

基于数据驱动的风电机组偏航位置异常预警方法研究

沈旭¹ 王海云¹ 黄晓芳²

(1. 新疆大学 可再生能源发电与并网技术教育部工程研究中心, 乌鲁木齐 830047)

(2. 北京金风科创风电设备有限公司, 北京 100176)

摘要: 在偏航过程中偏航位置异常一方面会导致偏航位置误差积累,影响偏航对风精准度或导致电缆过度扭绞而影响安全,另一方面频繁位置跳变或频繁短时位置保持均会产生一定的偏航误差、影响偏航控制稳定性,从而导致偏航系统故障发生频率升高和运维成本增加等问题,因此提出了一种基于数据驱动的故障诊断方法,用于对风电机组偏航位置的异常情况进行预警。首先,针对数据采集与监视控制(Supervisory Control and Data Acquisition, SCADA)系统中的海量数据,采用基于标准化交互增益的Relief-F(Standardized Interaction Gain and Relief-F, SIG-Relief-F)特征算法筛选出用于识别与目标变量(在这种情况下可能是偏航系统故障)具有最强关联性的多个特征参数。这种方法的优势在于能够有效地考虑到特征之间的相关性,最大程度地保留偏航系统故障相关特征与交互特征。其次,建立反向传播神经网络(Back Propagation Neural Network, BPNN)偏航位置预测模型,通过滑动窗口法对残差的分布进行统计,从而确定故障阈值。最后,通过实例验证了所提方法的有效性与准确性,并通过对比多元状态估计技术(Multivariate State Estimation Technique, MSET)和支持向量机(Support Vector Machine, SVM)算法,验证了其具有更优的异常预警性能。研究结果可为实际偏航系统的故障诊断提供参考。

关键词: 风电机组; 偏航位置; 交互信息; Relief-F; 反向传播神经网络; 异常预警

中图分类号: TK83 **DOI:** 10.16579/j.issn.1001.9669.2025.10.008

0 引言

根据中国国家能源局最新统计数据,截至2022年10月底,全国风电装机容量约3.5亿千瓦,比2020年增长16.6%。SANTOS等^[1]提出,相较于2022年,到2030年时全球海上风电产业将实现5倍增长。随着可再生能源的稳步发展,海上风电未来在清洁能源领域将扮演至关重要的角色,并展现出广阔的发展前景。风电机组主要包括叶片、齿轮箱、发电机、变桨系统、偏航系统等关键部件^[2]。其中,偏航系统是实现风电机组快速精准对风、捕获最大风能的关键执行机构之一^[3]。偏航系统的运行状态直接影响着风电机组的性能和发电量。如果偏航系统出现故障,可能会造成风电机组的失控或产生不必要的资源浪费,甚至危及机组的稳定运行。此外,不稳定的偏航系统会降低风电机组的安全性和可靠性,并增加运维成本。

风电机组多数工作在较为恶劣的自然环境中,长期受到温度、积雪、盐雾、沙尘等极端复杂工况的影响,且多变的风况致使偏航系统需要在不同运行状态

间频繁变化,机组将承受多重载荷从而导致绝缘强度、疲劳强度和运行性能等随运行时间增长而逐渐下降,最终导致故障发生^[4-5]。据不完全统计,风电机组发电总成本的25%~50%来自运营和维护费用^[6-7],其不同部件故障中偏航系统故障占比约6.7%,其所造成的停机时长占比达到13.3%。因此,对风电机组偏航系统异常状态进行准确识别从而进行提前预警以保障机组的安全稳定运行具有重要的现实意义。

现阶段对偏航系统的故障诊断及优化控制已提出了一些方法。张海涛等^[8]将风向特性提炼为Elite因子并结合优化蚁群智能算法,克服了反向传播神经网络(Back Propagation Neural Network, BPNN)易陷入局部寻优的不足,提高了处理有效数据的精度,大大缩短了偏航齿轮箱故障的诊断时间。冯俊恒等^[9]通过叶片气动分析并结合仿真方法,对不同偏航角下的风电机组载荷进行了仿真分析,得出正向的偏航偏差角会导致更大的风电机组轮毂载荷的结论,为大型风电机组偏航控制及保护提供了有效参考。邓子豪等^[10]利

收稿日期: 2024-02-21 修回日期: 2024-03-20

基金项目: 天山英才项目(2022TSYCJC0030); 自治区重点研发计划(2022B03031); 哈密高新区科技项目(HGX2023KJXM008)

作者简介: 沈旭,男,1997年生,江苏扬州人,硕士研究生; 主要研究方向为风电机组数据挖掘与故障诊断技术; E-mail: shenxu971011@163.com。

王海云(通信作者),女,1973年生,新疆乌鲁木齐人,教授,博士研究生导师; 主要研究方向为新能源发电与并网技术; E-mail: why@xju.edu.cn。

引用格式: 沈旭,王海云,黄晓芳. 基于数据驱动的风电机组偏航位置异常预警方法研究[J]. 机械强度, 2025, 47(10): 71-79.

