基于数据挖掘的风力发电机故障预测

何正军
安徽省新能创业投资有限责任公司 安徽 合肥 230051

摘 要 风力发电是最具规模化、最有潜力的可再生能源之一，其装机容量不断增长。风力发电机组的主要部件长时间在复杂多变的恶劣工作环境下运行，风机的故障率较高，使非计划停机损失和运行成本不断增长，严重影响了风电场的经济效益以及安全运行。传统的定期维护和事后维修方式存在盲目性较高，滞后性的问题。对于风力发电技术在特殊环境下出现的故障以及传统维护方法存在的问题，本文将以某风电场的实际基础为研究对象，配合SCADA系统运行的数据深入挖掘其运行的故障情况，构建其预测模型，重点分析风电机组的各类故障出现的主要原因及其带来的影响。随后，本文提出基于BAS-NSET算法的齿轮箱状态监测模型和基于BAS-BP神经网络的绕组温度预测模型，实现故障早预测。研究表明，此方法能够提升故障诊断准确率，减少因为非计划而引发的停机损失，为风力发电厂提供技术支持。基于此，本文将围绕该技术的数据处理流程、主流预测模型、实践挑战与评估方法方法展开讲解，并对未来的发展进程进行分析。

关键词 数据挖掘；风力；发电机故障；预测

Fault Prediction for Wind Turbines Based on Data Mining

He Zheng-jun
Anhui Xinneng Venture Investment Co., Ltd., Hefei 230051, Anhui Province, China

Abstract Wind power is one of the most large-scale and promising renewable energy sources, with its installed capacity continuously growing. The main components of wind turbine generators operate for extended periods under complex, variable, and harsh working environments, resulting in a high failure rate of wind turbines. This leads to increasing unplanned outage losses and operational costs, severely affecting the economic efficiency and safe operation of wind farms. Conventional scheduled maintenance and corrective maintenance approaches suffer from high levels of blindness and inherent lag. In response to the faults occurring in wind power generation technology under special environmental conditions and the limitations of traditional maintenance methods, this paper takes the actual operational foundation of a specific wind farm as the research subject, leveraging data from the SCADA system to deeply investigate its operational fault patterns and construct predictive models. The analysis focuses on the primary causes of various types of wind turbine faults and their associated impacts. Subsequently, this paper proposes a gearbox condition monitoring model based on the BAS-NSET algorithm and a winding temperature prediction model based on the BAS-BP neural network to achieve early fault prediction. The research demonstrates that this approach can improve fault diagnosis accuracy, reduce unplanned outage losses, and provide technical support for wind farms. Based on these findings, this paper elaborates on the data processing workflow of this technology, mainstream predictive models, practical challenges, and evaluation methods, while also analyzing its future development trajectory.

Key words Data Mining; Wind Power; Generator Fault; Prediction

随着风电监控和数据采集系统（SCADA）的普及，大量的运行状态数据为通过数据分析实现故障提前预判提供了可能。数据挖掘和机器学习的故障预测技术的目的在于从高维、时序的运行数据中自动挖掘故障演化规律和早期征兆，推动运行维护方法向着主动、准确的预测性维护转变。这不仅可以保证机组的运行稳定以及安全，还可以提升发电的可靠性，能够有效降低全生命周期的成本，是增强风电行业竞争力的有效途径。

1 工程概况

某风电场共安装64台联合动力UP82-1500型风电机组，总容量95.5 MW，机组采用水平轴、双馈感应发电机等设计，核心部件包括风轮、变桨系统、传动链、齿轮箱、发电机等，配套SCADA中央监控系统实时采集47项运行参数。因为此风电场建设早而且运行时间长，所以出现了各类设备问题，全年累计故障发生率较高，隔离突发故障导致其损失电量高达

1 139 kWh，因此本文将针对此风电场的情况展开探讨，希望能够构建起新的故障预测模型，有效保障该风电场的安全稳定运营，以下将展开详细的论证分析。

2 风电场设备故障现状与问题分析

2.1 故障发生总体情况

因为此风电场已经持续运行多年，各类设备老化问题越来越突出，整体故障类型分布较集中：齿轮箱温度类故障占30%，发电机温度类故障占25%，变桨系统类故障占15%，变频器类故障占10%，电网类、转速类故障各占5%，其他类型故障占10%。各类故障导致风电场出现多次非计划的停机，其中在7—9月超温问题最为严重，而且在大风天气之下也会出现大面积的机组停机；在3—5月齿轮箱油温超限故障也频发，1—4月和9—12月多出现叶轮转速故障。

