

基于Transformer的数据中心机房 三维温度场快速重构和预测方法

Fast Reconstruction and Prediction Method of Three-dimensional Temperature Field in Data Center Room Based on Transformer

朱旭¹,贺晓¹,高健¹,闫若飞¹,陈俊丞¹,姚贵策²(1. 中讯邮电咨询设计院有限公司,北京 100048;2. 北京航空航天大学,北京 100191)

Zhu Xu¹,He Xiao¹,Gao Jian¹,Yan Ruofei¹,Chen Juncheng¹,Yao Guice²(1. China Information Technology Designing & Consulting Institute Co.,Ltd.,Beijing 100048,China;2. Beihang University,Beijing 100191,China)

摘要:

机房内部气流组织的合理性关系到数据中心制冷系统的能耗、IT设备的性能和热环境的安全。传统气流组织模拟性能一般在小时级,使用场景受限,难以满足实时运营对机房温度场预测的时效性需求。提出一种基于Transformer的机房三维温度场快速重构和预测方法,通过融合深度学习模型与传统CFD,将气流组织预测时间降低至分秒级,全局平均预测精度误差控制在5%以内,从而使CFD仿真有效地从设计阶段应用到运维阶段,支撑数据中心的智慧运营业务。

关键词:

数据中心;温度场预测;Transformer;机器学习;计算流体力学

doi:10.12045/j.issn.1007-3043.2024.10.001

文章编号:1007-3043(2024)10-0001-06

中图分类号:TU248.7

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

The rationality of airflow organization inside the data center is related to the energy consumption of the cooling system, the performance of IT equipment, and the safety of the thermal environment. Traditional CFD simulation is generally at the hourly level, with limited usage scenarios, making it difficult to meet the timeliness requirements of real-time operation for temperature field prediction. It presents a fast reconstruction and prediction method for the three-dimensional temperature field of computer rooms based on Transformer. By integrating deep learning with traditional CFD, the prediction time can be reduced to the level of minutes and seconds, and the global average prediction error can be controlled within 5%, which enables CFD simulation to be effectively applied from the design stage to the operation and maintenance stage, supporting intelligent operation services.

Keywords:

Data center; Temperature distribution prediction; Transformer; Machine learning; CFD

引用格式:朱旭,贺晓,高健,等. 基于Transformer的数据中心机房三维温度场快速重构和预测方法[J]. 邮电设计技术,2024(10):1-6.

1 概述

随着物联网、大数据以及人工智能等新一代信息技术的推进和发展,全社会数字经济蓬勃发展,算力需求日益增长^[1]。据中国算力平台统计测算,2023年中国算力市场规模平均增速达到30%,同时数据中心用电量约为1500亿kWh,占全社会用电量的1.6%^[2]。

预计到2030年,我国数据中心能耗总量将超过4000亿kWh^[3]。近年来,为了推进数据中心等重点设施的绿色低碳升级,工信部等国家部门制定了相关政策,要求到2025年全国新建大型、超大型数据中心的PUE降到1.3以下,改建核心机房降到1.3以下^[4-5],数据中心的节能降碳面临巨大挑战。

数据中心冷却系统占数据中心总能耗的30%~55%^[6],由于数据中心各机房空间布局复杂,气流组织可能存在分布不均匀、冷热气流混掺等问题,进而导

收稿日期:2024-09-02

致冷却效果不佳,带来持续的能源浪费。通过人工测量、现场布置温度传感器等实验测量方法可以捕捉局部热点,但受限与空间和成本问题,难以获取机房全局的温度场信息。利用计算流体力学(Computational Fluid Dynamics, CFD)的方法,可有效地计算模拟出特定条件的机房温度场三维分布,从而为提高冷却效能提供参考^[7]。