SHEN Xu, WANG Haiyun, HUANG Xiaofang. Research on data-driven abnormal warning methods for wind turbine yaw positions [J]. Journal of Mechanical Strength, 2025, 47(10): 71-79.

用 Relief-F (Relevant Features-F) 算法和核密度-均值法提取了能反映偏航齿轮运行工况的 7 个特征参数, 结合神经网络来诊断偏航齿轮箱的正常状态、磨损状态以及断齿故障共 3 种运行状态, 试验证明该方法可以有效地辨别故障类型。ZHAO 等^[11]采用监控与数据采集 (Supervisory Control and Data Acquisition, SCADA) 系统对风电机组的偏航数据进行分析, 提出了一种基于激光测距的经济模型预测控制 (Economic Model Predictive Control, EMPC) 偏航策略, 来评估偏航过程、汇总偏航轴承的发电量和损伤当量载荷。最终利用 XE112-2000 型风力发电机组模型在仿真软件 Bladed 中验证了 EMPC 的有效性。ZHAO 等^[12]基于 GEZ 齿轮箱故障诊断试验平台, 提取了振动试验数据, 通过时域和频域分析方法, 利用 Matlab 软件通过特征值和波形等相关信息对信号进行有效分析, 再通过对比正常信号波形和数值参数得到故障信号, 最后通过试验给出故障结果。上述研究都为偏航系统的故障诊断提供了新的研究方向。

当前偏航系统的故障诊断方法大都是针对偏航齿轮相关故障的识别与分类, 而通过偏航位置的异常跳变来识别偏航系统是否产生异常则缺少相关的方法研究。毫无疑问, 偏航位置的准确性是机组实现有效偏航控制的基础, 风电机组实际运行过程中发生过因偏航位置异常 (凸轮跳齿) 导致机组扭摆故障的案例。此外, 偏航位置的异常也反映出某些偏航驱动机械故障, 如齿轮磨损啮合不准、偏航电动机故障、齿轮磨损断齿等。

本文基于 SCADA 系统的瞬态数据, 提出了一种基于数据驱动的风电机组偏航位置异常诊断方法。该方法通过对风电机组运行数据进行异常模式识别, 构建偏航位置异常预警模型, 并通过历史数据在大数据平台上进行持续迭代和优化。同时, 将异常识别结果进行预警标记, 并将结果发送至风电机组现场进行校核, 从而形成一个闭环的故障诊断流程。其中包括数据采集、异常诊断与预警、校核排查、反馈和事件处理。通过这一流程, 机组的运行风险大大降低, 机组的安全性得到提高, 同时减少了运维成本。通过平台为期 3 个月的实际运行结果证明, 该模型的异常识别准确率达到 92%。

1 预备知识

1.1 偏航系统基本机制

偏航是指风电机组通过机舱的水平调整使其朝向风向, 进而确保叶片始终朝向风的方向。从控制的

角度来看, 偏航系统分为被动式和主动式 2 种类型。被动式偏航系统利用风作用于叶片产生的转矩, 通过空气动力效应自动将风力发电机与风向对齐。当风向发生变化时, 风作用在风轮叶片上产生的力矩, 通过气动效应使风电机组自动对准风向。主动式偏航则是利用电力或者液压驱动来完成对风动作的偏航方式。目前风电机组偏航驱动结构大都为水平轴、主动型、电动驱动的滚动偏航^[13]。偏航系统通常位于风电机组的机舱顶部, 靠近风轮和塔架之间, 主要由偏航齿圈、偏航电动机 (含电磁制动器)、偏航卡钳、偏航小齿轮、偏航减速器等部分组成。偏航系统在发电状态下与风电机组的控制系统相互配合, 使风轮始终处于迎风状态, 充分利用风能, 提高风电机组的发电效率。在大风、暴风等超过切出风速的情况下, 机组停机, 通过偏航减少机组载荷, 保证机组的安全。偏航系统基本结构如图 1 所示。

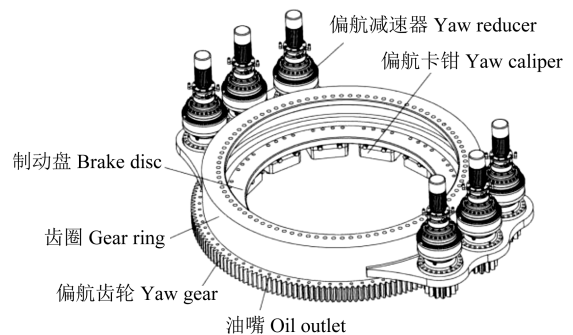


图 1 偏航系统基本结构

Fig. 1 Basic structure of the yaw system

1.2 偏航系统常见故障类型

在风电机组长时间的运行过程中, 加上长期处于各种恶劣的环境下, 偏航系统的关键部件极易造成疲劳损伤, 从而不可避免地出现一些机械故障, 常见的故障类型如图 2 所示。

