数模型的训练工作,更在推理任务中实现了能效的跨越式提升^[118]。相较于初代的英伟达芯片,Blackwell在训练过程中的能耗降至原来的1/350,在推理过程中的能耗更是降至原来的1/45 000^[119]。未来,光子集成电路与类脑计算等颠覆性范式会重新定义计算能效的物理极限。

4.3 绿色能源的源荷互动与时空调节

保障AI电力供应稳定性的重点在于形成“源荷互动”的精准协同机制,即在时空维度达成计算负荷和低碳能源的最佳匹配,从而完成从单纯“节能”向深度“绿能”的范式转变。

在能源结构选择方面,核能具备零碳排放与高稳定性的特征,成为支持AI算力扩张的理想能源。AI模型训练对电力供应稳定性的要求极高,这与核能发电的特性高度契合,有利于使算力摆脱对化石能源的依赖。在实践层面,BLOOM模型凭借法国核电有效缩减了自身的全生命周期碳足迹^[120];亚马逊位于宾夕法尼亚州的960 MW核能设施也为数据中心提供长期的低碳电力保障^[121]。这些科技巨头不仅将核能视为替代化石燃料的关键,积极推动现有核设施的重新启用^[122],同时探索小型模块化反应堆等新一代核能技术的可行性^[123]。

从空间策略角度看,核心逻辑为将高能耗的算力基础设施向风电、光伏等可再生能源富集区域迁移。由于电力传输过程存在拥堵瓶颈及线损问题,新建超大规模数据中心呈现出明显的“逐绿而居”趋势。例如,中国“东数西算”工程将算力枢纽建立在可再生能源充裕的西部地区^[124],以及利用北极地区风能与自然冷却条件建设数据中心^[125]。在实证层面,阿里巴巴与华北电力大学合作的“算力-电力”优化调度项目成功将部分算力负载从江苏南通调往河北张北,实现了数据中心与电力系统以促进

可再生能源消纳为目标的跨区域协同,在国内该领域具有标志性意义^[126]。

在时间策略层面,通过灵活调整计算任务的时间窗口并利用储能技术,可实现对波动性清洁电力的主动消纳。针对训练任务,通过将其调度至电力负荷低谷期或可再生能源发电高峰期,可达成本节约和碳减排的双重收益^[127-128]。具体技术路径包括延迟任务调度、增强工作负载的时间灵活性及按照实时碳强度选择能源来源^[129-131]。针对难以预测的推理任务,则需依靠专为系统冗余设计的储能系统,通过在发电高峰期存储绿电以应对用电低谷期的需求。构建绿色电力的需求响应机制,能够将传统“能耗大户”的数据中心转化为支撑新型电力系统稳定运行的“柔性资源”,实现AI模型运行成本与碳排放的协同优化,进而为构建可持续的数字基础设施提供系统性解决方案^[132-134]。

4.4 基于市场激励的经济约束机制构建

针对技术效率提升引发的能源回弹效应,仅依靠供给侧的技术改良难以从根本上遏制算力需求的无序扩张。根据环境经济学经典的外部性理论^[135],需尽快引入依托于市场的经济手段,利用价格信号将AI全生命周期的影响内部化,在需求侧形成有效成本约束,同供给侧的技术治理形成互补^[136]。

建立覆盖算力基础设施的碳定价及能耗配额交易体系。朱子恒等^[137]提出,将数据中心纳入全国碳排放权交易市场可以引导数据中心提升能效并采用清洁电力。通过设定严格的碳排放总量上限及各级递减的免费配额,企业将主动权衡碳配额购买与低碳技术投资之间的成本收益,从而有效对冲单位算力成本下降带来的能源回弹风险^[138]。值得注意的是,对于盲目追求参数规模却缺乏实际应用价值的非关键性AI训练任务,碳价成本将成为约束其无序扩张的有效经济杠杆,推动算力资源流向高价值、低能耗的领域。

推行绿色算力的政府采购。联合国贸易和发展会议在《2024数字经济报告》^[23]中提到,政府和公共部门应充分发挥市场引导作用,通过可持续公共采购来创造绿色算力的早期市场。在采购数字基础设施与服务时纳入全生命周期碳足迹评估,并优先

选择符合能效标杆和绿电消纳要求的服务商。此类需求侧的政策引导不仅能为绿色AI企业赋予稳定的市场预期，更能凭借供应链传导机制倒逼上游芯片制造商优化制程工艺及下游应用开发商采用更高效的算法模型，形成全产业链的绿色转型动力^[139]。

设计差异化的税收优惠与绿色金融激励机制。一方面，构建“奖惩并重”的财政政策体系。对采用液冷技术、余热回收系统及自建分布式光伏的AI企业，可以给予企业所得税减免或固定资产加速折旧等优惠；反之，对能效指标低于行业准入水平的落后产能，执行阶梯电价或惩罚性电价。另一方面，深化绿色金融工具在算力领域的应用，鼓励金融机构推出与碳减排绩效挂钩的可持续发展关联贷款^[140]。通过降低绿色AI项目的融资成本，引导资本要素向环境友好型技术路径集聚，从源头上利用经济理性来抑制粗放式的算力基础设施建设。

