

基于大数据分析的分布式光伏电站运行状态预测与优化调度研究

邱乾胜 肖凌超 张宏候

浙江芯能光伏科技股份有限公司, 中国·浙江 海宁 314000

摘要: 随着分布式光伏电站在能源结构转型中的重要性日益凸显, 其运行状态预测与优化调度成为研究热点。基于大数据技术构建的数据采集、存储与分析体系, 为光伏电站智能化运维提供了坚实基础。通过传统统计模型、机器学习方法及混合模型的深度融合, 实现了对发电量的精准预测。在此基础上, 能量管理系统整合需求响应机制, 结合智能调度算法和多目标优化方法, 建立了一套完整的优化调度策略, 有效提升了系统运行效率, 为分布式光伏电站的智能化发展提供了技术支持。

关键词: 分布式光伏电站; 大数据分析; 运行状态预测; 优化调度; 能量管理系统

Research on Operation Status Prediction and Optimization Scheduling of Distributed Photovoltaic Power Plants Based on Big Data Analysis

Qiansheng Qiu Lingchao Xiao Honghou Zhang

Zhejiang Xinneng Photovoltaic Technology Co., Ltd., Haining, Zhejiang, 314000, China

Abstract: With the increasing importance of distributed photovoltaic power plants in energy structure transformation, their operational status prediction and optimization scheduling have become research hotspots. The data collection, storage, and analysis system based on big data technology provides a solid foundation for the intelligent operation and maintenance of photovoltaic power plants. Through the deep integration of traditional statistical models, machine learning methods, and hybrid models, accurate prediction of power generation has been achieved. On this basis, the energy management system integrates the demand response mechanism, combines intelligent scheduling algorithms and multi-objective optimization methods, and establishes a complete optimization scheduling strategy, effectively improving the system's operational efficiency and providing technical support for the intelligent development of distributed photovoltaic power plants.

Keywords: distributed photovoltaic power station; big data analysis; prediction of operational status; optimize scheduling; energy management system

0 前言

随着全球能源需求的增长和环境问题的加剧, 各国都在积极推动清洁、可再生能源的发展。光伏发电作为其中的关键力量, 其分布式应用对于提高能源效率和减少碳排放具有显著作用。然而, 分布式光伏电站在实际运营中遇到了不少挑战, 包括发电不稳定、设备故障频发以及调度难度大等问题。幸运的是, 大数据技术的发展为这些挑战提供了解决方案。通过收集、存储、分析和挖掘大量运行数据, 可以精准预测光伏电站的运行状态, 并优化调度策略, 从而提高电站的运行效率和经济效益。

1 分布式光伏电站及其运行特点

分布式光伏电站作为新型清洁能源系统的重要组成部分, 在能源结构转型和可持续发展中发挥着关键作用。按照技术定义, 分布式光伏电站是指安装在用户侧或靠近用户

侧的小型光伏发电系统, 具有分散布局、就近消纳的特点。与传统集中式光伏电站相比, 分布式光伏电站在建设周期、投资规模和并网要求等方面具有显著优势: 建设周期较短可快速投入使用, 投资规模灵活便于分期建设, 对电网冲击小易于并网接入, 这些特点使其在城乡区域得到广泛应用。在运行特征方面, 分布式光伏电站呈现出明显的差异性和复杂性^[1]。发电量受光照强度、环境温度等自然因素影响显著, 导致输出功率存在较大波动和不确定性; 系统设备分布范围广泛, 各个光伏电站之间需要协调调度和统一管理, 增加了运维难度; 不同厂家生产的光伏组件、逆变器等核心设备在性能参数和故障特征上存在差异, 给设备管理和维护带来挑战; 系统运行过程中配置有大量的传感器和监控设备, 实时采集电压、电流、温度等运行参数, 产生海量运行数据。基于分布式光伏电站的技术特点和运行特征, 建立健全的运行管理体系显得尤为重要。

