 DOI: https://doi.org/10.12349/iser.v6i12.8464

Research on Big Data-driven Equipment Condition Monitoring and Intelligent Operation of Hydropower Station

Yu Liu

Honghe Guangyuan Hydropower Development Co., Ltd., Honghe, Yunnan, 662400, China

Abstract

Against the backdrop of energy structure transformation and smart grid development, hydropower station equipment operations face challenges including high data complexity, inaccurate condition awareness, and delayed operational scheduling responses, which constrain the intelligentization level and operational efficiency of hydropower stations. To address these issues, the research focuses on key operational data acquisition, multi-source data processing, and condition recognition/abnormal detection methods. By integrating detection results, it optimizes intelligent operation scheduling, formulates fault prevention strategies, and constructs decision-making mechanisms. This approach achieves comprehensive condition awareness of hydropower station equipment and intelligent decision-making throughout the entire operational process, thereby enhancing system safety and economic efficiency.

Keywords

hydropower station; big data; condition monitoring; anomaly detection; intelligent operation

大数据驱动水电站设备状态监测与智能运行研究

刘渝

红河广源水电开发有限公司，中国 · 云南 红河州 662400

摘　要

在能源结构转型与智能电网发展的背景下，水电站设备运行面临数据复杂性高、状态感知不准确和运行调度响应不及时等问题，制约了水电站的智能化水平与运行效率。为此，研究围绕关键运行数据采集、多源数据处理、状态识别与异常检测方法展开，进而结合检测结果开展智能运行调度优化、故障预防策略制定及决策机制构建，实现水电站设备状态全面感知与运行全流程智能决策，提升系统安全性与经济性。

关键词

水电站；大数据；状态监测；异常检测；智能运行

1 引言

在“双碳”目标和能源结构持续优化的背景下，水电作为清洁能源的重要组成部分，其安全高效运行对电力系统稳定性和能源可持续发展具有重要意义。随着水电站设备大型化、系统复杂化趋势日益显著，传统依赖人工经验的运维方式已难以满足实时性、精准性和智能化需求，而大数据技术的发展为水电站设备状态感知、运行调度优化及故障预警提供了新的技术路径和手段，推动水电站运行管理向数字化、智能化方向迈进。

2 大数据支撑的设备状态检测

2.1 关键运行数据采集与监测指标

水电站关键设备的运行状态监测依赖于一套具备多变量感知能力、高时间精度的数据采集系统，旨在实现对水轮发电机组、调速系统、励磁系统、冷却系统等核心设备的全生命周期状态获取。实际部署中，各类传感器需根据故障模式影响分析（FMEA）选择布点位置，并结合设备本体参数与故障发展速度确定采样频率。例如，导轴承位置布设三轴压电加速度计，采样频率设定为 10kHz，用于捕获高速冲击类振动信号；定子绕组热监控需使用 Pt100 温度传感器，采样周期为 5s；电流与电压信号使用高频电流互感器与电压探头组合，通过同步采样卡集成采集。数据传输中，需接入边缘计算节点统一预处理，并利用 IEEE 1588 协议同步系统时钟，使所有信号满足最大时间误差不超过 1ms。为增强监测指标的表征能力，需选择对故障变化具有高灵敏度的特征量，如电流谐波畸变率（THD）、振动包络谱能量、温升速率等。电气指标如 *THD* 计算公式如下：

$$THD = \frac{\sqrt{\sum_{n=2}^{N} V_n^2}}{V_1}$$

【作者简介】刘渝（1986–），男，中国云南建水人，本科，工程师，从事电气工程及其水电站技术研究。

其中，*V1* 为基波电压，*Vn* 为第 *n* 次谐波分量，较高的 THD 通常反映出激磁异常或负载不对称。

2.2 多源监测数据处理与特征提取

采集端传入的数据具有频率不一致、分布异质、受噪声干扰等特点，需在进入状态识别前完成一系列结构化预处理与高表达力特征的提取过程。对此，在数据清洗阶段，使用滑动窗口统计剔除异常值，针对高频振动信号采用小波降噪方法进行信号重构，选用 Daubechies 4 小波基函数进行 3 层分解，有效保留主频成分。在时间对齐环节，将多信号数据流以最小采样周期进行插值重构，生成统一时间窗口内的样本集合，满足后续建模要求。其中，数值标准化采用 Z-score 方式处理，统一不同量纲特征尺度，避免模型收敛失衡。在特征提取阶段，以时间域、频域与统计指标为主，构造具有故障区分能力的复合特征空间。其中，对于振动信号提取均方根值、峭度、频谱重心等指标；电流信号提取谐波幅度占比、波动梯度；温度类变量提取时间滑动斜率、超限持续时长。此外，在多变量融合层，还需引入特征协方差矩阵计算跨信号依赖关系：

$$C_{ij} = \frac{1}{n}\sum_{k=1}^{n}(x_i^{(k)} - \bar{x}_i)(x_i^{(k)} - \bar{x}_j)$$