5 主要结论和未来研究方向

5.1 主要结论

本文将AI置于全球气候治理与中国“双碳”目标的双重背景下，系统阐述了AI作为一种GPT所呈现的气候影响悖论。研究发现，AI技术的发展与能源系统之间存在着深刻的辩证关系。一方面，AI算力基础设施正以指数级扩张，成为新兴的能源消费巨头，给气候变化带来巨大的碳排放压力；另一方面，AI具有强大的优化和预测能力，可以作为助力能源体系重塑、促进产业低碳转型的关键赋能工具。在中国特定的情境下，这种“双刃剑”的特点更为突出：既要应对以煤为主的能源结构对算力清洁化造成的制约，又要肩负起利用AI推动经济高质量发展并实现“双碳”目标的关键使命。因此，科学评估AI技术的净环境影响，不仅同技术本身的可持续性有关，而且是落实国家绿色发展目标的关键课题。

从环境成本维度看，AI全生命周期的能源消耗和碳足迹具有总量激增和结构复杂的特征。在运营阶段，随着模型参数和算力规模的爆发式增长，数据中心对电力的需求攀升，且呈现出高度的空间集聚性，给局部电网造成了非对称冲击。尤其是在大语言模型的推理阶段，高频次的触发使其累积能耗

已逐渐超过训练阶段，成为AI电力需求中日益关键的组成部分。在全生命周期视角下，硬件制造阶段的隐含碳排放不容忽视。尽管先进制程芯片优化了单位算力能效，但其制造过程能源强度高且设备更新周期日益缩短，加剧了资源消耗和电子废物累积所带来的挑战。

从环境收益维度看，AI在赋能能源转型与供应链减排方面展现出明显的正外部性。研究表明，AI技术可以通过资源评估与气象预测大幅提高可再生能源的并网消纳能力，还可以加速低碳材料研发创新，推动产业链供应链的绿色转型。

然而，评估AI净气候影响的难点在于技术效率提升可能诱发的回弹效应。研究表明，AI在物质、经济和社会行为层面均存在复杂的回弹机制：物质层面体现为“规模效应”的增长会抵消数字化带来的“替代效应”；经济层面体现为算力成本的下降会引发总需求的增长；社会行为层面表现为AI诱导的消费与时间再分配增添了隐性的环境成本。

迈向可持续AI的关键在于，构建涵盖技术优化与市场约束的协同治理体系。面对数据透明度缺失、软硬件协同不足及能源结构制约等系统性挑战，单纯的技术改进已难以为继，必须采取“技术优化”与“市场约束”双轮驱动的协同治理路径。技术侧需形成统一而透明的全生命周期监管框架，冲破能耗“数据黑箱”，推进“源荷互动”的时空调节策略，达成算力与绿电的深度整合；市场侧则要通过碳定价、能耗配额交易及绿色政府采购等经济手段，将环境外部性内部化，从源头上遏制能源回弹。只有推动系统性治理变革，才能充分释放AI的低碳赋能潜力，使其真正成为实现全球碳中和目标的绿色引擎。

5.2 未来研究方向

尽管当前研究初步揭示了AI的气候影响悖论，但在构建真正可持续的数字未来方面，仍面临“数据黑箱”、因果机制模糊、成本收益分配不公及地缘政治干扰等挑战。未来，研究应着力在以下4个维度实现突破：

(1) 突破“数据黑箱”与核算壁垒，构建全生命周期的环境足迹核算体系。当前该领域面临的首要挑战是数据基础薄弱与核算边界模糊，环境评估

缺少统一标准,很大程度上依赖科技公司的自愿披露信息,这使得人们对硬件制造的隐含碳、水足迹等指标的关注度严重不足。针对这一难点,未来研究亟需打破单一的运营能耗视角,参考动态投入产出模型及内生资本核算的前沿方法^[141-142],将AI基础设施的资本形成和折旧过程纳入动态评估框架,开发标准化的核算方法论,建设开放透明的行业数据库。关键在于建立包含硬件制造、模型开发训练和推理应用的全链条监测体系。在此基础上,深入探究算力网络与能源系统的耦合机制,重点考察如何通过智能调度达成算力负载与可再生能源发电曲线在时空上的动态匹配,量化评估“东数西算”工程等跨区域算力调配模式的真实减排潜力,为将数据中心转化为支撑新型电力系统的“柔性调节资源”给予坚实的数据保障。

(2) 破解因果识别难题,加强AI系统性回弹效应的量化研究。尽管较多文献认可AI在微观层面提升能效的潜力,但对于其诱发的宏观回弹效应却缺乏实证研究支撑。未来,研究亟需从环境经济学、社会心理学和复杂系统科学的跨学科视角,拓展因果推断与系统模拟的方法论工具箱。一方面,使用准自然实验或结构方程模型,通过实施碳税、调整电价等外生政策冲击精准识别回弹效应的因果链条;另一方面,采用基于主体的建模,模拟微观主体的非线性适应性行为,动态推演从微观决策到宏观能耗的复杂涌现路径。核心任务在于,梳理“技术赋能减排”与“需求反弹增排”在不同时空尺度与行业场景下的动态博弈机制,精准测度AI应用的净环境效益边界。这有益于学术界冲破线性的“技术乐观主义”误区,从复杂系统论的角度客观评估AI的真实环境外部性,进而为制定有效抑制回弹效应的配套调节政策提供科学依据。

(3) 重视全球数字能源正义,重构包容性的全球价值链。现实情况表明,数字转型带来的成本与收益在全球范围内分配失衡,许多资源丰富的发展中国家因矿物开采和电子废物处理承受了沉重的环境代价,却未能公平分享数字红利。针对此种结构性不平等,未来研究需引入“环境正义”视角,探索如何弥合“绿色数字鸿沟”。重点在于,规划跨国碳责任分担机制及绿色供应链体系,帮助提升“全