根据机组的实际运行情况可知, 偏航系统的各部件从异常到最终导致故障停机是一个累积的过程, 这点从 SCADA 数据层面剖析尤为明显。初级的偏航部件受损一般表现为齿轮磨损、电磁刹车片磨损、电动机温度异常等; 接着就是中级磨损, 如轴承开裂、凸轮跳齿等, 如果在此阶段未能对偏航系统的相关部件进行及时排查与维护, 将导致更加严重的系统故障, 甚至危及整机运行安全^[14]。

在偏航系统发生初级损伤时, 其偏航位置的变化较为敏感。因此本文将偏航位置作为偏航系统的预测参数, 通过其变化的表现形式识别偏航系统的异常。

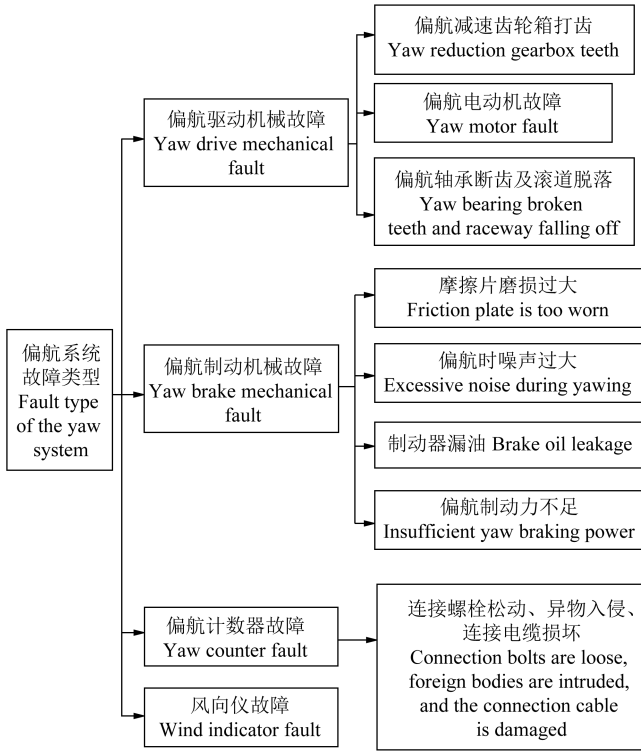


图 2 偏航系统常见故障类型

Fig. 2 Common fault types of yaw systems

1.3 改进的 Relief-F 算法

1.3.1 Relief-F 特征选择理论

Relief-F 算法是一种常见的特征选择算法,用于从给定的特征集中选择具有较高相关性和重要性的特征。其主要思想^[15]是通过计算特征之间的距离和样本之间的距离来评估特征的重要性。简而言之,就是通过特征变量区分不同样本的能力来赋予权重,对于区分能力较强的特征变量增加权重,而对于区分能力较弱的变量则减小权重。从训练集 S 中随机挑选出一个样本 A , 然后分别计算 A 的不同类最近邻 (NearMiss) 和同类最近邻 (NearHit), 最后通过以下规则迭代每个特征的权重^[16], 即

$$W(X) = w(X) - \frac{\sum_{j=1}^k f_{\text{diff}}(X, A_i, H_j)}{nk} + \sum_{C \neq \text{Class}(A_i)} \frac{P(C)}{1 - P[\text{Class}(A_i)]} \times \frac{\sum_{j=1}^k f_{\text{diff}}(X, A_i, S_j)}{nk} \quad (1)$$

式中, X 为某一特征参数; $w(X)$ 为特征 X 的权重, 其初始值通常设为 0; n 为循环次数; A_i 为第 i 次选择的随机样本; k 为选取的近邻个数; H_j 为样本 A_i 同种类别中的最近邻样本; S_j 为样本 A_i 不同类别中的最近邻样本; $\text{Class}(A_i)$ 为样本 A_i 的类别; $P(C)$ 为类别 C 所占的比例; $f_{\text{diff}}(X, A_i, H_j)$ 为同种类别的故障样本在特征参数 X 上的差值。

1.3.2 特征交互

交互信息是一种用于衡量随机变量之间相互作用程度的概念。它基于信息论的原理,用于描述变量一起出现时所包含的信息量,以及这种信息量相对于单独出现时的信息量的增益。因此,交互信息又被称为交互增益 (Interaction Gain, IG)。本文中的 IG 指的是在特征选择中用于衡量多种特征之间相互作用对于目标变量的预测性能提升程度。其中,三方交互增益的定义^[17]为

$$IG(f_i; f_j; C) = I(f_i; f_j; C) - I(f_i; C) - I(f_j; C) \quad (2)$$

式中, $I(f_i; f_j; C)$ 为在类别 C 中特征变量 f_i 与特征变量 f_j 之间的交互信息。当 $IG(f_i; f_j; C) > 0$ 时,说明特征变量 f_i 与特征变量 f_j 联合所提供的信息量大于 2 个单体特征变量分别提供的信息量之和,说明了两者的之间存在交互性,否则,说明了两者的联合所提供的信息量与类别 C 无关或者信息量冗余重复。