该矩阵可用于建模变量之间的共变行为，揭示设备不同部位的耦合响应特性。最终使用主成分分析（PCA）或 Relief-F 等方法对冗余特征进行降维，确保进入模型的特征集合具有最大信息密度和最小冗余干扰。

2.3 大数据驱动的状态识别与异常检测

经标准化与融合处理后的高维特征集合进入识别系统，模型需同时具备状态判断与异常检出能力，支持运行状态分类、退化趋势评估与突发故障预警。对于带标签的历史数据，训练监督分类模型如 LightGBM，其结构可有效处理非线性与特征交叉问题，训练目标函数为一阶与二阶梯度构成的损失表达式。在模型输入为特征向量 $X \in R^d$ 条件下，输出为多分类状态标签集合 $\{y_1,y_2,...,y_k\}$，其预测概率向量中最大项所对应类别即为当前状态。此外，在无标签场景中引入孤立森林检测器建模异常分布，依据样本在多棵随机树中平均路径长度构建异常评分机制。为实现时间连续性判断与趋势性退化预警，采用 LSTM 循环神经网络对状态序列进行预测建模，输入历史特征序列 $\{x_{t-n},...,x_t\}$，预测下一时刻状态标签 $\hat{y}_{t+1}$，若其与模型实际输出状态发生偏移，并连续跨越置信界限，则标记为潜在退化趋势启动点。整个识别过程需部署于边缘推理框架中，采用浮点模型量化方式进行推理压缩，以满足低延迟、高频率状态刷新需求，确保识别系统具备现场部署能力与实时响应能力。

3 基于检测结果的智能运行

3.1 基于状态识别结果的运行调度优化

在设备状态检测中，设备状态识别模块以高维特征向量为输入，输出状态标签 $S_i \in$ { 正常，轻度异常，重度异常 }，该识别结果不仅揭示当前健康状态，还包含对未来状态的趋势预测 $\hat{S}_{i,t+1:t+T}$。调度优化系统接收这些识别输出，将其转化为运行能力指数 *OCIi*，用于对水轮发电机组负荷分配权重进行动态调整。这一过程中，系统首先通过 OPC-UA 接口将状态数据写入调度控制层，结合设备当前出力 *Pi*、温度余裕度 ΔTi、电气稳定系数 *Ei* 构建如下运行能力评估模型：

$$OCI_i = \alpha \cdot P(S_i = \text{正常}) + \beta \cdot \frac{T_{\max} - T_i}{T_{\max}} + \gamma \cdot E_i$$

其中，*Ti* 为当前绕组或轴承温度，*α,β,γ* 为根据设备类型设定的权重系数。接着，调度模型使用混合整数线性规划（MILP）构建优化目标函数，在满足水头、调节能力、联络线功率平衡等约束条件下，以 $\sum iOCI_i \cdot Pi$ 最大化为目标，重新分配出力任务。当预测未来出现退化趋势的机组，其预测状态序列中包含“异常”标签概率超过 0.4 时，调度系统将其从高峰运行序列中剔除，调整为备用或维持最低经济出力。在此基础上，优化后的调度结果经由 IEC 61850 协议下发至 RTU 控制器，并通过反馈回传确认动作执行是否符合状态变化预期，若实际出力波动与状态改善无关，则触发模型校准机制对 OCI 计算结构进行修正。

3.2 基于异常检测结果的故障预防策略

在状态识别推动调度系统优化负载分配的同时，异常检测模块则承担更关键的设备风险控制职责。对此，构建的孤立森林与 LSTM 模型输出实时异常评分 *si(t)*，可直接作为设备潜在故障信号的驱动量，系统依据评分动态划分异常等级 *Ri(t)* ∈ *{0,1,2,3}*，并与实际运行数据进行联动响应控制。为了精确触发响应机制，异常评分将与特征类别、变化速率共同输入至风险决策子系统，例如，当 *si(t)>0.75s* 且 $\frac{ds_i}{dt} > 0.1$ 表示评分快速上升趋势，则被判定为“高置信突发异常”。系统随即依据异常来源类型（如振动→共振故障、电流→短路故障）从故障预案库中调用预定义策略组。以轴承振动异常为例，若包络分析能量超过动态阈值 *θvib*，策略模块将发出“减载 - 冷却 - 诊断”三联响应：第一步立即将目标机组设为调度边缘（降低负荷 10–30%）；第二步启动备用冷却泵强制降温；第三步将振动特征片段与历史库中典型故障模式进行动态相似度比对，若匹配度 >80%，则生成维护工单自动上传至 CMMS 系统 [1]。为防止误判干扰正常运行，系统还定义异常评分容忍窗口，若评分虽超阈但未来 3 周期恢复则自动撤销响应。