球南方”国家在数字价值链中的地位。同时,探讨通过技术转移、资金支持及国际标准制定,确保全球AI产业的演进进程兼具生态可持续性与社会公正性,避免将环境外部性单向转移至欠发达地区。

(4) 探索地缘政治约束下的供应链韧性与绿色自主创新路径。当前国际形势复杂多变,AI芯片断供与技术封锁均为AI气候效应无法轻视的外部影响因素。硬件供应链的断裂可能迫使部分国家长期陷入制程落后造成的高能耗算力设施困境,进而加剧碳排放锁定效应。未来研究应结合定向技术变革理论,透彻剖析如何在外部约束下重塑具有环境友好属性的技术轨迹。重点在于探究如何通过算法架构优化、软硬件协同设计等非制程依赖型创新来改善计算能效,探索一条既能保障供应链韧性,又能规避技术脱钩带来额外环境成本的绿色自主创新道路,为后发国家在有限条件下实现低碳数字化转型给予理论依据。

参考文献

- [1] JONES C I. The AI dilemma: Growth versus existential risk[J]. *American Economic Review: Insights*, 2024, 6(4): 575–590.
- [2] ZHANG A B, DANISH, CHEN R. The paradox of artificial intelligence and environmental performance: The critical role of energy transition and institutional quality[J]. *Journal of Environmental Management*, 2025, 394: 127330.
- [3] FRIMPONG V. The sustainability paradox of artificial intelligence: How AI both saves and challenges resource management efforts[EB/OL]. [2026-01-04]. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5176930.
- [4] International Energy Agency. Energy and AI[R]. Paris, 2025.
- [5] Human-Centered Artificial Intelligence. Artificial intelligence index report 2025[R]. Stanford, 2025.
- [6] SHEHABI A, SMITH S J, HUBBARD A, et al. 2024 United States Data Center energy usage report[R]. Berkeley: Lawrence Berkeley National Laboratory, 2024.
- [7] 国家数据局. 数字中国发展报告(2024年)[EB/OL]. 北京, 2025.
- [8] 中国信息通信研究院. 绿色算力发展研究报告(2025年)[EB/OL]. (2025-07-24)[2025-12-13]. https://www.caict.ac.cn/kxyj/qwfb/ztbg/202507/t20250724_686833.htm.
- [9] 卫嘉, 白宇. 截至今年6月底我国智能算力规模达788亿亿次/秒[N]. 人民日报, 2025-09-02(10).
- [10] CHEN Z M, XIONG Q Y, DUAN J H, et al. AI carbon footprint in China sets to double post-2030 carbon peaking[J]. *Energy Economics*, 2025, 150: 108880.
- [11] United Nations Educational, Scientific and Cultural Organiza-