然而,传统的CFD技术通常需要进行复杂的数值模拟和计算,耗费大量的计算资源与时间,针对数据中心机房气流组织的性能模拟通常需要小时级的时间成本,使用场景受限,难以满足数据中心实时运营阶段对于仿真数据的时效性需求。随着信息技术的发展,大量数据可被收集和积累,智能预测模型的高速性优势日益凸显。部分研究尝试将机器学习方法与传统CFD仿真技术结合,在给定的边界条件以及初始条件下,利用机器学习方法对流体域的三维温度场进行快速重构,从而实现温度场等物理量的智能预测^[8-11]。近年来,智能预测模型主要采用卷积神经网络、生成对抗网络、自编码器等方法。考虑到数据中心机房智能预测具备时序性强、复杂构型和强非线性的特点,上述方式的效果有限,亟需研发针对数据中心机房的三维温度场智能预测方法。

Transformer通过将注意力引入到自编码器模型中,允许模型在解码时关注输入序列的不同部分,可以有效提升模型性能。Transformer摒弃了RNN和CNN的结构,依赖于自注意力机制和残差前馈神经网络,在自然语言处理领域获得了广泛关注和应用。Transformer在处理长程依赖、并行计算、任务适应性和稳定性等方面具有一定的优势,其自注意力、上下文建模等机制不仅可以有效地捕捉复杂机房空间特征中的全局温度场趋势,还能关注到关键的局部热点信息。因此,本文基于Transformer深度学习模型和架构,通过对机房三维温度场信息的挖掘和学习,实现对数据中心三维温度场的快速重构和智能预测,为数据中心的智慧运营提供技术支撑。

2 基于Transformer的温度场快速重构和预测方法

2.1 三维温度场数据获取技术

本文以某大型数据中心机房为研究对象,对其建立CFD模型(见图1),并通过该模型进行多组运行工况的模拟,生成用于AI-CFD模型训练和验证的三维

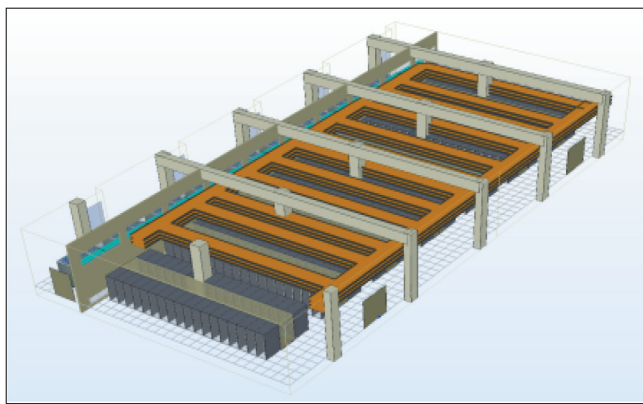


图1 机房CFD仿真模型示意

温度场数据。机房长为17.8 m,宽为35.8 m,高为6 m,包含7个封闭式冷通道、244个机柜和12个房间级空调,气流组织方式为封闭冷通道下送风和上回风。距离地板0.9 m高度以下为架构地板下方区域,0.9~3.1 m为机柜区域,3.1 m以上为机柜上方区域。

为了模拟数据中心机房的真实应用场景,本文设计了3种功率模式和2种空调模式作为工况配置,用于匹配和模拟实际数据中心机房空调的运作情况。3种功率模式分别为A(all)、S(single)与I(inverse),模式A表示当前列与之前所有列机柜功率同步变化,之后所有列功率均为零;模式S表示之前列所有机柜功率均为零,当前列功率变化,之后列均为满功率;模式I与S相反,表示之前列所有机柜均为满功率,当前列功率变化,之后列机柜功率均为零。2种空调模式分别为送风与回风模式。

机柜功率的变化可分为5档,机柜满负荷功率为5.5 kW,可调节的机柜功率为满功率的0%、25%、50%、75%以及100%;空调送风温度为18℃、20℃、22℃、24℃以及26℃5档。根据数据中心机房的二维平面(见图2),结合上述功率和温度参数设置模式,共