2 大数据技术在分布式光伏电站中的应用

大数据技术的发展为分布式光伏电站的智能化运维提供了强大的技术支撑,在数据采集、存储、处理及分析等环节发挥重要作用。分布式光伏电站运行过程中产生的实时数据种类繁多,包括电压、电流、温度、辐照度等基础运行参数,以及设备状态、故障信息等衍生数据,这些数据通过物联网设备进行采集并存储于分布式数据库中,形成完整的数据采集与存储体系。面对采集到的海量运行数据,建立规范的数据处理流程至关重要,因原始数据普遍存在噪声干扰、数据缺失、异常值等质量问题,需要借助专业的大数据处理框架进行系统化处理。Hadoop、Spark 等分布式计算平台为大规模数据的清洗、过滤和补全提供了高效的技术支持,通过并行计算和分布式存储显著提升了数据处理效率^[2]。在数据分析与挖掘环节,通过对处理后的高质量数据进行深入分析,可挖掘出潜在的规律和关联关系。关联规则分析有助于发现不同运行参数之间的内在联系,聚类分析能够识别相似的运行模式和故障特征,分类分析则可用于设备状态评估和故障诊断。这些数据挖掘技术的综合应用使光伏电站运行效率的关键影响因素得以准确识别,为系统优化和故障预警奠定了坚实的数据基础。

3 运行状态预测方法

3.1 传统统计模型

传统统计模型在分布式光伏电站发电量预测领域占据重要地位,其中以时间序列分析方法最为典型且应用广泛。在众多统计模型中,ARIMA(自回归积分移动平均)模型通过对历史数据进行差分处理实现数据平稳化,继而建立数学模型进行预测分析,在短期发电量预测中表现出较高的准确性。基于回归分析的预测方法则从变量间关系的角度出发,通过建立自变量与因变量之间的数学关系进行预测,如线性回归能够反映光照强度与发电量之间的线性相关性,多元回归则可综合考虑温度、风速等多个影响因素。这类传统统计模型凭借其扎实的数学理论基础和相对简单的计算过程,在实际应用中具有较强的可解释性和易实现性。然而,由于光伏发电系统受多种环境因素的复杂影响,发电特性往往呈现出明显的非线性特征,而传统统计模型在处理这类非线性关系时存在明显局限性:当系统输入与输出之间的关系复杂多变时,简单的线性模型难以准确描述系统特性;在面对突发天气变化等异常情况时,预测精度会显著降低,这些不足促使研究者探索更为先进的预测方法。

3.2 机器学习方法

机器学习算法凭借其强大的数据处理能力和优异的预测性能,在光伏发电预测领域展现出广阔的应用前景。支持向量机(SVM)作为一种经典的机器学习算法,通过核函数将低维数据映射到高维特征空间,在处理小样本、高维数据时表现出色,且具备较强的泛化能力,能有效避免过拟合

问题。人工神经网络(ANN)则依托其独特的网络结构,通过多层神经元之间的复杂连接模拟人脑的信息处理机制,在捕捉光伏发电系统中复杂的非线性关系方面展现出显著优势,特别适合处理大规模历史数据的建模分析。在实际应用中,基于决策树的集成学习方法如随机森林、梯度提升树凭借其良好的可解释性和较高的预测精度受到广泛关注,这类算法通过组合多个基学习器的预测结果,有效提升了模型的泛化性能和鲁棒性^[3]。随着深度学习技术的快速发展,长短期记忆网络(LSTM)和卷积神经网络(CNN)等先进算法在光伏发电预测中得到深入应用:LSTM 凭借其特殊的门控机制能够有效处理长序列时间数据,捕捉发电量的时序特征和长期依赖关系;CNN 则通过卷积运算提取数据的空间特征,在处理气象图像等空间数据时表现突出。

3.3 混合模型与集成学习

混合模型和集成学习方法作为光伏发电预测领域的前沿研究方向,通过有机融合不同算法的优势特点实现了预测性能的显著提升。在混合模型的构建过程中,传统统计方法可用于处理数据的线性特征,而机器学习算法则负责捕捉非线性成分,两者优势互补形成了更为完善的预测体系。基于集成学习的预测方法通过构建多个基础模型并采用不同的组合策略进行集成,有效降低了单一模型的局限性:通过 Bagging 策略可降低预测方差,提高模型稳定性;运用 Boosting 方法则着重于减小预测偏差,提升模型精度;Stacking 技术通过多层模型堆叠实现了更为复杂的非线性映射关系。这些集成策略的灵活运用使预测模型在面对复杂多变的光照条件和天气因素时表现出更强的适应性。在实际应用中,混合模型和集成学习方法通过合理分配不同算法的权重,充分发挥各类预测方法的优势,在保持较高预测精度的同时也确保了模型的泛化能力和鲁棒性,为分布式光伏电站的精准预测和智能调度提供了可靠的技术支撑。