- tion. Recommendation on the ethics of artificial intelligence [M]. Paris: United Nations Educational, Scientific and Cultural Organization, 2022.
- [12] KOOMEY J G. Estimating total power consumption by servers in the U.S. and the world[EB/OL]. (2007-02-15)[2025-12-13]. <https://fepropaz.com/wp-content/uploads/2019/10/Koomey2007.pdf>.
- [13] MASANET E, SHEHABI A, LEI N A, et al. Recalibrating global data center energy-use estimates[J]. *Science*, 2020, 367 (6481): 984–986.
- [14] KOOMEY J G. Growth in data center electricity use 2005 to 2010[EB/OL]. (2011-08-01)[2025-12-10]. https://www.alexandrobarras.com/wp-content/uploads/old/4363/Growth_in_Data_Center_Electricity_use_2005_to_2010.pdf.
- [15] International Energy Agency. Electricity 2024: Analysis and forecast to 2026[R]. Paris, 2024.
- [16] International Energy Agency. Electricity 2025: Analysis and forecast to 2027[R]. Paris, 2025.
- [17] International Energy Agency. Data centres and data transmission networks[EB/OL]. (2023-07-11)[2025-12-13]. <https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks>.
- [18] International Energy Agency. World energy outlook 2025[R]. Paris, 2025.
- [19] ISTRATE R, TULUS V, GRASS R N, et al. The environmental sustainability of digital content consumption[J]. *Nature Communications*, 2024, 15: 3724.
- [20] 张翱, 孙久文. 数据要素空间黏性的区域经济增长极化效应: 来自国家超级计算中心建设的证据[J]. *中国软科学*, 2025(7): 97–108.
- [21] LI T, YU L, MA Y B, et al. Carbon emissions of 5G mobile networks in China[J]. *Nature Sustainability*, 2023, 6(12): 1620–1631.
- [22] MAJUMDER S, DONG L, DOUDI F, et al. Exploring the capabilities and limitations of large language models in the electric energy sector[J]. *Joule*, 2024, 8(6): 1544–1549.
- [23] United Nations Conference on Trade And Development. Digital economy report 2024[R]. Geneva, 2024.
- [24] KAACK L H, DONTI P L, STRUBELL E, et al. Aligning artificial intelligence with climate change mitigation[J]. *Nature Climate Change*, 2022, 12(6): 518–527.
- [25] Google Sustainability. Google environmental report 2024[R/OL]. (2024-07-10)[2025-12-13]. <https://sustainability.google/reports/google-2024-environmental-report>.
- [26] Baidu. Baidu 2024 environmental, social and governance (ESG) report[R/OL]. (2025-04-28)[2025-12-14]. <https://esg.baidu.com/ESGReport>.
- [27] Microsoft. 2025 environmental sustainability report: Accelerating progress to 2030[R/OL]. (2025-05-29)[2025-12-14]. <https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/msc/documents/presentations/CSR/2025-Microsoft-Environmental-Sustainability-Report-PDF.pdf>.
- [28] O'BRIEN I. Data center emissions probably 662% higher than big tech claims. Can it keep up the ruse?[EB/OL]. (2024-09-15)[2025-12-08]. <https://www.eceee.org/all-news/news/news-2024/data-center-emissions-probably-662-higher-than-big-tech-claims-can-it-keep-up-the-ruse>.
- [29] VERDECCHIA R, SALLOU J, CRUZ L. A systematic review of green AI[J]. *WIREs Data Mining and Knowledge Discovery*, 2023, 13(4): e1507.
- [30] SEVILLA J, HEIM L, HO A, et al. Compute trends across three eras of machine learning[C]//2022 International Joint Conference on Neural Networks (IJCNN). Padua, Italy, 2022.
- [31] EPOCH AI. Data on AI models[EB/OL]. (2024-06-19)[2025-12-13]. <https://epoch.ai/data/ai-models>.
- [32] PATTERSON D, GONZALEZ J, LE Q, et al. Carbon emissions and large neural network training[EB/OL]. (2021-04-23)[2025-12-13]. <https://arxiv.org/abs/2104.10350>.
- [33] STRUBELL E, GANESH A, MCCALLUM A. Energy and policy considerations for deep learning in NLP[C]//57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy, 2019.
- [34] 宋德勇, 华斯琦, 陈梁, 等. 从“绿色悖论”到“双化协同”: 智能化转型与企业碳排放[J]. *中国人口·资源与环境*, 2025, 35(7): 34–46.
- [35] THOPPILAN R, DE FREITAS D, HALL J, et al. LaMDA: Language models for dialog applications[EB/OL]. (2022-02-10)[2025-12-20]. <https://arxiv.org/abs/2201.08239>.
- [36] RAE J W, BORGEAUD S, CAI T, et al. Scaling language models: Methods, analysis & insights from training gopher[EB/OL]. (2022-01-21)[2025-12-20]. <https://arxiv.org/abs/2112.11446>.
- [37] PATTERSON D, GONZALEZ J, HÖLZLE U, et al. The carbon footprint of machine learning training will plateau, then shrink[J]. *Computer*, 2022, 55(7): 18–28.
- [38] LUCCIONI A S, VIGUIER S, LIGOZAT A L. Estimating the carbon footprint of bloom, a 176B parameter language model[J]. *The Journal of Machine Learning Research*, 2023, 24(1): 11990–12004.
- [39] CHOWDHURY A, NARANG S, DEVLIN J, et al. Palm: Scaling language modeling with pathways[J]. *The Journal of Machine Learning Research*, 2023, 24(1): 11324–11436.
- [40] DeepSeek-AI, LIU A X, FENG B, et al. DeepSeek-V3 technical report[EB/OL]. (2025-02-18)[2025-12-20]. <https://arxiv.org/abs/2412.19437>.
- [41] DE VRIES A. The growing energy footprint of artificial intelligence[J]. *Joule*, 2023, 7(10): 2191–2194.
- [42] AVELAR V, DONOVAN P, LIN P, et al. The AI disruption: Challenges and guidance for data center design[EB/OL]. [2025-12-13]. https://media.datacenterdynamics.com/media/documents/Schneider_Electric_-_Modernization_WP110_-_V1.1_EN.pdf.
- [43] JI Z Y, JIANG M. A systematic review of electricity demand for large language models: Evaluations, challenges, and solutions