1.3.3 基于特征交互的 Relief-F 算法

标准化交互增益 (Standardized Interaction Gain, SIG), 可以使得不同特征之间的交互增益具有可比性,更好地反映它们对模型性能提升的贡献程度。其定义为

$$SIG(f_i; f_j; C) = \frac{IG(f_i; f_j; C)}{H(f_i) + H(f_j)} \quad (3)$$

式中, $H(f_i)$ 为任意 f_i 的熵, 其定义为 $H(f_i) = -p(f_i) \log[p(f_i)]$, $p(f_i)$ 为 f_i 发生的概率; $H(f_j)$ 亦如此。将式(1)与式(3)相结合,可以得到改进后特征 X_i 的权重计算式为

$$\bar{w}(X_i) = \frac{1}{n} \sum_{j=1}^N SIG(X_i; X_j; C) + w(X_i) \quad (4)$$

将式(4)作为新的算法来更新权重,这里定义为标准化交互增益的 Relief-F (Standardized Interaction Gain and Relief-F, SIG-Relief-F) 算法。相比于传统的 Relief-F 算法,它可以消除特征之间不同量纲的影响,提高特征权重计算的准确性,并更全面地考虑特征之间的相互作用,从而提高特征选择的效果、改善模型的性能。

2 偏航位置异常预警模型的建立

2.1 模型输入参数的选择

本文的数据来源是某风场 2022 年 1 月—12 月份机组的运行数据,机组编号为 $F_1 \sim F_6$ 。风电机组的基本参数为:额定功率为 5 200 kW;叶轮直径为 165 m;切入风速为 3 m/s;切出风速为 24 m/s;数据间隔时间为 7 s。所记录运行参数包括时间、风速、功率、偏航位置、对

风角度、偏航速度等。

在SCADA系统中,通常会监测和记录大量的状态参数,这些参数可以用于监测设备或系统的运行状态。然而,对于状态参数预测模型的构建和应用来说,使用所有可用参数可能会导致模型过于复杂,同时也可能引入不相关或冗余的信息,从而降低识别精度。

使用SIG-Relief-F算法对历史数据进行特征筛选,与风力发电机偏航系统异常相关的特征参数如表1所示。

表1 偏航系统状态参数特征权重

Tab. 1 Feature weight of status parameters of the yaw system

序号 Number	状态参数 State parameter	特征权重 Feature weight
1	偏航位置 Yaw position	0.089
2	偏航角度 Yaw angle	0.062
3	偏航速度 Yaw speed	0.078
4	偏航系统压力 Yaw system pressure	0.063
5	偏航电动机回路1L ₁ 相电流 1L ₁ phase current of the yaw motor loop	0.079
6	偏航电动机回路2L ₁ 相电流 2L ₁ phase current of the yaw motor loop	0.082
7	偏航电动机制动器磨损状态 Wear state of the brake of the yaw motor	0.084
8	风速 Wind speed	0.086
9	有功功率 Active power	0.065
10	偏航电动机工作时间 Operating time of yaw motors	0.061
⋮	⋮	⋮

通过计算历史数据中与偏航系统有关的38个特征参数的权重可得,偏航位置、偏航速度、偏航电动机电流和偏航系统压力等参数具有相对较高的权重。因此,本文选择表1所示的10个特征参数形成观测向量集,作为异常预警模型的输入。这种方法有助于实现数据降维和交互式特征选择。

2.2 基于BPNN的预测模型

BPNN是一种多层前馈神经网络,根据误差反向传播算法进行训练。它具有全局逼近和高精度的优点,具有很强的非线性映射能力以及良好的自组织学习能力^[18-19]。标准BPNN由输入层、隐藏层(可以有多个隐藏层)、输出层组成^[20]。输入层接收输入数据,隐藏层负责对输入数据进行特征提取和映射,输出层给出最终的预测结果。其训练过程主要分为2个阶段:第1阶段为前向传播,输入信号通过各层之间的连接传递并经过激活函数处理,得到每1层的输出结果;第2阶段为反向传播,根据输出结果与实际标签之间的误差,

通过链式法则计算每个神经元的梯度,并利用梯度下降法来更新权重参数。在整个训练过程中,通过多次迭代来不断调整权重参数,使得网络对训练样本的预测结果逼近实际标签,并最小化损失函数。典型的BPNN原理架构如图3所示。其具体训练步骤^[21]为:

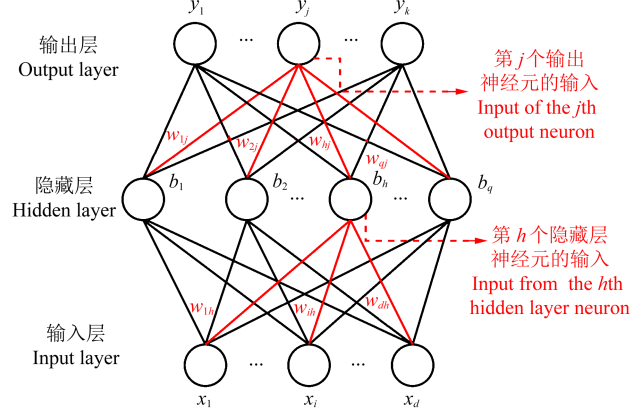


图3 反向传播神经网络基础架构

Fig. 3 BPNN infrastructure

1) 选取偏航系统正常运行状态下的数据作为训练集。

2) 初始化参数,初始化权重和偏置。

3) 前向传播。给定输入 $x_d = [x_1, x_2, \dots, x_{12}]^T$ 和实际输出 y_k 。计算预测输出 T_k , 并与实际输出(本文输出指偏航位置的时序数据)进行比较,计算误差。

