

基于标准组串模型的光伏电站 智能运维系统研究

翁得汉¹, 李宪章¹, 束秋节², 林颖若³

(1. 华润电力投资有限公司华南分公司, 广东 广州 511455;
2. 华润电力控股有限公司, 广东 广州 511455;
3. 华润(清远)太阳能发电有限公司, 广东 清远 511500)

摘要 传统光伏监控系统多依赖阈值告警, 难以精准识别低效运行单元, 导致运维滞后、发电损失。本文提出一种基于标准组串模型的智能运维系统, 通过构建性能基准组串, 结合离散率分析与机器学习算法, 实现对异常组串的自动识别、定位与诊断。系统采用“云一边一端”协同架构, 在边缘侧完成实时数据清洗与初步判别, 云端进行大数据建模与趋势预测。实际应用于多个百兆瓦级光伏场站的结果表明, 该方法可将低效组串识别准确率提升至 92% 以上, 年均发电量提升 1.8% ~ 3.2%, 显著提高运维效率与经济收益。

关键词 光伏电站; 标准组串; 离散率分析; 智能运维; 云边协同

中图分类号: TM61; TM76

文献标志码: A

DOI: 10.3969/j.issn.2097-3365.2025.36.009

0 引言

光伏发电作为清洁能源的核心支柱, 其装机容量在全球范围内持续高速增长。截至 2024 年底, 我国光伏发电累计装机容量已突破 10 亿千瓦, 占全国发电装机总量的比重超过 35%。然而, 在光伏电站规模化发展的同时, 大型光伏场站普遍面临设备数量庞大(单百兆瓦级电站组串数量可达数万组)、故障隐蔽性强等突出问题, 传统“被动响应式”运维模式已难以满足高效、精益化的管理需求。现有光伏监控系统虽能实现逆变器停机、通信中断等严重故障的告警功能, 但对于组件衰减、局部遮挡、接线松动、灰尘覆盖等导致的“亚健康”运行状态缺乏有效的评估手段。

1 标准组串模型构建

1.1 模型定义与筛选机制

标准组串并非固定不变的特定组串, 而是在同一光伏方阵内, 满足特定条件的动态组串集合。其核心筛选条件包括四个维度: 一是安装条件一致性, 要求组串的安装朝向、倾角完全一致, 确保接收的太阳辐射条件相同; 二是物理状态完好性, 无明显的物理遮挡、组件污渍、玻璃破损等问题; 三是运行历史清洁性, 近 3 个月内无故障记录及维修历史; 四是性能指标优越性, 其归一化发电效率需处于全场组串的前 20% 分位。

为实现标准组串的动态筛选与更新, 本文采用 K-

means 聚类算法对历史发电数据进行处理。具体流程如下: 首先提取同一方阵内所有组串连续 30 天的归一化功率数据; 然后以日均发电效率、功率波动系数为特征变量, 通过 K-means 算法将组串划分为高效稳定、中效波动、低效异常三个聚类簇; 最后从高效稳定簇中筛选出满足物理状态条件的组串, 构成初始标准组串集合。

通过该筛选机制, 标准组串集合的发电性能稳定性得到有效保障。实验数据显示, 同一标准组串集合内各成员的归一化功率离散系数始终低于 3%, 远低于普通组串集合 15% ~ 20% 的离散水平, 为后续的异常诊断提供了可靠的基准依据。

1.2 环境归一化处理方法

光伏组串的输出功率受辐照度、组件温度等环境因素影响显著, 相同运行状态的组串在不同环境条件下输出功率差异可达数倍。为消除环境因素干扰, 精准对比组串的固有性能, 需对实测功率数据进行环境归一化处理。本文采用的归一化公式如下:

$$P_{norm} = \frac{P_{meas}}{G_{meas}} \times G_{ref} \times [1 + \alpha(T_{cell} - T_{ref})] \quad (1)$$

式(1)中, P_{meas} 为组串实测功率(kW); G_{meas} 为实测太阳辐照度(W/m²); G_{ref} 为标准辐照度, 取值 1 000 W/m²; T_{cell} 为组件实际温度(℃); T_{ref} 为标准温度, 取值 25 ℃。

为确保归一化处理的准确性,系统在每个光伏方阵设置1组高精度环境监测单元,同步采集辐照度、环境温度数据,并通过组件温度估算模型(结合辐照度^[1]、环境温度、风速及组件运行时间)计算组件实际温度 T_{cell} ,替代传统的直接测量方式,降低了硬件部署成本,同时提升了温度数据的覆盖范围。实验验证表明,该估算模型的温度误差可控制在 ± 2 °C以内,对归一化结果的影响小于1%,满足工程应用精度要求。