- [J]. *Renewable and Sustainable Energy Reviews*, 2026, 225: 116159.
- [44] LEPIKHIN D, LEE H J, XU Y Z, et al. GShard: Scaling giant models with conditional computation and automatic sharding [EB/OL]. (2020-06-30)[2025-12-20]. <https://arxiv.org/abs/2006.16668>.
- [45] O'DONNELL J. DeepSeek might not be such good news for energy after all [EB/OL]. (2025-01-31)[2025-12-06]. <https://www.technologyreview.com/2025/01/31/1110776/deepseek-might-not-be-such-good-news-for-energy-after-all>.
- [46] GUO S, LI Y L, HU Y H, et al. Embodied energy in service industry in global cities: A study of six Asian cities [J]. *Land Use Policy*, 2020, 91: 104264.
- [47] LORENZINI R. Digital & environment: How to evaluate server manufacturing footprint, beyond greenhouse gas emissions? [EB/OL]. (2021-11-26)[2025-12-15]. <https://boavizta.org/en/blog/empreinte-de-la-fabrication-d-un-serveur>.
- [48] Dell Technologies. Life cycle assessment of Dell R740 [EB/OL]. (2019-06-07)[2025-12-14]. https://www.delltechnologies.com/asset/en-us/products/servers/technical-support/Full_LCA_-_Dell_R740.pdf.
- [49] BARDON M G, WUYTENS P, RAGNARSSON L Å, et al. DTCO including sustainability: Power-performance-area-cost-environmental score (PPACE) analysis for logic technologies [C]//2020 IEEE International Electron Devices Meeting (IEDM). San Francisco, USA, 2020.
- [50] SCHNEIDER I, XU H, BENECKE S, et al. Life-cycle emissions of AI hardware: A cradle-to-grave approach and generational trends [EB/OL]. (2025-02-01)[2025-12-06]. <https://arxiv.org/abs/2502.01671>.
- [51] VARAS A, VARADARAJAN R, GOODRICH J, et al. Strengthening the global semiconductor supply chain in an uncertain era [R/OL]. (2021-04-01)[2025-12-14]. <https://www.semiconductors.org/strengthening-the-global-semiconductor-supply-chain-in-an-uncertain-era>.
- [52] WANG P, ZHANG L Y, TZACHOR A, et al. E-waste challenges of generative artificial intelligence [J]. *Nature Computational Science*, 2024, 4(11): 818–823.
- [53] 张涛, 陈志顺. 生态补偿视角下的电子垃圾治理策略 [J]. *中国行政管理*, 2019, 35(6): 156–157.
- [54] ROLNICK D, DONTI P L, KAACK L H, et al. Tackling climate change with machine learning [J]. *ACM Computing Surveys*, 2023, 55(2): 1–96.
- [55] KRECHOWICZ A, KRECHOWICZ M, POZETA K. Machine learning approaches to predict electricity production from renewable energy sources [J]. *Energies*, 2022, 15(23): 9146.
- [56] SHARMA N, SHARMA P, IRWIN D, et al. Predicting solar generation from weather forecasts using machine learning [C]//2011 IEEE International Conference on Smart Grid Communications (SmartGridComm). Brussels, Belgium, 2011.
- [57] 周剑, 聂孝婷, 庞可欣, 等. 基于深度强化学习算法的跨区域多微电网系统扩展规划研究 [J]. *中国管理科学*, 2025, 33(11): 41–53.
- [58] XIE L, HUANG T, ZHENG X T, et al. Energy system digitization in the era of AI: A three-layered approach toward carbon neutrality [J]. *Patterns*, 2022, 3(12): 100640.
- [59] 屠西伟, 史丹. 数字产业集聚与企业能源效率改进 [J]. *数量经济技术经济研究*, 2025, 42(10): 70–89.
- [60] DING C, KE J, LEVINE M, et al. Potential of artificial intelligence in reducing energy and carbon emissions of commercial buildings at scale [J]. *Nature Communications*, 2024, 15: 5916.
- [61] 杨震, 赵静洲, 林依挺, 等. 数据中心PUE能效优化的机器学习方法 [J]. *系统工程理论与实践*, 2022, 42(3): 801–810.
- [62] LAL A, YOU F Q. Climate sustainability through AI-crypto synergies and energy transition in the digital landscape to cut 0.7 GtCO₂e by 2030 [J]. *Environmental Science & Technology*, 2025, 59(12): 6061–6073.
- [63] 罗雄. “人工智能+”驱动中国式产业现代化: 理论逻辑与模式创新 [J]. *经济学家*, 2025(11): 47–56.
- [64] ZHANG Z X, CHEN M, ZHONG T, et al. Carbon mitigation potential afforded by rooftop photovoltaic in China [J]. *Nature Communications*, 2023, 14: 2347.
- [65] VESPIGNANI J, SMYTH R. Artificial intelligence investments reduce risks to critical mineral supply [J]. *Nature Communications*, 2024, 15: 7304.
- [66] MERCHANT A, BATZNER S, SCHOENHOLZ S S, et al. Scaling deep learning for materials discovery [J]. *Nature*, 2023, 624(7990): 80–85.
- [67] ZENI C, PINSLER R, ZÜGNER D, et al. A generative model for inorganic materials design [J]. *Nature*, 2025, 639(8055): 624–632.
- [68] WANG H S, YANG G Q, YUE Z Y. Breaking through ingrained beliefs: Revisiting the impact of the digital economy on carbon emissions [J]. *Humanities and Social Sciences Communications*, 2023, 10: 609.
- [69] 鲁锦涛, 宋祖昕, 宋马林. 人工智能对区域低碳创新的影响: 基于环境规制的调节效应 [J/OL]. *中国管理科学*, 1–16(2025-10-10)[2025-12-10]. <https://doi.org/10.16381/j.cnki.issn1003-207x.2025.0155>.
- [70] VERENDEL V. Tracking artificial intelligence in climate inventions with patent data [J]. *Nature Climate Change*, 2023, 13(1): 40–47.
- [71] EYRING V, COLLINS W D, GENTINE P, et al. Pushing the frontiers in climate modelling and analysis with machine learning [J]. *Nature Climate Change*, 2024, 14(9): 916–928.
- [72] VOROTYNTSEV P, GORDIENKO Y, ALIENIN O, et al. Satellite image segmentation using deep learning for deforestation detection [C]//2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON). Lviv, Ukraine, 2021.
- [73] SCHUIT B J, MAASAKKERS J D, BIJL P, et al. Automated detection and monitoring of methane super-emitters using satellite data [J]. *Atmospheric Chemistry and Physics*, 2023, 23(16): 10061–10073.