2.2 主要故障类型及损失

发电机与齿轮箱是最为核心的部件，如果出现故障会影响发电的效率，依据故障统计的数据得知，因为电机超温、齿轮箱超温以及叶轮转速出现问题等导致电量损失高达1 139 kWh，其中，齿轮箱相关故障是损失的主要来源。



图1 高速轴前轴承内圈损坏状态

所以近期对多台机组齿轮箱的损坏情况进行了深入的分析，其根本原因是高速轴前承轴的损坏（图1），进一步检测发现其原因在于：首先，根本原因是轴承损坏，并不是高速轴加工装配的问题，也不是材料成分的问题，经过权威机构检测得知，在损坏轴承内圈滚道次表层处都发现了特殊的圆形斑点，其成分与本体一致，未来会成为应力集中点和裂纹发生处。其次，失效的原因是在风机运行时，高速轴前轴承承受了复杂的径向力和周期，另外如果风况不稳定，轴承内圈会受到异常的冲击或挤压，此时表层的圆形斑点在交变应力的作用下，会出现细微的裂纹，这个裂纹会逐步扩展，最终会导致内圈断裂和滚道剥落，所以会引发齿轮箱失效。最后，对于此次故障，由于维修成本超过了80万元，而且导致电量损失约54万kWh，这表明在承轴出现初始微裂纹到完全失效的过程中，SCADA系统中的齿轮箱、油温、轴温等参数都可能是潜在风险和预警，如果错过了早期的干预，后期会引发严重的损失[1]。

3 基于BAS-NSET算法的齿轮箱状态监测与故障早期预警

3.1 BAS-NSET齿轮箱温度预测模型建立

基于天牛须搜索算法（BAS）与非线性状态估计方法（NSET），构建齿轮箱温度预测模型，具体步骤如下。

首先，进行模型变量变形筛选，从SCADA系统中将采集到的47项运行参数与所选取的齿轮箱温度强相关的5组核心数据[机组功率（P）、风速（u）、环境温度（T）、当前时刻齿轮箱温度（T_G）及前一时刻齿轮箱温度（T_{G1}）]组成观测向量。在这组向量中，机组功率、风速决定着齿轮箱荷载与转速，环境温度会影响散热效率，历史温度则体现温度变化的关系，多变量协同能够精准、全面地反映整体的运行状态。

其次，对选取的2018年3月的数据进行基础建模，在剔除停机的无效数据后，保留了3 731组有效的10 min均值数据，对于不同变量量纲差异，使用归一化处理方法将数据投射[0,1]区间，确保向量距离计算的准确性；对风速变量设定边界阈值，切入风速3 m/s、额定风速12 m/s，超出范围数据按边界值修正。随后确定好数据之后，要进行算法融合优化，利用BAS算法优化NSET模型权值向量。以预测残差最小化为目标构建适应度函数；初始化BAS算法参数，空间维度k设为5，两须距离衰减系数η_d与步长衰减系数η_δ均设为0.95，通过随机向量生成、左右须位置计算、适应度值对比及位置迭代更新，获取最优权值向量，解决传统NSET模型权值求解复杂的问题。

最后，优化完成后，生成最终模型。将优化后的权值向量代入NSET模型，通过非线性算子（欧氏距离）实现输入向量与记忆矩阵的映射，生成齿轮箱温度预测值，最终形成BAS-NSET齿轮箱温度预测模型，实现对齿轮箱运行状态的实时监测。

3.2 轴承故障的模型构建与验证

为了有效验证模型对齿轮箱轴承早期故障预警的性能，选取发生故障的17#机组F13881在故障前3个月的具体数据，以及同型号的健康机组的对比数据。首先，对数据进行了预处理，提取时域和频域特征。其次，使用健康机组数据及故障机组数据进行前期的数据分析，构建NSR过程记忆矩阵，并运用BAS算法对数据进行优化。最后，进行验证，将故障机组失效前数周的数据输入训练好的模型计算状态估计残差分析发现，在传

