

风电机组数据采集与监控系统异常数据识别方法研究

李忠义¹, 赵长江¹, 相童巍¹, 王乐¹, 默婧²

1. 华能新能源股份有限公司河北分公司, 河北 石家庄 050011

2. 石家庄机动车排污管理中心, 河北 石家庄 050000

DOI:10.61369/EPTSM.2026020010

摘 要 : 随着风电规模化发展, 风电机组数据采集与监控系统采集的海量运行数据中混杂的异常数据, 严重影响设备状态评估与故障预警的准确性。本文基于华能新能源股份有限公司生产实践, 系统分析了风电场 SCADA 系统中异常数据的典型类型及其产生机理, 重点研究了基于统计阈值、规则逻辑与机器学习的三类异常识别方法, 并提出了一种分层融合的工程化应用框架。实际应用表明, 该框架能够有效识别传感器故障、通信异常及设备早期劣化等多种异常模式, 为提升风电场智能化运维水平提供了技术支撑。

关 键 词 : 风电机组; 数据采集与监控系统; 异常数据识别; 智能运维

Research on Abnormal Data Identification Methods for Wind Turbine Data Acquisition and Monitoring Systems

Li Zhongyi¹, Zhao Changjiang¹, Xiang Tongwei¹, Wang Le¹, Mo Jing²

1. Huadian New Energy Co., Ltd. Hebei Branch, Shijiazhuang, Hebei 050011

2. Shijiazhuang Motor Vehicle Emission Management Center, Shijiazhuang, Hebei 050000

Abstract : With the large-scale development of wind power, abnormal data mixed within the massive operational data collected by wind turbine data acquisition and monitoring systems severely impacts the accuracy of equipment condition assessment and fault early warning. Based on the production practices of Huaneng New Energy Co., Ltd., this paper systematically analyzes the typical types and generation mechanisms of abnormal data in wind farm SCADA systems. It focuses on three anomaly identification methods—statistical threshold-based, rule logic-based, and machine learning-based—and proposes a hierarchical fusion engineering application framework. Practical applications demonstrate that this framework can effectively identify various anomaly patterns such as sensor failures, communication anomalies, and early-stage equipment degradation, providing technical support for enhancing the intelligent operation and maintenance of wind farms.

Keywords : wind turbine; data acquisition and monitoring system; anomaly detection; intelligent operations

引言

风电机组数据采集与监控系统 (SCADA) 是风电场运行的“神经中枢”, 持续采集机组运行状态、环境参数及发电性能等数百个测点数据。然而, 在复杂运行环境及设备自身因素影响下, 采集数据中不可避免地包含各类异常值。这些异常数据不仅干扰运行人员对机组状态的正常监视, 更会严重影响基于数据驱动的性能分析、功率预测及故障诊断等高级应用的可靠性。如何从海量、高维、强噪声的 SCADA 数据流中, 准确、高效地识别并剔除异常数据, 已成为风电行业数字化、智能化转型过程中待解决的基础性关键问题。

本文立足于华能新能源股份有限公司多个风电场的实际运维需求与数据基础, 针对工程实践中常见的异常数据类型, 对当前主要的识别方法进行系统性梳理与技术经济性比较, 并结合具体案例, 探讨适用于规模化风电场群的、具备工程可行性的异常数据识别技术路线, 以期为行业同类问题的解决提供参考。

作者简介:

李忠义, 职称: 助理工程师, 研究方向: 新能源发电集控运行;

赵长江, 职称: 助理工程师, 研究方向: 新能源发电集控运行;

相童巍, 职称: 助理工程师, 研究方向: 电力人工智能;

王乐, 职称: 助理级工程师, 研究方向: 新能源发电运行;

默婧, 职称: 中级工程师, 研究方向: 机动车污染防治。

一、SCADA系统异常数据类型及产生机理分析

风电机组 SCADA 数据异常表现形式多样，其背后关联着不同的产生机理与潜在风险。从工程实践角度，可主要归纳为以下四类：

（一）点异常

表现为单个或连续少量数据点严重偏离参数合理物理范围或历史正常波动区间。其典型产生原因包括：传感器瞬时故障或受到强电磁干扰；数据传输过程中因信号衰减或协议错误导致的突变；控制系统指令异常引发的状态骤变。例如，风速仪瞬时输出超大风速值，或发电机转矩在毫秒级时间内跳变为零。这类异常易于通过阈值判断发现，但需注意阈值设定的合理性，避免对正常启停、穿越等工况的误判。