2 离散率分析与异常诊断

2.1 基于标准组串的离散率计算

离散率分析是识别低效组串的核心手段,其核心思想是通过计算待测组串与标准组串集合的功率差异程度,判断其是否偏离正常运行范围。对于每个待测组串*i*,其与标准组串集合的功率离散率 D_i 计算方式如下:

$$D_i = \frac{1}{N} \sum_{j=1}^N \left| \frac{P_{norm,i} - P_{norm,j}}{P_{norm,j}} \right| \quad (2)$$

式(2)中, N 为标准组串集合的组串数量(通常取值20~30组,确保统计显著性); $P_{norm,i}$ 为待测组串*i*的归一化功率; $P_{norm,j}$ 为标准组串*j*的归一化功率。

离散率 D_i 反映了待测组串相对于标准基准的偏离程度, D_i 越大,说明该组串的运行状态越异常。系统设定动态阈值delta(通常取值8%~12%),当 $D_i > \text{delta}$ 时,判定该组串为潜在低效组串^[2]。阈值delta的动态调整机制基于季节变化与方阵老化情况,夏季因辐照度稳定,阈值取下限8%;冬季因辐照度波动大,阈值取上限12%;同时,随着方阵运行年限增加,每年将阈值上调0.5%,以适应组件自然衰减的影响。

2.2 多维特征融合的故障诊断模型

离散率分析可实现低效组串的初步识别,但无法区分具体的故障类型。为提升诊断的精细化水平,本文构建了基于随机森林的多分类模型,结合多维电气特征实现故障类型的精准判别。

模型的输入特征包括三类:一是基础性能特征,如离散率 D_i 、归一化功率均值、功率波动标准差;二是电气参数特征,如组串电流一致性(与标准组串电流的差异率)、电压波动系数、早中晚功率比值^[3];三是IV曲线特征,通过逆变器内置算法提取的曲线斜率、填充因子、开路电压偏差等特征参数。共选取12个维度的特征变量,构成故障诊断的特征向量。

模型的训练数据集来源于多个光伏场站的历史故障记录,涵盖遮挡、热斑、二极管失效、接线松动、

组件衰减等5类典型故障,共计12 000组样本数据(其中正常样本6 000组,故障样本6 000组)。通过随机森林算法构建多分类模型,采用5折交叉验证进行模型优化,最终模型的故障诊断准确率达到89.5%,其中对遮挡、热斑等显性故障的识别准确率超过95%,对接线松动等隐性故障的识别准确率也达到82%以上,为针对性的运维提供了精准依据。

3 基于标准组串模型的光伏电站智能运维系统架构设计与实现

3.1 “云一边一端”协同架构设计

为平衡实时性与扩展性需求,系统采用“端一边一云”三级协同架构,将数据采集、分析计算、决策管理等功能进行合理分配,具体架构如下:

1. 端侧:部署组串级智能采集终端,采用LoRa无线通信技术(通信距离可达3 km,功耗低至5 mA),实时采集每组串的电流、电压、功率数据,采样频率为1分钟/次,并将数据上传至边缘节点。同时,端侧设备具备本地数据缓存功能,可在通信中断时缓存7天内的数据,保障数据完整性。

2. 边缘侧:每个光伏方阵部署1台边缘计算网关,主要承担三项功能:一是数据预处理,对端侧上传的数据进行清洗,剔除因通信干扰导致的异常值(采用 3σ 准则);二是实时分析,执行环境归一化处理与离散率计算,完成低效组串的初步筛选,将识别出的潜在异常数据上传至云端,正常数据仅进行本地存储(存储周期为30天)^[4],降低云端数据传输与存储压力;三是本地响应,对于离散率超过20%的严重异常组串,直接触发本地声光告警,确保运维人员可快速响应紧急故障。

3. 云端:基于阿里云服务器构建云端管理平台,核心功能包括:一是标准组串库管理,定期执行K-means聚类算法,更新标准组串集合;二是深度诊断,运行随机森林故障诊断模型,对边缘侧上传的异常数据进行故障类型判别;三是趋势预测,基于历史数据构建ARIMA时间序列模型,预测组串性能变化趋势,实现从“被动维修”向“主动预警”的转变;四是运维管理,自动生成运维工单,通过APP推送至运维人员手机端,并跟踪工单执行进度,形成运维闭环。

3.2 系统实现与部署情况

系统的硬件部分采用国产化元器件,组串采集终端成本控制在150元/台以内,边缘计算网关成本约5 000元/台,相比进口同类设备成本降低40%以上。

软件部分基于 Python 语言开发，采用 Django 框架构建云端管理平台，前端采用 Vue.js 实现可视化展示，支持组串运行状态实时监控、故障告警、数据统计分析等功能。