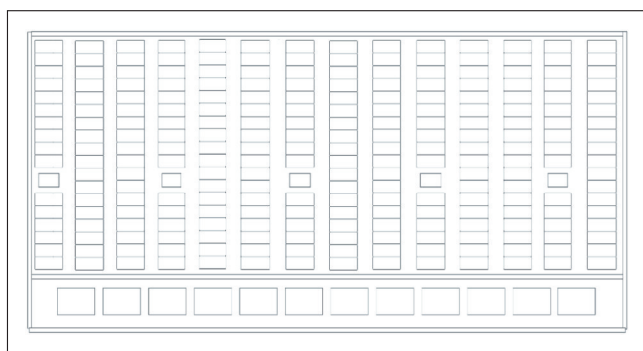


图2 数据中心机房二维平面示意

模拟了840多组机房三维温度场数据作为模型训练和验证数据集。

2.2 AI-CFD模型学习和预测架构

2.2.1 ITM和SDF特征构建

为了将输入功率与温度嵌入到输入向量,构建初始温度场(见图3),具体为将数据中心机房剖分为若干区域,再基于不同区域与空调对应的位置进行赋初值^[12-13]。

$$\text{ITM}(\text{area}) = \begin{cases} f(\text{AC}) + g(\text{power}), & \text{area} \in \text{cabinet} \\ f(\text{AC}) - g(\text{power}), & \text{area} \in \text{inlet} \\ f(\text{AC}), & \text{area} \in \text{out} \end{cases} \quad (1)$$

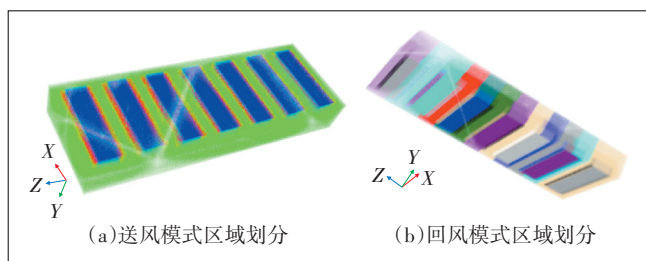


图3 送风和回风模式的初始温度场

其中,空间 area 分为机柜 cabinet、进风口 inlet 以及外部环境 out,当空间位于机柜内或进风口时,初始温度场由空调温度与机柜功率的函数共同决定;当空间位于外部环境时,初始温度场仅由空调温度的函数决定。

符号距离图又可以称为定向距离函数,在空间中的一个有限区域上确定一个点到区域边界的距离并同时对其符号进行定义,点在区域边界内部为正,外部为负,位于边界上时为零。考虑到机柜中常有热点存在,因此将机柜模型进行抽象处理,当空间点在机柜内部时,定义为正距离,当点在机柜外部时,定义为负距离(见图4)。

$$\text{SDF}(x) = \begin{cases} d(x, \partial\Omega), & x \in \Omega \\ -d(x, \partial\Omega), & x \in \Omega^c \end{cases} \quad (2)$$

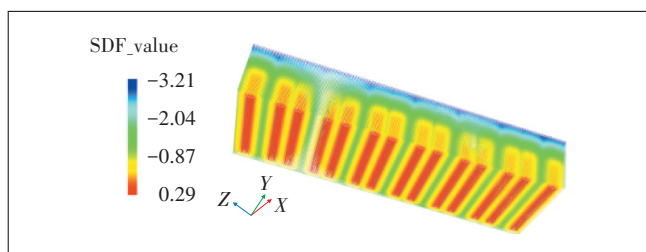


图4 机房符号距离图

其中, x 即为空间点坐标, Ω 为机柜内部空间, Ω^c 为机柜外部空间。

2.2.2 Transformer网络架构

本文结合Transformer的注意力机制,使用Transformer的Encoder模块,利用残差块进行深度模型搭建,构建改进的Transformer模型,改进的Transformer网络架构如图5所示。该网络架构既利用了生成式模型的优势,又降低了累计误差发生的可能性。通过读取输入特征,经过含有 N 个Block的Encoder后可将提取到的ITM和SDF特征传入全连接神经网络进行进一步映射,最终得到输出温度图。