输入层-隐藏层:

$$\alpha_h = \sum_{i=1}^d w_{ih} x_i + \theta_h \quad (5)$$

隐藏层-输出层:

$$\beta_j = \sum_{h=1}^q w_{hj} b_h + \theta_j \quad (6)$$

4) 反向传播。根据式(7)~式(9)更新权值和偏置,即

$$E_k = \frac{1}{2} \sum_{k=1}^m (y_k - T_k)^2 \quad (7)$$

$$w_{ij}^{(l)} = w_{ij}^{(l)} - \eta \frac{\partial E_k}{\partial w_{ij}^{(l)}} \quad (8)$$

$$\theta_{ij}^{(l)} = \theta_{ij}^{(l)} - \eta \frac{\partial E_k}{\partial \theta_{ij}^{(l)}} \quad (9)$$

式中, α_h 为第 h 个隐藏层神经元的输入; w 为权重; θ 为偏置; β_j 为第 j 个输出神经元的输入; E_k 为误差函数; $w_{ij}^{(l)}$ 为第 $l-1$ 层的第 j 个神经元与第 l 层的第 i 个神经元之间的连接权重; η 为学习率, $\eta \in (0, 1)$ 。

5) 反复执行前向传播和反向传播的步骤[步骤3)和步骤4)], 直到达到预设的停止条件(迭代次数、误差阈值)。

6) 异常检测。在训练完成后,使用训练好的BPNN预测模型对新的未标记数据进行异常检测。将

数据输入到已训练的网络中,根据输出层的输出值与实际值的残差进行异常判断。

BPNN 网络结构直接影响模型的表示能力、学习能力和计算效率。通过经验公式(10)确定初始隐藏层节点数,然后通过“试错法”逐步增加和减少节点数,找到使神经网络误差最小化的最终隐藏层节点数为 12。

$$p = \sqrt{m + n} + a \quad (10)$$

式中, p 为隐藏层节点数; m 为输入层节点数; n 为输出层节点数; a 为常数, $a \in (1, 10)$ 。

BPNN 具有良好的鲁棒性和泛化能力,但其在训练过程中容易受到学习率等参数的影响而陷入局部最优解,从而导致无法收敛到全局最优解。为提高 BPNN 预测模型的精确度,本文构建了 BPNN 预测模型的单隐藏层和双隐藏层,借鉴国内外常用的模型预测精度评估标准^[22],分别计算其均方误差(Mean Squared Error, MSE)、均方根误差(Root Mean Squared Error, RMSE)、平均绝对误差(Mean Absolute Error, MAE)、平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)以及相关系数 R^2 ,如表 2 所示。

表 2 单隐藏层和双隐藏层反向传播神经网络模型的预测精度
Tab. 2 Prediction accuracy of BPNN models with the single hidden layer and double hidden layers

隐藏层层数 Number of the hidden layer	MSE	RMSE	MAE	MAPE/%	R^2
1	0.130 4	0.356 2	0.276 1	5.12	0.816 5
2	0.098 7	0.214 2	0.201 4	4.02	0.902 3

由表 2 可知,相较于单隐藏层的 BPNN 模型,双隐藏层的 BPNN 模型在所有误差指标上都取得了更好的结果。例如,双隐藏层的模型相对于单隐藏层的模型,MSE、RMSE、MAE、MAPE 均减小了,而 R^2 增加了。这表明双隐藏层的模型具有更高的精度和更好的拟合效果。因此,本文选取隐藏层节点数为 12 的双隐藏层 BPNN 作为偏航系统状态参数预测模型。

2.3 偏航位置异常状态的量化

偏航系统在正常运行状态下的状态参数的残差分布特性反映了其处于正常状态下的特征。当偏航系统处于正常运行状态时,这些状态参数的残差应该在一个小范围内波动^[23]。然而,如果偏航系统发生异常,实际值与期望值之间的差异将增大,残差会变得较大。因此,通过监测偏航系统的状态参数残差,可以识别其异常特征。本文主要从偏航位置的残差分布特性出发,并结合其他状态参数辅助验证分析,识别偏航系统的异常模式。

如图 4 所示,以分位数 0.025 和 0.975 对预测参数

残差范围进行划分(0.025 和 0.975 对应着标准正态分布中的临界值,即正态分布的上、下 2.5% 的累积概率,如此可以更全面地考虑异常值情况,并保证模型在不同情况下的稳定性和有效性)。