- 9071–9098.
- [74] BURNS C, BOLLARD B, NARAYANAN A. Machine-learning for mapping and monitoring shallow coral reef habitats[J]. *Remote Sensing*, 2022, 14(11): 2666.
- [75] 钞小静, 沈路. 创新价值链视角下人工智能技术对制造业企业绿色创新效率的影响[J]. *经济学动态*, 2025(4): 50–67.
- [76] HONG Z X, XIAO K. Digital economy structuring for sustainable development: The role of blockchain and artificial intelligence in improving supply chain and reducing negative environmental impacts[J]. *Scientific Reports*, 2024, 14: 3912.
- [77] XIE X M, HAN Y H, TAN H. Greening China's digital economy: Exploring the contribution of the East–West Computing Resources Transmission Project to CO₂ reduction[J]. *Humanities and Social Sciences Communications*, 2024, 11: 466.
- [78] SORRELL S. Jevons' paradox revisited: The evidence for backfire from improved energy efficiency[J]. *Energy Policy*, 2009, 37(4): 1456–1469.
- [79] JEVONS W S. The coal question[M]//*The Economics of Population*. London: Macmillan and Co., 1866: 193–204.
- [80] DUMONT A, MAYOR B, LÓPEZ-GUNN E. Is the rebound effect or Jevons paradox a useful concept for better management of water resources? Insights from the irrigation modernisation process in Spain[J]. *Aquatic Procedia*, 2013, 1: 64–76.
- [81] GIAMPIETRO M, MAYUMI K. Unraveling the complexity of the Jevons paradox: The link between innovation, efficiency, and sustainability[J]. *Frontiers in Energy Research*, 2018, 6: 26.
- [82] DIMITROPOULOS A, OUESLATI W, SINTEK C. The rebound effect in road transport: A meta-analysis of empirical studies[J]. *Energy Economics*, 2018, 75: 163–179.
- [83] ONAT N C, MANDOURI J, KUCUKVAR M, et al. Rebound effects undermine carbon footprint reduction potential of autonomous electric vehicles[J]. *Nature Communications*, 2023, 14: 6258.
- [84] 史丹, 张瑾, 史可寒, 等. 新型工业化的典型特征及其对能源消费的影响[J]. *中国人口·资源与环境*, 2025, 35(1): 1–16.
- [85] TOMLINSON B, BLACK R W, PATTERSON D J, et al. The carbon emissions of writing and illustrating are lower for AI than for humans[J]. *Scientific Reports*, 2024, 14: 3732.
- [86] KAPLAN J, MCCANDLISH S, HENIGHAN T, et al. Scaling laws for neural language models[EB/OL]. (2020-01-23)[2025-12-06]. <https://arxiv.org/abs/2001.08361>.
- [87] HOEFLER T, HENDEL A, ROWETH D. The convergence of hyperscale data center and high-performance computing networks[J]. *Computer*, 2022, 55(7): 29–37.
- [88] TREMAYNE-PENGELLY A. Amid the A.I. boom, these states have become data center hubs[EB/OL]. (2024-12-04)[2025-12-14]. <https://observer.com/2024/12/ai-demand-where-data-centers-using-most-energy>.
- [89] MASANET E, LEI N A, KOOMEY J. To better understand AI's growing energy use, analysts need a data revolution[J]. *Joule*, 2024, 8(9): 2427–2436.
- [90] HÖTTE K, SOMERS M, THEODORAKOPOULOS A. Technology and jobs: A systematic literature review[J]. *Technological Forecasting and Social Change*, 2023, 194: 122750.
- [91] ITTEN R, HISCHIER R, ANDRAE A S G, et al. Digital transformation: Life cycle assessment of digital services, multifunctional devices and cloud computing[J]. *The International Journal of Life Cycle Assessment*, 2020, 25(10): 2093–2098.
- [92] TANTAU A, NICULESCU E. The role of Power Purchase Agreements for the promotion of green energy and the transition to a zero carbon economy[J]. *Proceedings of the International Conference on Business Excellence*, 2022, 16(1): 1237–1245.
- [93] 石峰, 杨扬, 袁韵, 等. 人工智能驱动下的营销变革[J]. *中国管理科学*, 2025, 33(1): 111–123.
- [94] KUMAR V, ASHRAF A R, NADEEM W. AI-powered marketing: What, where, and how?[J]. *International Journal of Information Management*, 2024, 77: 102783.
- [95] HARTMANN P, MARCOS A, CASTRO J, et al. Perspectives: Advertising and climate change—Part of the problem or part of the solution[J]. *International Journal of Advertising*, 2023, 42(2): 430–457.
- [96] NIINIMÄKI K, PETERS G, DAHLBO H, et al. The environmental price of fast fashion[J]. *Nature Reviews Earth & Environment*, 2020, 1(4): 189–200.
- [97] BUHL J, ACOSTA J. Work less, do less: Working time reductions and rebound effects[J]. *Sustainability Science*, 2016, 11(2): 261–276.
- [98] CHEN P, MALIK A, SONG J H, et al. Disparities and flows analysis of global leisure carbon footprints[J]. *Resources, Conservation and Recycling*, 2026, 224: 108535.
- [99] LUCCIONI A S, STRUBELL E, CRAWFORD K. From efficiency gains to rebound effects: The problem of jevons' paradox in AI's polarized environmental debate[C]//2025 ACM Conference on Fairness, Accountability, and Transparency. Athens, Greece, 2025.
- [100] ARORA N, CABANNES T, GANAPATHY S, et al. Quantifying the sustainability impact of Google Maps: A case study of Salt Lake City[EB/OL]. (2021-10-21)[2025-12-14]. <https://arxiv.org/abs/2111.03426>.
- [101] ERTEL W, BONENBERGER C. Rebound effects caused by artificial intelligence and automation in private life and industry[J]. *Sustainability*, 2025, 17(5): 1988.
- [102] FAIZ A, KANEDA S, WANG R H, et al. LLMCarbon: Modeling the end-to-end carbon footprint of large language models[EB/OL]. (2024-01-19)[2025-12-14]. <https://arxiv.org/abs/2309.14393>.
- [103] COURTY B, SCHMIDT V, KAMAL G, et al. Mlco2/codecarbon: V2.4.1[CP/OL]. (2024-05-10)[2025-12-14]. <https://zenodo.org/records/11171501>.
- [104] LACOSTE A, LUCCIONI A, SCHMIDT V, et al. Quantifying the carbon emissions of machine learning[EB/OL]. (2019-11-