目前，该系统已在中车集团下属的 5 个百兆瓦级光伏电站完成部署，总接入容量达 520 MW，覆盖组串数量超过 45 000 组。系统运行稳定，数据采集成功率达 99.2%，边缘侧分析延迟小于 5 秒，云端诊断响应时间小于 30 秒，完全满足大型光伏电站的运维需求^[5]。

4 基于标准组串模型的光伏电站智能运维系统应用效果分析

为验证系统的实际应用效果，选取某 100 MW 地面光伏电站作为典型案例，对比系统部署前后 6 个月的运维与发电数据。该电站位于内蒙古鄂尔多斯市，装机容量 100 MW，共包含 416 个光伏方阵，8 640 组光伏组串，年均辐照度约 1 700 kW·h/m²，属于优质光伏资源区。

对比分析结果显示，系统部署后运维与发电指标均得到显著提升：一是低效组串识别效率大幅提升，部署前依赖人工巡检，低效组串平均识别时间为 15 天，部署后通过系统自动识别，平均识别时间缩短至 2 小时，识别效率提升 180 倍；二是运维成本显著降低，运维人员数量由原来的 12 人减少至 8 人，人力成本下降 35%，同时因故障处理及时，维修材料损耗减少 20%；三是发电效益明显增长，电站年等效利用小时数由原来的 1 580 小时提升至 1 603 小时，提升 23 小时，折合增发电量约 2.1%，按当地上网电价 0.32 元/度计算，每年可增加经济收益约 71.68 万元。

同时，对 5 个已部署场站的整体数据统计显示，系统对低效组串的识别准确率平均达 92.3%，最高可达 94.1%，不同场站的年均发电量提升幅度在 1.8%~3.2% 之间，其中资源条件较差、运维基础薄弱的场站提升效果更为显著，增幅均超过 3%，充分验证了系统的普适性与有效性。客户反馈表明，系统的应用使运维工作从“事后补救”转变为“事前预警”，运维主动性与精细化水平得到质的提升。

5 结论与展望

5.1 研究结论

本文提出的基于标准组串模型的光伏电站智能运维系统，通过构建动态更新的标准组串集合作为性能基准，结合环境归一化处理与离散率分析，实现了对低效组串的精准识别；进一步融合多维电气特征与随

机森林算法，完成了故障类型的精细化诊断；“云一边一端”协同架构的应用，既保障了实时数据处理能力，又实现了大数据分析与趋势预测功能。

工程实践验证表明，该系统具有以下优势：一是识别精度高，低效组串识别准确率超过 92%，故障诊断准确率达 89.5%；二是运维效率高，将低效组串识别时间从数天级缩短至小时级；三是经济收益显著，可使光伏电站年均发电量提升 1.8%~3.2%，同时降低 35% 的运维人力成本；四是成本可控，采用国产化硬件与轻量化算法，部署成本低于同类系统 40%，具备大规模推广应用的条件。该系统有效解决了传统运维模式的痛点，为光伏电站的精益化管理提供了可靠的技术支撑。

5.2 未来展望

尽管系统已实现了较好的应用效果，但仍有进一步优化的空间。未来将从三个方向开展深入研究：一是融合气象预测数据，结合数值天气预报（NWP）的辐照度、温度预测结果，构建更精准的性能预测模型，实现故障的提前预警（预测时间窗口 1~3 天）；二是引入清洗优化策略，基于组串积灰程度评估（通过功率衰减趋势分析）与降水预测数据，制定最优的组件清洗计划，避免无效清洗作业；三是拓展多源数据融合，结合无人机巡检的图像数据与红外热成像数据，弥补电气参数分析在物理故障识别上的不足，构建“电气参数+图像特征”的多维度诊断模型，进一步提升故障识别的全面性与准确性。随着技术的不断完善，有望推动光伏运维从“预测性维护”向“主动性优化”发展，为光伏产业的高质量发展提供更强有力的技术保障。

参考文献：

- [1] 刘星.光伏电站运维中的智能监测与故障诊断技术应用[J].电气工程与自动化,2025,04(01):85-88.
- [2] 刘东,李田泽,刘开石,等.GA-BP 神经网络在光伏阵列故障检测中的应用研究[J].电源技术,2024(03):24-26.
- [3] 刘开石,李田泽,刘东,等.基于 ABC-SVM 算法的光伏阵列故障诊断[J].电源技术,2025(09):20-22.
- [4] 陈世群,高伟,陈孝琪,等.一种基于极限学习机和皮尔逊相关系数的光伏阵列故障快速诊断方法[J].电气技术,2024(10):10-12.
- [5] 马纪梅,张志耀,张启然.基于改进 RBF 神经网络的光伏组件故障诊断[J].电测与仪表,2024(02):19-21.