（二）上下文异常

指数据值本身在全局统计上看未超限，但在特定运行工况下违背物理规律或控制逻辑。

功率与风速失配：在风速持续高于额定风速（如 $>12\text{m/s}$ ）的工况下，变桨系统应调节桨距角以限制功率，此时风机功率应稳定在额定功率附近。若此时功率反而随风速增加而波动或下降，可能暗示变桨系统响应迟缓、存在卡涩或传感器反馈偏差。

热力学异常：在环境温度较低（如 -10°C ）且机组处于待机状态时，齿轮箱油温应与环境温度接近。若此时油温异常偏高（如 $>40^{\circ}\text{C}$ ），可能预示着冷却系统未关闭、内部摩擦增大或加热器粘连。

此类异常往往是设备性能下降的早期征兆，单纯依靠单点阈值无法识别，必须结合工况上下文进行分析。

（三）集体异常（模式异常）

这类异常最为隐蔽，也是智能运维的重点难点。其特征是单个参数序列可能无明显异常，但在时间维度上或多个相关参数组合所呈现的联合分布模式，偏离了历史正常机组或同类机组的标准模式。例如：

传动链早期磨损：主轴承振动频谱中特定频率分量（如故障特征频率）的能量并未超过报警线，但呈现出缓慢上升的趋势；同时伴随着齿轮箱油温波动特征的细微改变（如温升速率加快）。这种多参数协同演化的异常模式，往往是传动链部件出现早期磨损、不对中或微点蚀的征兆。

偏航系统迟滞：风向变化时，机舱角度的跟随存在系统性延迟，且这种延迟随时间推移逐渐增大。

集体异常通常预示着设备的早期劣化，是实现“治未病”式运维的关键切入点。

（四）系统性与通信异常

主要由数据采集硬件、网络传输链路故障引起，而非设备本体故障。其表现形式包括：

数据长时间缺失：采集卡通道损坏或光纤链路中断，导致某一测点长时间无数据更新。

常数输出：通信链路卡死或传感器死区，导致某一测点持续输出固定不变的值（如一直显示 25.5°C ），缺乏正常的波动。

大规模乱码：网络风暴或协议转换器故障，导致大规模测点数据同时出现跳变或无效值。

这类异常直接影响 SCADA 系统的可用性，通常需要通过系统健康自诊断与通讯质量监测来发现，是保障数据完整性的第一道防线。

二、异常数据识别的主要方法

针对上述不同类型的异常，目前已发展出多种识别方法。在工程实践中，没有一种万能的方法，通常需要根据异常类型、实时性要求及计算资源进行综合选择。

（一）基于统计与阈值的方法

这是工程应用中最基础、最广泛的方法。其核心是根据历史正常运行数据，计算各参数在不同工况分区（如按风速、环境温度分区）内的统计特征（如均值、标准差、分位数），并结合设备制造商提供的技术允许范围，设定静态或动态报警阈值。动态阈值可根据滑动窗口内的统计量自适应调整，对工况变化具有一定的鲁棒性。该方法计算简单、实时性强，适用于快速检测明显的点异常和部分上下文异常。但其主要依赖历史经验，对未知类型的异常和复杂的模式异常识别能力有限，且阈值设置过严易导致误报，过松则导致漏报。

（二）基于规则与逻辑的方法

该方法将风机运行的物理原理、控制策略及专家经验编码为一系列判别规则。例如，“若风速大于切入风速且持续超过 T_1 时间，发电机转速仍为零，则判定为启动失败”；“若发电机有功功率 >0 ，而三相电流中有任一相为零或严重不平衡，则怀疑电流测量回路故障”。这类规则通常以“IF-THEN”的形式集成在 SCADA 系统或专家系统中。其优势在于可解释性强，运维人员易于理解和接受，能有效识别违反基本物理定律的异常。不足之处在于规则库的构建和维护高度依赖领域专业知识，难以穷举所有可能的故障模式，且对新机型、新问题的适应性较差。

（三）基于机器学习的方法

近年来，该方法成为研究热点，尤其适用于发现复杂、隐蔽的模式异常。主要包括：

无监督学习：如采用聚类算法对正常运行工况进行聚类，落在稀疏聚类簇或不属于任何簇的数据点可视为异常。主成分分析等降维方法可通过重构误差来发现偏离主体分布的数据。

有监督学习：在有大量标注好的正常与异常样本前提下，可训练分类器（如支持向量机、随机森林）进行判别。但在实际中，获取全面且准确的异常样本标签成本高昂。

时间序列预测模型：这是目前最具前景的方向之一。利用长短时记忆网络（LSTM）、门控循环单元等模型，学习关键参数历史序列的动态规律，预测其下一时刻的取值范围。当实际观测值持续偏离预测区间时，则判定为异常。此类方法能够捕捉参数间复杂的非线性时序依赖关系，对渐变故障的早期预警表现出良好潜力。然而，其模型性能严重依赖训练数据质量与数量，且具有“黑箱”特性，预测结果的可解释性有待提升。