- 04)[2025-12-14]. <https://arxiv.org/abs/1910.09700>.
- [105] European Union. Artificial intelligence act[EB/OL]. (2024-07-12)[2025-12-14]. <https://artificialintelligenceact.eu/the-act>.
- [106] CHANG Y P, WANG X, WANG J D, et al. A survey on evaluation of large language models[J]. *ACM Transactions on Intelligent Systems and Technology*, 2024, 15(3): 1–45.
- [107] LEISERSON C E, THOMPSON N C, EMER J S, et al. There's plenty of room at the top: What will drive computer performance after Moore's law[J]. *Science*, 2020, 368(6495): eaam9744.
- [108] HUANG Y P, CHENG Y L, BAPNA A, et al. Gpipe: Efficient training of giant neural networks using pipeline parallelism[EB/OL]. (2019-07-25)[2025-12-14]. <https://arxiv.org/abs/1811.06965>.
- [109] NARAYANAN D, SHOEYBI M, CASPER J, et al. Efficient large-scale language model training on GPU clusters using megatron-LM[C]//International Conference for High Performance Computing, Networking, Storage and Analysis. St. Louis, USA, 2021.
- [110] CHEN T Q, XU B, ZHANG C Y, et al. Training deep nets with sublinear memory cost[EB/OL]. (2016-04-22)[2025-12-14]. <https://arxiv.org/abs/1604.06174>.
- [111] MICKEVICIUS P, NARANG S, ALBEN J, et al. Mixed precision training[EB/OL]. (2018-02-15)[2025-12-14]. <https://arxiv.org/abs/1710.03740>.
- [112] GOZALO-BRIZUELA R, GARRIDO-MERCHAN E C. ChatGPT is not all you need. A state of the art review of large generative AI models[EB/OL]. (2023-01-11)[2025-12-14]. <https://arxiv.org/abs/2301.04655>.
- [113] GALE T, NARAYANAN D, YOUNG C, et al. Megablocks: Efficient sparse training with mixture-of-experts[J]. *Proceedings of Machine Learning and Systems*, 2023, 5: 288–304.
- [114] PODDAR S, KOLEY P, MISRA J, et al. Towards sustainable nlp: Insights from benchmarking inference energy in large language models[C]//2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies. Albuquerque, USA, 2025.
- [115] WU C J, RAGHAVENDRA R, GUPTA U, et al. Sustainable AI: Environmental implications, challenges and opportunities[J]. *Proceedings of Machine Learning and Systems*, 2022, 4: 795–813.
- [116] YARALLY T, CRUZ L, FEITOSA D, et al. Uncovering energy-efficient practices in deep learning training: Preliminary steps towards green AI[EB/OL]. (2023-03-24)[2025-12-14]. <https://arxiv.org/abs/2303.13972>.
- [117] JOUPPI N P, YOUNG C, PATIL N, et al. In-datacenter performance analysis of a tensor processing unit[C]//44th Annual International Symposium on Computer Architecture. Toronto, Canada, 2017.
- [118] UCHIYAMA K. Nvidia Blackwell platform arrives to power a new era of computing[EB/OL]. (2024-03-18)[2025-12-14]. <https://nvidianews.nvidia.com/news/nvidia-blackwell-platform-arrives-to-power-a-new-era-of-computing>.
- [119] Nvidia. Power efficiency[EB/OL]. (2025-09-13)[2025-12-14]. <https://www.nvidia.com/en-eu/glossary/power-efficiency>.
- [120] HEIKKILÄ M. We're getting a better idea of AI's true carbon footprint[EB/OL]. (2022-11-14)[2025-12-14]. <https://www.technologyreview.com/2022/11/14/1063192/were-getting-a-better-idea-of-ais-true-carbon-footprint>.
- [121] LEWIS M. Amazon just bought a 100% nuclear-powered data center[EB/OL]. (2024-03-05)[2025-12-14]. <https://electrek.co/2024/03/05/amazon-just-bought-a-100-nuclear-powered-data-center>.
- [122] CRAWFORD K. The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence[M]. New Haven: Yale University Press, 2021.
- [123] TERRELL M. New nuclear clean energy agreement with Kairos Power[EB/OL]. (2024-10-14)[2025-12-16]. <https://blog.google/company-news/outreach-and-initiatives/sustainability/google-kairos-power-nuclear-energy-agreement>.
- [124] 国家发展改革委, 国家数据局, 中央网信办, 等. 国家发展改革委等部门关于深入实施“东数西算”工程加快构建全国一体化算力网的实施意见: 发改数据〔2023〕1779号[A/OL]. (2023-12-25)[2025-12-14]. https://www.gov.cn/zhengce/zhengceku/202401/content_6924596.htm.
- [125] NSF Arctic Data Center. Understanding the arctic[EB/OL]. (2021-05-26)[2025-12-14]. <https://arcticdata.io/all/blog/2021/05/understanding-the-arctic>.
- [126] 中国信息通信研究院. 算力电力协同发展研究报告(2025年)[R]. 北京, 2025.
- [127] AL KEZ D, FOLEY A M, AHMED F W, et al. Potential of data centers for fast frequency response services in synchronously isolated power systems[J]. *Renewable and Sustainable Energy Reviews*, 2021, 151: 111547.
- [128] JAGANNADHARAO A, BECKAGE N, NAFUS D, et al. Timeshifting strategies for carbon-efficient long-running large language model training[J]. *Innovations in Systems and Software Engineering*, 2025, 21(2): 517–531.
- [129] HANAFY W A, LIANG Q L, BASHIR N, et al. CarbonScaler: Leveraging cloud workload elasticity for optimizing carbon-efficiency[J]. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2023, 7(3): 1–28.
- [130] CHIEN A A, LIN L Z X, NGUYEN H, et al. Reducing the carbon impact of generative AI inference (today and in 2035)[C]//2nd Workshop on Sustainable Computer Systems. Boston, USA, 2023.
- [131] RADOVANOVIĆ A, KONINGSTEIN R, SCHNEIDER I, et al. Carbon-aware computing for datacenters[J]. *IEEE Transactions on Power Systems*, 2023, 38(2): 1270–1280.
- [132] RUAN J Q, ZHU Y F, CAO Y J, et al. Privacy-preserving bi-level optimization of internet data centers for electricity-carbon collaborative demand response[J]. *IEEE Internet of Things Journal*, 2024, 11(14): 24948–24959.
- [133] WAN T, TAO Y C, QIU J, et al. Internet data centers participat-