图 1 为该章节简要说明图。

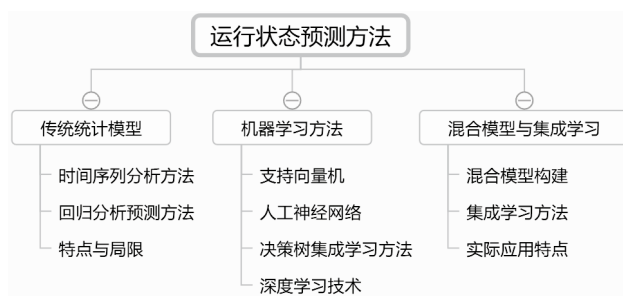


图 1 运行状态预测方法简要说明图

4 优化调度策略与方法

4.1 能量管理系统

能量管理系统(EMS)作为分布式光伏电站智能化运行的核心控制平台,通过构建完整的监控体系和优化决策机制,实现了对发电系统的全方位管理。在系统架构设计上,

EMS 集成了数据采集、状态监测、优化分析和控制执行等多个功能模块，形成了闭环的管理体系。通过实时采集光伏组件输出功率、电网负荷需求、储能设备状态等关键运行参数，EMS 建立了系统运行状态的实时监控机制，为优化决策提供数据支撑。在此基础上，结合大数据分析技术对采集的历史数据进行深度挖掘，发现系统运行规律并预测未来发电和用电趋势，进而制定科学合理的调度策略^[4]。在实际运行过程中，EMS 通过灵活调节光伏发电、储能充放电和负荷用电，实现了系统供需平衡的动态优化：在光照充足时优先消纳光伏发电，多余电量存入储能系统；在用电高峰期则通过释放储能电量满足负荷需求，确保用电可靠性。这种基于预测的智能调度机制不仅提高了可再生能源的消纳比例，也显著改善了系统的经济性和稳定性。

4.2 需求响应

需求响应作为智能电网中的重要调节手段，通过引导和激励用户主动参与电力系统调节，在平衡供需关系、提升系统效率方面发挥着关键作用。基于价格信号和激励机制，需求响应能够有效引导用户改变用电行为，将部分可调节负荷从用电高峰期转移至低谷期，从而实现电力需求的时空转移。在分布式光伏系统中，需求响应与光伏发电预测的深度融合开创了用电侧管理的新模式：通过对光伏发电量的精准预测，系统可提前规划负荷调节方案，在光照充足、发电量较大时鼓励用户增加用电负荷；而在光照不足或阴雨天气条件下，则通过需求响应机制引导用户减少非必要用电，保障系统供需平衡。这种基于预测的需求响应策略通过建立发电端与用电端的协调互动机制，显著提升了分布式光伏系统的运行灵活性：负荷可根据发电预测结果进行提前调整，避免因光伏出力波动导致的供需失衡；同时，通过合理设计需求响应补偿机制，调动用户参与积极性，形成良性的互动格局。

4.3 智能调度算法

智能调度算法作为分布式光伏电站运行优化的关键技术，通过融合大数据分析和人工智能技术实现了调度效率的显著提升。在众多智能优化算法中，遗传算法（GA）基于生物进化理论构建了一套完整的优化框架：通过编码、

选择、交叉和变异等遗传算子模拟种群进化过程，在解空间中不断探索和优化，最终获得满足约束条件的最优调度方案。粒子群优化算法（PSO）则借鉴了自然界中生物群体的智能特性，利用群体中个体之间的协作关系指导搜索方向：每个粒子通过更新自身位置和速度，在个体最优解和群体最优解的共同作用下快速逼近全局最优解，体现出收敛速度快、实现简单等优势特点^[5]。蚁群算法（ACO）通过对蚂蚁觅食行为的深入研究，建立了基于信息素的路径选择机制：在搜索过程中通过信息素的积累和挥发动态更新最优路径，有效解决了复杂的组合优化问题。强化学习则从智能体与环境交互的角度出发，构建了一种自适应的调度优化方法：通过定义合理的状态空间、动作空间和奖励函数，系统能够在持续交互过程中不断积累经验并改进决策策略，逐步形成适应环境变化的智能调度方案。