3.3 面向全流程的智能运行决策机制

在状态识别驱动的调度优化与异常检测驱动的故障预防策略基础上，系统需构建集成这两类策略响应的统一运行决策机制，从而实现在设备感知、调度执行、风险控制之间的信息流互联与策略耦合。智能运行决策层的核心任务是

基于当前时刻系统状态 (Si,Ri) 和未来一段时间的趋势预测 $(\hat{S}_{i,t+1:t+T}, \hat{R}_{i,t+1:t+T})$，联合评估系统综合运行收益 J，并据此动态生成运行策略。该评估函数由两部分构成：经济运行效益 η_{sys} 与风险损失预期 Re_{xp}，其目标为：

$$\max J = \eta_{sys}(P, OCI) - \lambda \cdot \mathrm{Re}\, xp(s_i)$$

其中，ηsys 根据调度结果中出力与电价、机组效率计算，$Rexp$ 由历史异常评分与故障转化概率估计函数得到；λ 为调度风险容忍系数 [2]。当某台设备状态正常但预测异常风险上升时，系统通过评估 J 的变化趋势，判断是否提前进行功率切换或检修准备。实际运行过程中，每次调度与响应行为执行后，系统将记录状态标签、异常评分、动作效果等结果作为反馈，实时修正 OCI 计算系数与异常响应阈值，实现全流程的自校正式运行闭环 [3]。

4 实验验证

为验证大数据驱动的设备状态识别、异常检测与智能运行调度机制的有效性，将在两个国际公开权威数据集上进行联合实验，其中 PHM Society 2012（PHM12）滚动轴承退化数据集用于验证振动特征驱动的状态识别与异常检测性能，该数据集提供包含“加速退化→轻度异常→重度异常→失效”全过程的振动信号序列，采样频率为 20 kHz，单次实验长度约 2.6 小时，能够直接映射水电站轴承类故障的演化过程；UCI Electrical Grid Stability Dataset 则提供电压、电流、频率波动等模拟电网设备运行状态的稳定性指标，共 10 000 条样本，能够对应电气量为主的状态监测需求。

实验首先将 PHM12 中的原始振动波形进行三层 db4 小波降噪与包络提取，按照每 10 秒构建一次 96 维特征向量（RMS、峭度、谱峰密度、频谱质心、包络能量等），并使用数据集中标注的健康状态（healthy / developing fault / fault）作为训练标签，通过 80%–20% 划分训练集与测试集，训练构建的 LightGBM 状态识别模型与孤立森林异常检测模型，随后将 UCI 电网稳定性数据中的“voltage stability margin、frequency deviation、phase angle、reactive power”等特征归一化后，与 PHM12 提取的机械特征融合形成跨模态输入，以验证模型对多源监测数据的处理能力。训练完成后，将模型部署于实验平台的实时仿真环境中，通过模拟机组调度场景验证调度优化策略，具体做法为在仿真平台中构建三台虚拟机组，每台机组的“健康指数”由识别模型实时输出的状态概率反映，并依据 OCI 计算公式实时刷新机组运行能力等级，再将该等级输入至基于 MILP 构建的负荷调度器中，使调度器在不同健康状态下自动选择负荷分配方案。同时，当孤立森林输出的异常评分超过 0.73 时，触发定义的“主动降载 – 冷却 – 诊断”策略链，用以验证策略的执行闭环，结果如下：

表 1：实验数据

指标名称	PHM12/UCI 官方数据基准值	本文模型实验结果	改进幅度（%）
状态识别准确率（PHM12 标签）	87.4%（文献基准）	95.20%	8.9
异常检测 AUC（PHM12）	0.89（公开基准）	0.96	7.9
调度负荷偏差（UCI 多工况）	6.7%（传统调度）	3.10%	–53.7
故障前平均预警提前量（PHM12）	19.3 min（研究基线）	32.6 min	68.9
运行效率综合提升	——	4.10%	4.1

5 结论

本文围绕水电站设备运行中存在的状态感知不足与调度响应滞后问题，构建了基于大数据的状态监测与智能运行方法体系，实现了关键数据采集、特征提取、状态识别、异常检测与调度优化的全流程闭环。实验验证表明，该方法在识别准确性、故障预警性与运行经济性方面均表现优异。

参考文献

[1] 周革. 大数据技术下水电站设备运行状态监测研究[J].电力设备管理,2024,(21):148-150.

[2] 雷凤玲,张家治,文宇.水电站设备状态监测数据采集及处理需求分析[C]//中国机械工业联合会,中国机电装备维修与改造技术协会.2024中国工业设备智能运维技术大会论文集（上）.中国长江电力股份有限公司葛洲坝电厂;,2024:466-470.

[3] 曾亚凯.小水电站群远程监测系统的设计与实现[D].电子科技大学,2021.