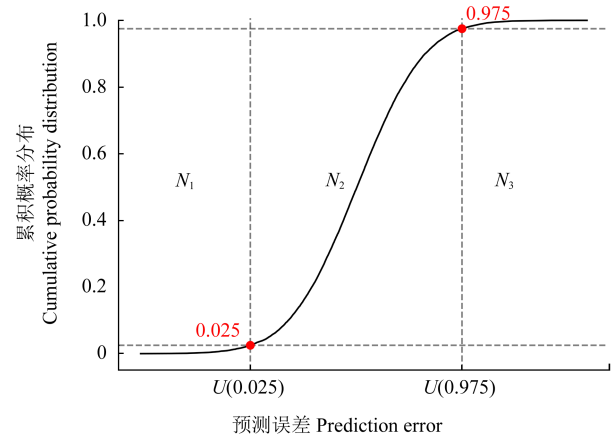


图 4 状态参数预测残差分布

Fig. 4 Residual distribution predicted by the state parameter

与时间序列分析、神经网络残差分析等方法相比,滑动窗口方法具有实时性、适应性,在实际工程应用中更加具有优势,且其原理相对简单。本文利用滑动窗口方法,通过分析残差从而实现对异常时序的精准定位。设定窗口长度为 10,步长为 3,分别统计预测参数在 N_1 、 N_2 、 N_3 中的个数,定义异常程度计算式为

$$e_{AI} = \frac{N_1 + N_3}{N_1 + N_2 + N_3} \quad (11)$$

式中, N_1 为残差落入区域 1 中的个数; N_2 、 N_3 同理。

偏航位置的异常程度与 e_{AI} 成正比,根据专家经验,本文判定当 e_{AI} 大于 0.6 时,偏航位置异常。

2.4 偏航位置异常识别算法流程

1) 数据采集。从 SCADA 系统中获取偏航位置数据,以及表 1 中提到的偏航系统的特征参数。这些将作为 BPNN 预测模型的输入。

2) 特征参数选择。从 SIG-Relief-F 筛选出特征参数作为模型的输入。

3) 状态参数预测。基于 BPNN 预测模型,利用输入参数预测偏航位置。计算 e_{AI} ,识别偏航位置异常时序。

4) 二次特征处理,归类真实的、特定的异常特征,对偏航位置真实异常情况进行总结。

5) 基于最终模型,通过大数据平台训练得到异常诊断结果。

6) 异常结果验证与分析。如确定为异常,应及时给现场预警并给出对应解决方案。

3 实例分析

3.1 基于BPNN模型的预测效果分析

通过SCADA数据对基于BPNN模型的预测效果进行有效性分析(将残差超出阈值的时序标记为异常片段)。

图5所示为某风场机组某时刻正常运行状态下共5000个样本点数据实际值与预测值的时序。其大部分数据的残差分布控制在 1.5° 以内,表明BPNN异常识别模型能较好地跟踪实际的运行状态。通过6台机组在2个月内正常运行状态下的数据结果得出该模型 F_1 值(准确率和召回率的综合评价指标)为0.912,说明该模型具有较高的识别精度。

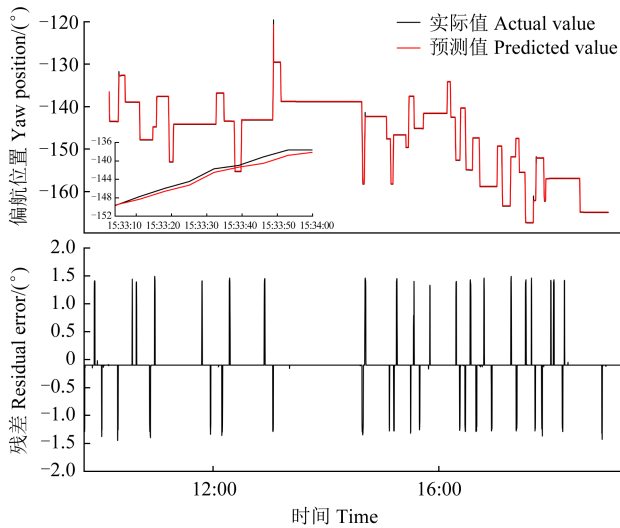


图5 模型预测效果

Fig. 5 Prediction effect of the model

3.2 异常案例分析及有效性验证

通过排查风电机组历史故障记录可知,在2022年10月6日,某风场3号机组偏航齿圈发生过损坏。为了验证模型的有效性,通过对偏航位置的异常指数 e_{AI} 计算可知,从2022年9月开始, e_{AI} 值大于0.6的次数为4次,在10月5日达到最大值0.841,由此可以初步判断,从2022年9月开始,偏航位置有异常趋势。初步

筛选出来的异常时序与实际异常存在部分“误差”,通过对现场工况的校核发现,部分“异常”仅仅是人工操作,导致数据与实际不一致,本文将这种情况称为“异常误报”。因此,通过构建具体异常二次特征得到偏航位置异常模式。图6为在正常情况下偏航位置的变化情况。纵轴表示偏航位置和偏航速度的变化率波动情况,横轴为这段数据的长度。