- ing in electricity network transition considering carbon-oriented demand response[J]. *Applied Energy*, 2023, 329: 120305.
- [134] CHEN M, GAO C W, SONG M, et al. Internet data centers participating in demand response: A comprehensive review[J]. *Renewable and Sustainable Energy Reviews*, 2020, 117: 109466.
- [135] PIGOU A. *The Economics of Welfare*[M]. New York: Routledge, 2017.
- [136] Organisation for Economic Co-operation and Development. Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint[EB/OL]. (2022-11-15)[2025-12-16]. https://www.oecd.org/en/publications/measuring-the-environmental-impacts-of-artificial-intelligence-compute-and-applications_7babf571-en.html.
- [137] 朱子恒, 张策, 丁肇豪, 等. 数据中心纳入全国碳排放权交易市场机制研究[J]. *中国电机工程学报*, 2024, 44(14): 5562–5574.
- [138] SANDALOW D B, MCCORMICK C, KUCUKELBIR A, et al. Artificial intelligence for climate change mitigation roadmap (second edition)[EB/OL]. [2026-01-04]. <https://academiccommons.columbia.edu/doi/10.7916/z2ws-k833/download>.
- [139] SIMCOE T, TOFFEL M W. Government green procurement spillovers: Evidence from municipal building policies in California[J]. *Journal of Environmental Economics and Management*, 2014, 68(3): 411–434.
- [140] CHEN J, MENG W F, CHEN Y L, et al. To be an eco-and tech-friendly society: Impact research of green finance on AI innovation[J]. *Journal of Cleaner Production*, 2024, 466: 142900.
- [141] CHEN Z M, OHSHITA S, LENZEN M, et al. Consumption-based greenhouse gas emissions accounting with capital stock change highlights dynamics of fast-developing countries[J]. *Nature Communications*, 2018, 9: 3581.
- [142] CHEN Z M, CHEN P L, LENZEN M, et al. Global embodied energy flow and stock analysis with endogenous fixed capital[J]. *Environmental Science & Technology*, 2022, 56(23): 17197–17205.

The Climate Impact Paradox of Artificial Intelligence: Literature Review and Research Prospects

CHEN Zhanming¹, XIONG Qiyang¹, HU Zitong^{1,2}, LIU Shiyun¹,
MA Jianhong^{1,*}

1. School of Applied Economics, Renmin University of China, Beijing 100872, China;

2. School of Economics and Management, China University of Geosciences (Wuhan), Wuhan 430078, China

Abstract : In the context of global climate governance, artificial intelligence (AI) exerts a dual effect on climate change: acting as both a carbon-intensive energy consumer and an enabler of low-carbon transitions, which creates a complex climate impact paradox. By conducting a systematic literature review, this paper aims to clarify the relationship between AI, energy consumption, and carbon emissions. The review reveals that the exponential expansion of computing infrastructure not only exacerbates the impact of operational energy consumption on regional power grids but also generates embodied carbon emissions through hardware manufacturing and electronic waste disposal. Although AI demonstrates significant positive externalities in optimizing energy systems, driving green technology innovation, and reshaping supply chains, the rebound effects triggered by efficiency improvements across material, economic, and socio-behavioral dimensions often diminish or even offset its decarbonization benefits, thereby inducing the Jevons Paradox. Consequently, this paper argues that relying solely on exogenous technological progress is insufficient to mitigate the environmental risks of AI. Instead, it is imperative to construct a comprehensive governance framework that encompasses full life-cycle carbon footprint accounting, software-hardware synergetic innovation, and market-based incentive mechanisms. This mechanism aims to achieve the deep integration of computing power and green electricity, positioning AI as a powerful engine for realizing sustainable development goals.

Key Words : artificial intelligence; data center; carbon footprint; rebound effect; low-carbon transition