统温度监测没有超限的早期阶段，模型输出的大多数参数联合残差序列已经出现持续缓慢上升的趋势和间歇性尖峰，这与风况突变荷载事件高度相符。残差的变化模式比单一的轴承温度参数提前了2～3周，进行了预警，此次模拟了轴承次表层圆形斑点在突变荷载下出现裂纹的早期阶段[2-3]。

3.3 模型有效性分析

为验证模型性能，选取某风电场UP82-1500型风电机组的实测数据进行仿真验证，具体过程如下：选取488组输入向量作为测试样本，涵盖不同风速、功率及环境温度工况，确保验证场景的全面性；将测试样本输入BAS-NSET模型，计算齿轮箱温度预测值与实测值的残差。

$$\varepsilon GT=X_4-x_4$$

式中，X_4—实测温度；x_4—预测温度。

在机组正常运行工况下，模型预测残差绝对值均低于0.005，均值接近0且标准差极小，表明模型在动态工况下仍具备极高的预测精度，能够精准拟合齿轮箱温度变化规律。模拟齿轮箱油超温停机、指令停机及重启等故障场景，结果显示故障发生时，模型预测残差出现显著突变，如第72点（齿轮箱油超温停机）、第203点（指令停机）的残差明显高于正常工况，且残差突变与故障发生时刻完全同步，能够有效捕捉故障引发的温度异常变化。与传统NSET模型对比分析，BAS-NSET模型的预测残差更小，对温度微小变化的响应更灵敏，在低载荷、环境温度波动等复杂工况下，预测精度提升显著，证明BAS算法对权值向量的优化效果显著。验证结果表明，BAS-NSET模型能够精准预测齿轮箱温度，在故障发生时通过残差突变实现早期预警，可有效识别齿轮箱潜在故障，为风电场预防性维护提供可靠的数据支撑，具备较强的工程实用性[4]。

4 未来发展趋势

在将来，风力发电机故障预测技术将呈现以下发展趋势。在多元信息融合和数字孪生方面：将来的系统不再局限于SCADA的数据，而是融合更高频的振动监测、声学监测、热成像、无人机巡检图等多源异构信息，建设更加全面的设备健康画面。在这些基础上，再结合物理模型和数据模型，建设风机的高保真数字孪生体，实现故障的仿真推演以及动态预测。在小样本学习和迁移学习方面：对于新机型、新场站故障样本稀缺的问题，利用迁移学习将已有大数据集上训练的知识迁移到新的场景中，或者利用小样本学习等技术使模型具有快速适应新任务的能力，这是解决数据问题和泛化问题的重点。在边缘计算与云边协同方面：由于边缘计算设备能力的不断提升，部分轻量级模型可以部署在风机或者变电站本地进行推理，有效实现了毫秒级的异常检测以及快速保护，复杂模型训练和大数据分析在云端进行，形成了云边协同的智能化运维框架。在标准化和开源生态建设方面：学术界和工业界需要共同努力，推动数据标注规范、模型评估基准和系统接口向着标准化的方向发展[5-6]。

5 结束语

对于风力发电厂机组故障预测需要基于现有的数据进行深度的挖掘，找到数据运行中存在的规律，要让故障维修从被动转向主动。本文基于原有系统所有的数据构建出两种预测模型，进一步解决了风电厂核心部件故障引发的各类问题，降低了风电场的运维成本与发电损失。未来，随着多技术的发展，风力发电机故障会朝着更加精准、高效、智能的方向发展，持续为风力产业注入动力。

参考文献

[1]魏书荣,周海林,符杨,等.基于温度热模型与数据融合驱动的海上风力发电机故障早期预警[J].高电压技术,2025,51(10):4945-4956,I0001,I0002.

[2]杨伟新,赵洪山,张扬帆,等.基于PKFF-Transformer的风力发电机故障预测[J].机床与液压,2025,53(4):221-229.

[3]刘杰,李政浩,乔旭东.风力发电机组电气系统的自动化故障预测模型研究[J].全面腐蚀控制,2025,39(8):60-62.

[4]刘吉祥,达琪,陈军,等.大数据背景下风力发电机组故障诊断研究[J].新潮电子,2025(14):28-30.

[5]王建伟.基于多模型集成算法的风力发电机组电气控制系统优化与故障预测研究[J].电力系统装备,2025(7):80-81,140.

[6]李明,曹洁,刘宗礼,等.多度量下ResGAT的风力发电机齿轮箱故障诊断[J].太阳能学报,2025,46(6):683-690.

（本文编辑　曹　阳）