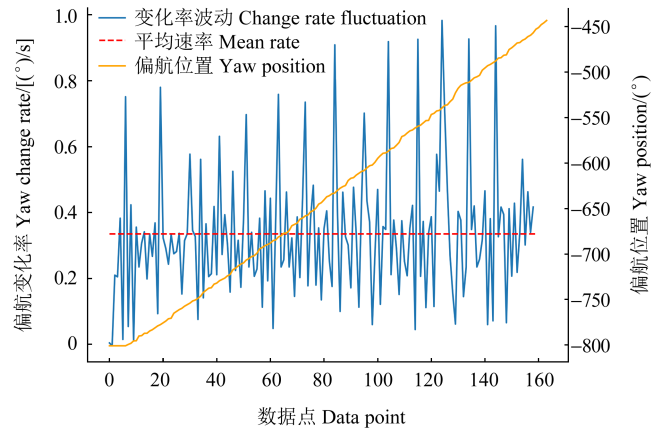


图6 正常情况下的偏航位置与速度趋势

Fig. 6 Yaw position and velocity trend under normal conditions

图7为模型经过二次特征加工后归类的4种偏航位置异常模式。经实际核查发现,这些时序下偏航系统部件均出现不同程度的损伤。由图7(a)可知,异常模式a的偏航位置在 310° 左右保持,偏航速率由 $0.3^\circ/s$ 跳变为 $0^\circ/s$ 附近。此模式异常特征逻辑为偏航任意位置长时间保持,且一个月内出现次数超过2次,保持时长超过60s。经过实际现场校核发现,异常产生原因为偏航齿圈与扭缆开关小齿啮合太紧导致偏航齿圈受到过大的力作用,这种过紧的啮合会增加摩擦和应力,最终导致偏航齿圈磨损。由图7(b)、图7(c)可知,异常模式b和c为电位计卡顿或电阻丝磨损,由图7(d)可知,异常模式d疑似为某机械故障。此模型通过识别具体的异常模式并进行提前预警,有效保障了偏航系统的安全稳定。模型的一些诊断结果如表3所示。

表3 异常诊断结果示例

Tab. 3 Example of abnormal diagnostic results

机组编号 Unit number	日期 Date	开始时间 Start time	结束时间 End time	异常起始位置 Abnormal start position/(°)	异常结束位置 Abnormal end position/(°)	异常模式 Abnormal mode
12 [#]	2022/11/04	06:48:33	06:51:07	-8.62	24.12	b
21 [#]	2022/11/12	11:17:44	11:18:21	9.21	9.32	a
49 [#]	2022/12/13	14:55:17	14:58:01	-17.95	-17.96	a
51 [#]	2022/12/21	09:47:04	09:49:04	-7.69	-13.04	d
72 [#]	2022/12/18	02:03:29	02:05:29	-2.6	-9.91	d
80 [#]	2022/12/31	20:02:20	20:02:34	10.1	60.7	c