注意力机制允许模型选择性地关注输入信息的重要部分,忽略其他不太相关的部分。这种选择性关注可以帮助模型更好地理解输入信息,并提高其处理信息的效率,在处理数据中心机房三维温度场高度非线性问题时具备优势。同时,因为注意力机制只关注输入信息的重要部分,而不是对所有信息进行同等处理,其计算效率较高,有助于提高模型的计算效率和性能。

进一步,将神经网络算法与残差链接融合,通过单独的神经网络将注意力机制的输出进行最后一次非线性映射,通过残差链接将前部层的特征直接迁移至后部层,使得浅层映射特征与深层映射特征进一步融合,从而消除梯度弥散。

2.3 三维温度场快速重构与预测

在训练阶段, AI-CFD模型输入工况主要包括机柜功率以及空调温度^[12-13],将计算出的初始温度图、符号距离图以及训练数据集输入至Transformer架构中进行模型训练,不断将预测结果与训练数据进行对比,反向传播进行更新,最终形成训练好的Transformer三维温度场预测模型(见图6),实现对数据中心三维温度场的快速重构。

在预测阶段,通过输入空调温度、机柜功率的工况信息,利用训练好的Transformer模型即可预测出机房实时的三维温度场信息(见图7)。

3 三维温度场预测结果和分析

3.1 学习率的影响

考虑到三维温度场智能预测的训练数据量较大,为节省模型训练时间和计算资源,以及防止模型出现过拟合现象,当 N 个 epoch 下的验证集损失没有减少时, AI-CFD模型训练会提前终止。分别设置以10个