三、工程应用实践：分层融合识别框架

基于华能新能源的实践，我们认识到单一方法难以应对所有异常类型。因此，我们提出并实践了一套“三层漏斗式”分层融合识别框架，旨在兼顾实时性、准确性与工程可行性。

（一）第一层：实时阈值与逻辑校验

在 SCADA 数据接入流处理层即进行快速过滤。首先，应用基于设备物理极限的“硬阈值”和基于历史统计的“软阈值”，滤除明显不可能的值（如负的功率、超速转速）。同时，运行一组核心的、高置信度的物理逻辑规则（如功率-风速曲线合理性检查），快速拦截低层异常。该层目标是高效处理大部分简单的点异常和系统性异常，减轻后端处理压力。

（二）第二层：工况划分与关联规则挖掘

对通过第一层的数据，按风速、温度、机组状态等进行精细化工况划分。在每个工况箱内，利用关联规则挖掘、相关性分析等方法，建立关键参数间的常态关联模型。例如，研究特定风速段下，发电机转速与轮毂转速的固定传动比关系，叶片角度与功率的关系等。当实时数据显著偏离这些关联模型时，则发出预警。该层专注于发现上下文异常和部分集体异常。

（三）第三层：智能模型深度分析

针对核心部件（如主轴承、齿轮箱、发电机）的健康状态监测，构建基于机器学习的深度分析模型。利用 LSTM 等时序模型，对各部件关键温度、振动、电流等多维参数序列进行联合建模，学习其正常模式。该模型以前两层清洗后的数据作为输入，运行周期可为小时或天级别，用于捕捉那些发展缓慢、在多参数中综合体现的早期模式异常。模型的输出是部件的或异常概率健康指数。

（四）闭环反馈与优化

建立异常案例库，将现场运维人员最终确认的故障原因、处理措施与对应的原始数据关联存储。定期利用这些经过验证的案例，对阈值、规则库以及机器学习模型进行反馈优化与迭代更新，形成“数据驱动识别-现场验证-模型优化”的持续改进闭环。

四、应用案例与效果分析

在华能新能源某装机容量 200MW 的沿海风电场，我们部署了上述融合框架进行试点应用。

案例一：风向标结冰异常识别。在冬季低温潮湿时段，第二层关联规则模块发现多台机组“风速-功率”曲线正常，但“风向-功率”曲线呈现异常：在风速足以发电的情况下，对风角度持续存在小而稳定的偏差。规则库结合气象数据（温度、湿度）判断为风向标潜在结冰，自动触发除冰装置激活指令并报警。事后登机检查证实了风向标轻微覆冰。该方法避免了因对风不准导致的发电量损失，平均单台机组在易结冰期挽回电量损失约 1.5%。

案例二：发电机冷却风扇共性缺陷预警。第三层 LSTM 模型在分析一批同型号机组的发电机数据时发现，在夏季高温时段（环境温度 $>30^{\circ}\text{C}$ ），发电机绕组温度预测值与实际值的残差系统性增大，且与冷却风扇电流的特定波动模式强相关。进一步对比分析该型号所有机组，确认了此共性模式。经与制造商联合排查，确认为该批次冷却风扇的散热设计余量在极端高温高湿环境下不足。基于此预警，推动了所有相关机组的预防性技改计划，将潜在的批量性发电机过热故障消除在萌芽状态。实践表明，该融合框架相比单一阈值方法，在误报率相当的情况下，对早期、复杂异常的检出率提升了约 40%，有效支撑了状态检修工作的开展。

五、结论

本文结合风电场实际运维需求，探讨了风电机组 SCADA 系统异常数据识别的方法与实践。研究认为，异常数据具有多模态特性，需根据其产生机理与表现形式，综合运用统计阈值、规则逻辑与机器学习等多种技术。提出的“三层漏斗式”融合识别框架，体现了从快速过滤到深度分析、从单一判据到多源融合、从静态规则到动态学习的工程化设计思路，在华能新能源的实践中取得了良好效果。

参考文献

- [1] 华能新能源股份有限公司技术标准. 风能发电系统智能风力发电场数据采集技术规范 (GB/T 44360-2024)
- [2] 李宝. 人工智能技术在风电机组智能巡检中的应用研究 [J]. 应用能源技术. 2025 (09): 115-117
- [3] 高鹏. 基于多线程并行的风电机组偏航滑移智能预警模型研究 [J]. 电力设备管理. 2025 (20): 153-155