4.4 多目标优化

多目标优化在分布式光伏电站调度系统中具有重要应用价值，通过建立科学完备的目标评价体系和优化权衡机制，实现了系统运行的全局最优。在实际运行过程中，光伏电站调度面临着经济性、可靠性和环保性等多维度的优化目标：经济性包括运维成本最小化、发电效益最大化等经济指标，可靠性涵盖电压质量控制、功率波动抑制等技术要求，环保性则体现在碳排放减少和清洁能源消纳等环境效益方面。针对这些存在矛盾和制约关系的优化目标，多目标优化方法采用科学的数学建模手段将各项指标统一到可比较的评价框架中：通过构建合理的目标权重体系反映不同指标的重要程度，运用规范化处理消除量纲差异影响，建立综合评价函数作为优化求解的依据。在具体实施过程中，基于帕累托最优理论分析目标之间的折中关系，采用多目标进化算法搜索非支配解集，最终结合决策者偏好选择最佳的调度方案。这种多目标协同优化的方法不仅确保了经济效益、系统稳定性和环境效益的协调发展，也为分布式光伏电站的智能调度提供了更具实用价值的解决方案。通过建立完整的多目标优化框架，分布式光伏电站的运行效率和综合效益得到显著提升。

图 2 为该章节思维导图。

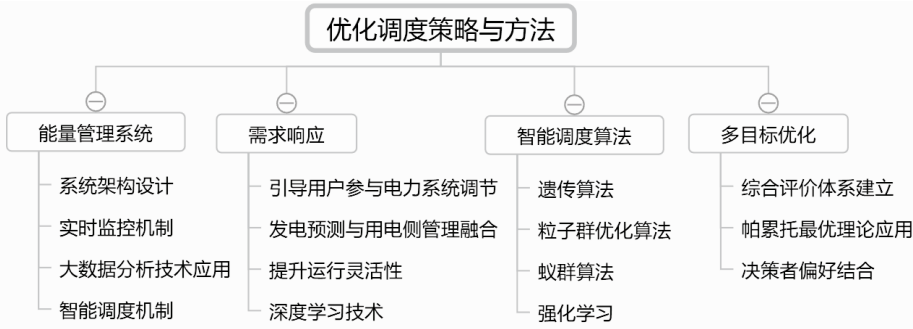


图 2 优化调度策略与方法思维导图

5 结语

大数据技术与分布式光伏电站的深度融合开创了智能化运维新局面。在数据采集与分析层面，建立了完整的数据处理体系，实现了海量运行数据的高效管理与深度挖掘。预测方法的创新发展体现在三个维度：传统统计模型奠定了数学基础，机器学习方法增强了非线性特征提取能力，混合模型与集成学习进一步提升了预测精度。优化调度策略通过能量管理系统实现全方位监控，需求响应机制促进了供需平衡，智能调度算法保障了运行效率，多目标优化方法则实现了经济性、可靠性和环保性的协同发展。这些技术创新与实践探索显著提升了分布式光伏电站的运行水平，推动了新能源产业的健康发展。

参考文献：

[1] 吴世健,廖强明,邱建东.基于大数据分析的分布式光伏预测与控制策略[J].电器工业,2024(6):45-49.

[2] 刘洋,刘吉成.大数据联盟下一种分布式光伏资源共享算法设计[J].沈阳工业大学学报,2024,46(1):109-114.

[3] 鲜恒超.分布式光伏发电系统的状态监测及故障诊断[J].通讯世界,2024,31(6):97-99.

[4] 尹国龙.基于分布式光伏发电量预测分析的运行优化策略研究[J].电测与仪表,2021(3):21-22.

[5] 郑芳,王欢,陈国平,等.基于大数据的分布式光伏接入配电网影响与功率预测研究[J].光源与照明,2023(8):138-140.

作者简介：邱乾胜（1990-），男，硕士，高级工程师。