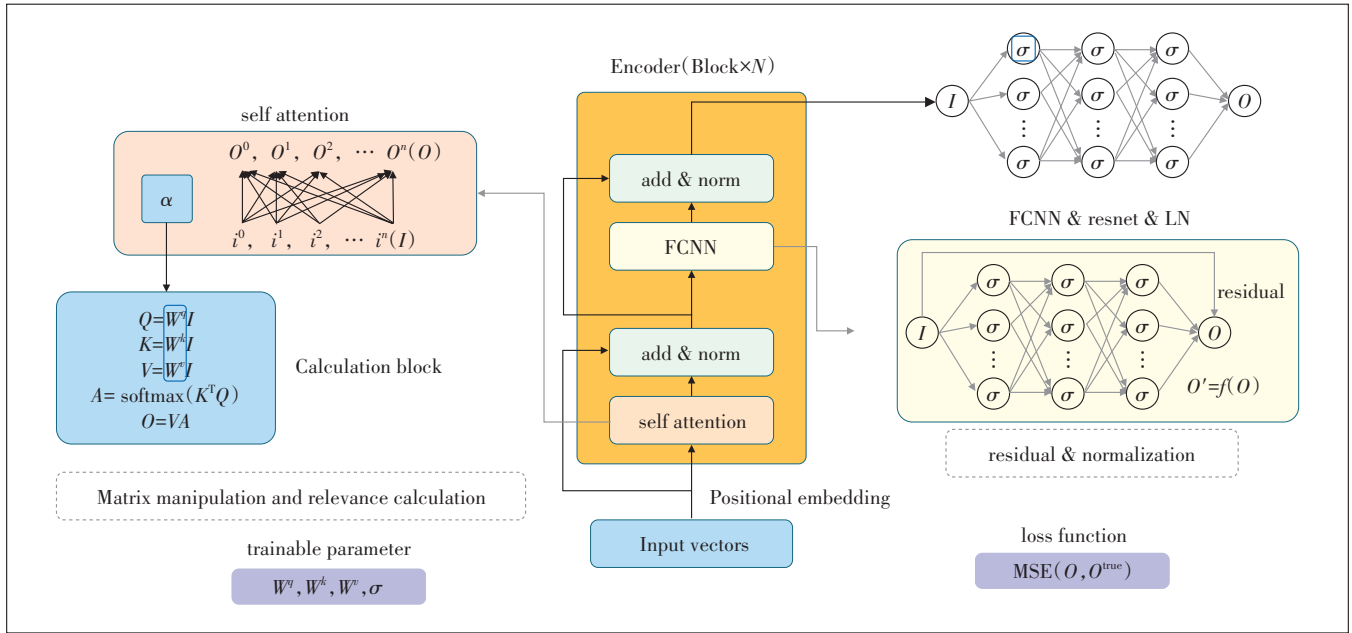


图5 改进的Transformer网络架构

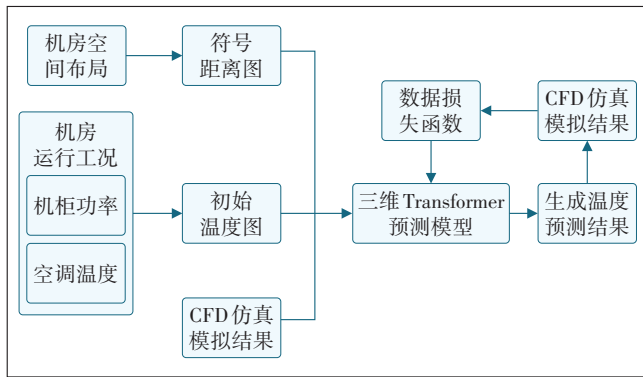


图6 AI-CFD模型离线训练

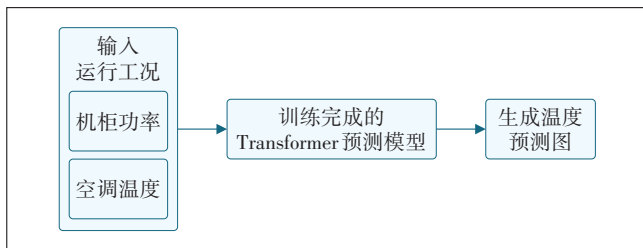


图7 AI-CFD模型在线预测

epoch(组别A)和20个epoch(组别B)为模型训练收敛的终止条件的分组,每个分组包含0.05、0.001、0.005和0.0001这4个学习率,共计8组参数设置工况。图8为8个工况下训练集和验证集的训练损失分布情况,整体上看,当学习率为0.001时,AI-CFD模型质量相

对较好。

3.2 自注意力机制头数量的影响

设置学习率为0.001,分别设置自注意力机制头的数量为3、4、5,共3组工况,研究自注意力头对模型性能的影响,结果如图9所示。整体上看,自注意力头数量对模型质量的影响不大,3组训练模型的损失基本一致。

3.3 不同工况验证

本节设置3组验证测试工况,具体如表1所示。

工况1、工况2、工况3的温度预测图与误差图分别如图10、11、12所示,每个工况在1.5 m处、2.5 m处以及三维空间的温度分布情况满足正常规律,本次用于比较结果的量化标准主要有整场平均绝对百分比误差GME以及整场均方误差MSE,分别如式(3)与式(4)所示。

$$GME = \frac{\sum \left| \frac{T_{pre} - T_{sigma}}{T_{sigma}} \right|}{N_{point}} \quad (3)$$

$$MSE = \frac{\sum (T_{pre} - T_{sigma})^2}{N_{point}} \quad (4)$$

其中, T_{pre} 为AI-CFD模型在空间某一点的预测温度结果, T_{sigma} 为同工况下6SigmaRoom模型某一点的仿真结果, N_{point} 为整个三维空间的点数量。

根据上述公式,工况1的整场平均绝对百分比误

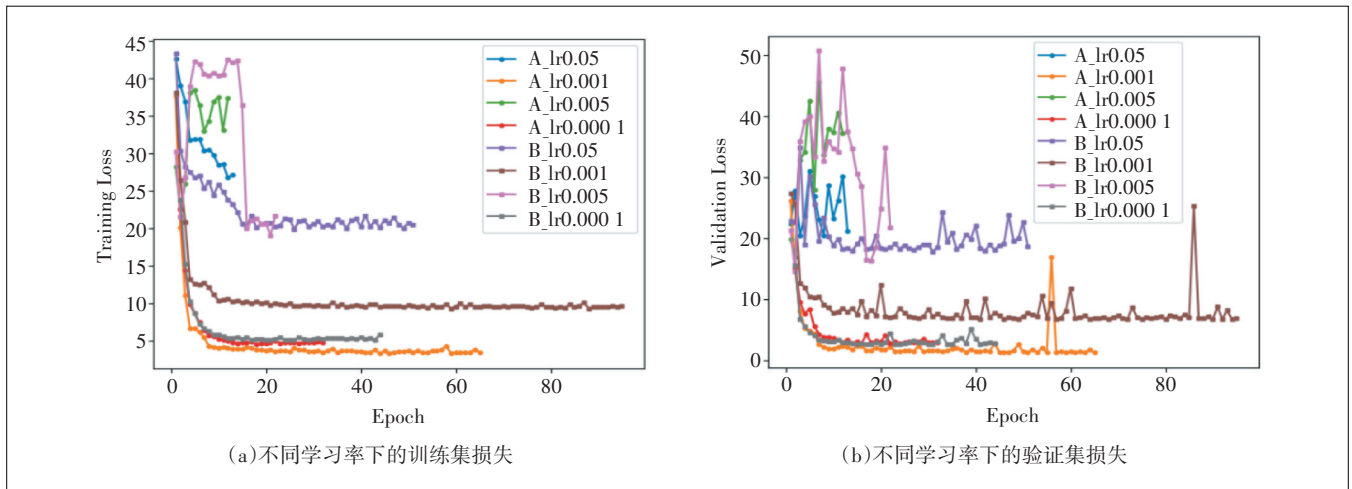


图8 8个工况下训练集和验证集的训练损失分布情况

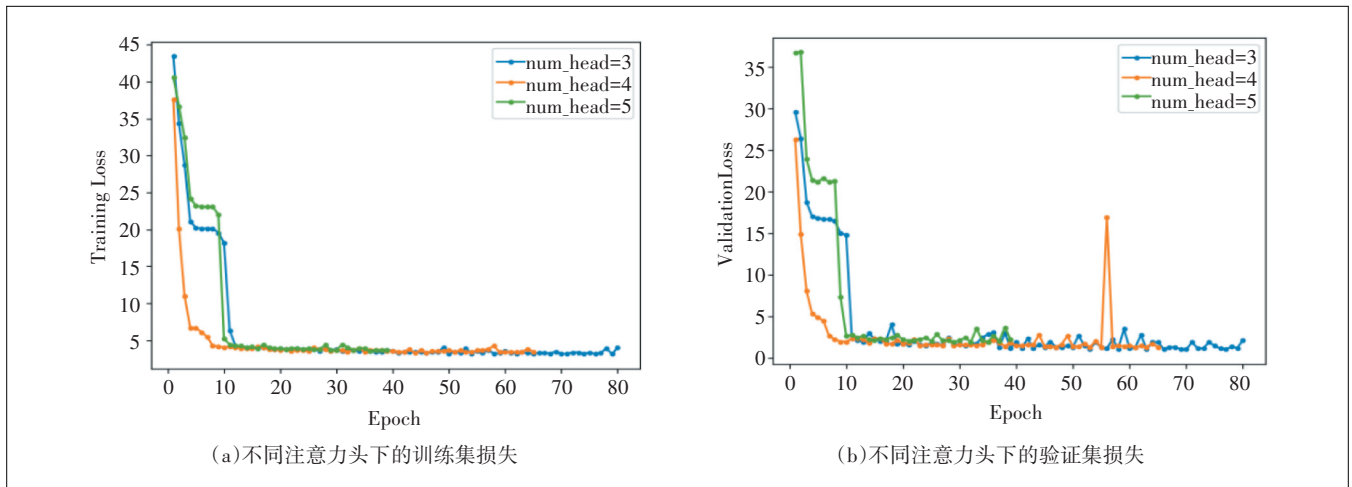


图9 不同注意力头下训练集损失和验证集损失情况

表1 测试验证工况

工况编号	机柜功率/kW	空调温度/°C
工况1	3.2	19
工况2	2	23
工况3	5.5	18

差为1.9%,最大处误差为3.5%,整场温度均方误差为0.65;工况2的整场平均绝对百分比误差为1.6%,最大处误差为3.21%,整场温度均方误差为0.56;工况3的整场平均绝对百分比误差为2.2%,最大处误差为3.50%,整场温度均方误差为0.76。值得关注的是,工况3是训练数据集中的训练工况,其GME为2.2%,工况1与工况2为不在训练数据集中的泛化工况,但GME分别为1.9%与1.6%,构建的三维温度场预测模

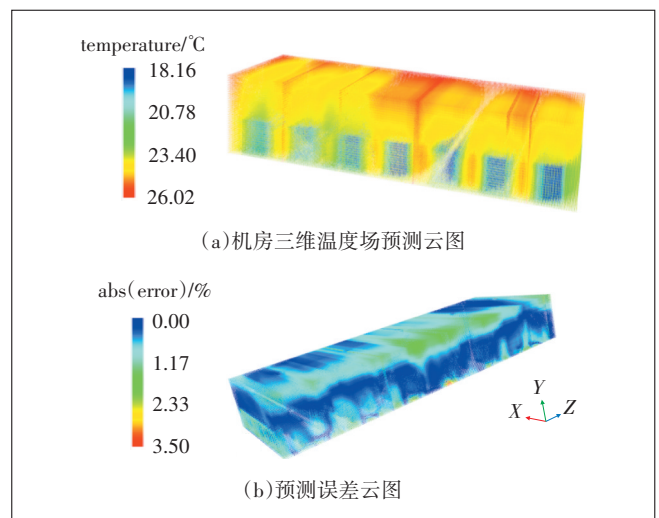


图10 工况1预测结果示意

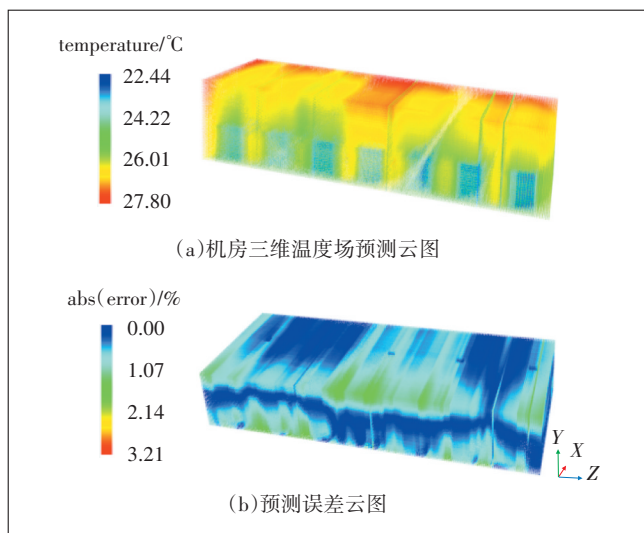


图11 工况2预测结果示意

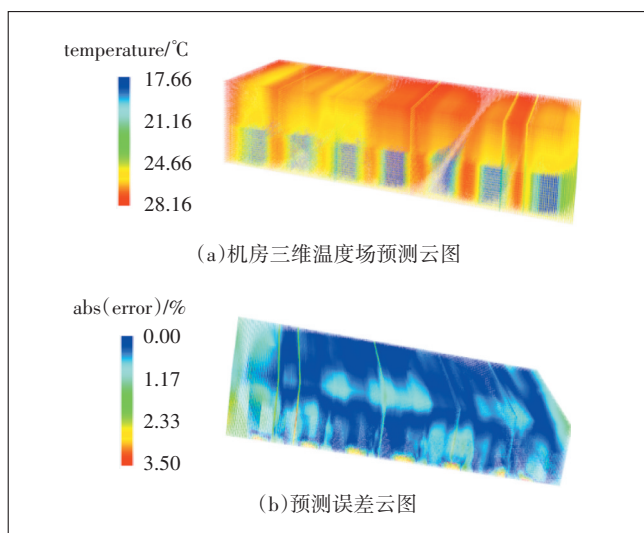


图12 工况3预测结果示意

型具有较好的泛化性能和鲁棒性。

4 结论

本文提出一种基于多特征输入Transformer架构的数据中心机房三维温度场快速重构和预测方法,实现分秒级预测性能,无需对空间布局相同的机房进行划分网格、仿真计算等长时间计算。本文研究了学习率、自注意力头等相关参数对模型预测性能的影响,根据实际需求选取最优参数,提出的方法可将全局平均预测精度误差控制在5%以内,具有较高的预测精度,并具有较好的工况泛化能力和鲁棒性。数据中心机房AI-CFD技术可以与AI节能调优等智能决策网络

相结合,支撑实现数据中心的节能、降碳与精细化运营。

参考文献:

- [1] 方正梁. 加速数据中心低碳绿色高效发展[N]. 人民邮电, 2023-04-12(3).
- [2] 佚名. 电算协同发展 机遇挑战并存(圆桌对话)[N]. 中国能源报, 2024-09-02(3).
- [3] 董梓童. 数据中心绿色化有了新标准[N]. 中国能源报, 2024-07-08(10).
- [4] 工业和信息化部. 工业和信息化部关于印发《新型数据中心发展三年行动计划(2021-2023年)》的通知[EB/OL]. [2024-08-25]. https://www.miit.gov.cn/jgsj/txs/wjfb/art/2021/art_12cc04dc9daf4d57a7038811a57383b6.html.
- [5] 工业和信息化部, 国家发展改革委, 财政部, 等. 工业和信息化部等七部门关于印发信息通信行业绿色低碳发展行动计划(2022-2025年)的通知[EB/OL]. [2024-07-14]. https://www.miit.gov.cn/jgsj/txs/wjfb/art/2022/art_f843aa7e249445bbbf3e6796f6ab7236.html.
- [6] ZHANG X, LINDBERG T, XIONG N, et al. Cooling energy consumption investigation of data center IT room with vertical placed server[J]. Energy Procedia, 2017, 105: 2047-2052.
- [7] 中华人民共和国住房和城乡建设部, 中华人民共和国国家质量监督检验检疫总局. 数据中心设计规范: GB 50174-2017[S]. 北京: 中国计划出版社, 2017.
- [8] PRIYADHARSHINI P, DIVYA P. Machine learning design thinking for fluid models[J]. Journal of physics: Conference series, 2021, 1947(1): 012056.
- [9] BRUNTON S L, NOACK B R, KOUMOUTSAKOS P. Machine learning for fluid mechanics[J]. Annual Review of Fluid Mechanics, 2020(52): 477-508.
- [10] USMAN A, RAFIQ M, SAEED M, et al. Machine learning computational fluid dynamics[C]//2021 Swedish Artificial Intelligence Society Workshop (SAIS). Manhattan: IEEE, 2021: 1-4.
- [11] 贺晓, 许俊, 胡孝俊, 等. 数据中心机房设计运行全过程数字孪生关键技术及体系架构[J]. 邮电设计技术, 2022(12): 9-13.
- [12] 许俊, 程序, 胡孝俊, 等. 基于人工智能算法的数据中心机房气流组织温度预测研究[J]. 邮电设计技术, 2022(12): 52-56.
- [13] 许俊, 胡孝俊, 高健, 等. 基于机器学习方法的三维气流组织温度场重构[J]. 科技资讯, 2024, 22(6): 227-230.

作者简介:

朱旭, 助理研究员, 博士, 主要从事数据中心与AI前沿交叉研究工作; 贺晓, 正高级工程师, 硕士, 主要从事数据中心研究咨询管理工作; 高健, 工程师, 硕士, 主要从事数据中心数字孪生、AI-CFD研发工作; 闫若飞, 助理工程师, 硕士, 主要从事数据中心相关AI算法研发工作; 陈俊丞, 工程师, 硕士, 主要从事数据中心相关AI算法研发工作; 姚贵策, 副教授, 博士, 主要从事智慧能源低碳节能管理技术研究工作。