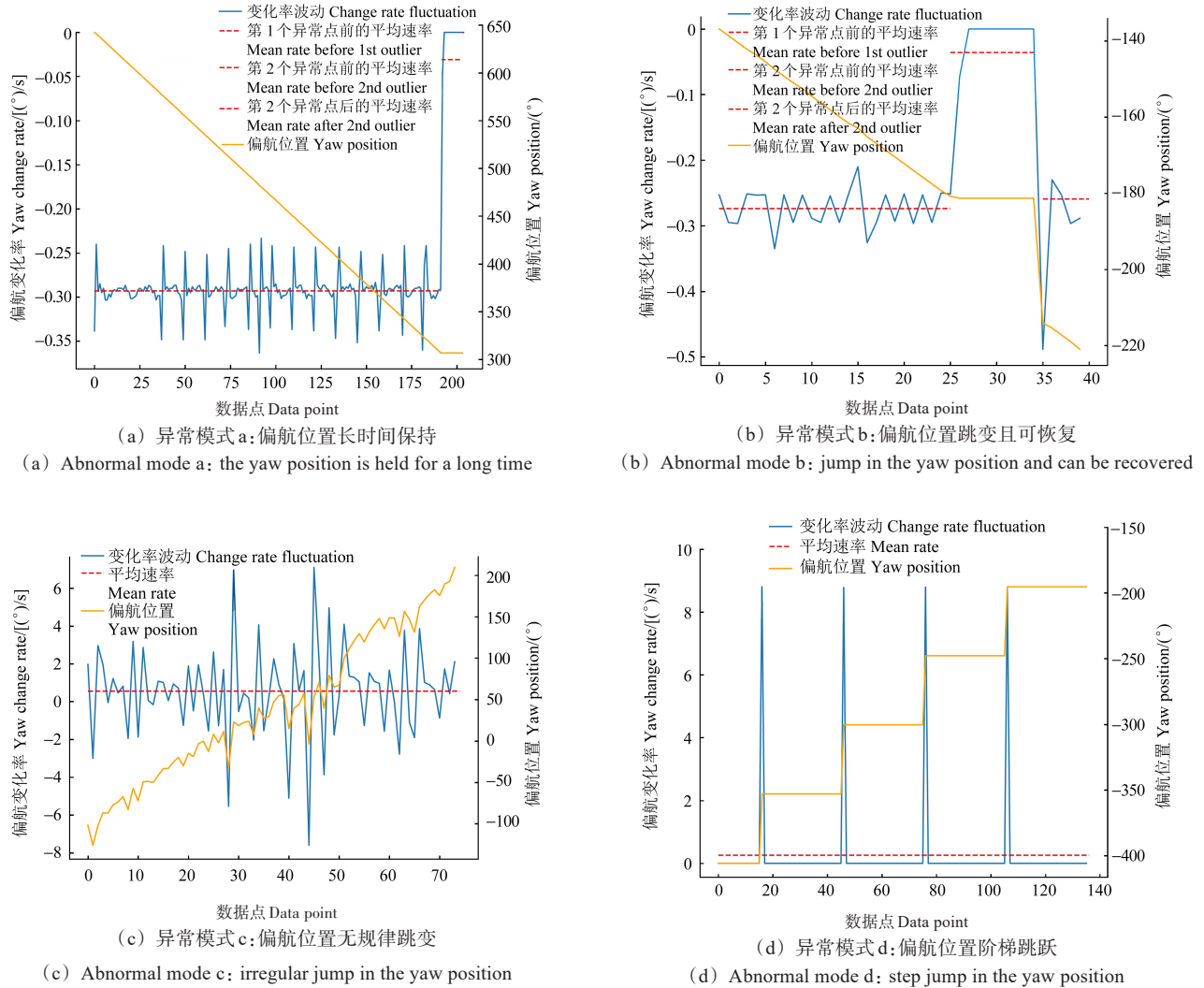


图 7 偏航位置异常模式时序

Fig. 7 Timing diagram of anomaly modes of the yaw position

3.3 不同方法的性能对比

为进一步验证所提方法的有效性,从该风场历史故障数据报表中选取了与偏航系统故障相关的 200 个异常样本,并与采用 MSET 模型、SVM 模型的传统方法进行对比,结果如表 4 所示。由表 4 可知,所提出的 SIG-Relief-F-BPNN 方法的 R^2 显著高于其他方法。经过实际工程应用,证明了该模型在机组运行过程中能够通过偏航位置的变化情况发现某些与偏航系统相关的潜在风险,从而进行提前预警,起到降低机组故障率并减少运维成本的作用。

表 4 不同方法性能评估

Tab. 4 Performance evaluation of different methods

方法 Method	MAE	RMSE	MAPE	R^2
BPNN	332.72	361.13	0.371	0.713
Relief-F-BPNN	310.21	335.28	0.281	0.763
SIG-Relief-F-SVM	241.31	262.17	0.133	0.851
SIG-Relief-F-MSET	258.44	287.49	0.161	0.833
SIG-Relief-F-BPNN	183.77	214.91	0.081	0.912

4 结论

考虑到偏航系统在部件发生异常的情况下,其相关状态参数会有一个异常数据的累积过程,针对偏航位置的状态变化表现,提出了一种基于数据驱动的偏航位置异常识别模型。经过实际案例验证,证明了其有效性与准确率。主要得出以下结论:

1) 与传统的数据降维方法相比,使用改进的 SIG-Relief-F 算法进行特征选择能够有效减少冗余特征对模型性能的影响,从而提高模型的准确性和可靠性。

2) 使用 BPNN 算法构建模型,可以充分考虑时间序列数据的非线性关系,进一步提高模型的预测准确性和可靠性。在实际应用中,与 MSET 和 SVM 方法相比,SIG-Relief-F-BPNN 方法表现出更高的准确性和更低的误报率。

3) 基于数据驱动的偏航位置异常诊断模型弥补了偏航系统故障诊断方法的单一性,提高了模型的泛

性。偏航位置的异常变化累积可以有效对相关机组进行提前预警,并根据风险等级安排相关现场人员进行安全排查,积累不同模式下的故障原因及解决经验,有效地降低了相关故障的运维成本,具有较高的实用性。

展望未来,基于数据驱动的异常识别技术将在风电机组故障诊断中发挥更重要的作用。首先,异常识别模型的精度和可靠性将得到进一步提高。目前已经有许多基于机器学习和深度学习的方法应用于风电机组故障诊断,但仍然需要进一步改善模型的性能。未来的研究可以通过结合多种算法和优化模型架构,提高异常识别模型的准确性和鲁棒性。其次,将模型与实际应用相结合,发展完善的预警和维护策略也是未来的方向。通过分析和识别风电机组的异常数据,可以提前预警潜在故障,并制订相应的维护计划。同时,还可以利用积累的故障数据和经验,优化运维流程、降低运维成本、提高风电机组的可靠性和效益。综上所述,基于数据驱动的风电机组异常诊断模型具有广阔的应用前景。未来的研究将继续改进算法和模型的性能,加强数据的处理和分析能力,为风能行业提供更准确、可靠的故障诊断和维护支持。

参考文献(References)

- [1] SANTOS A C, SOUZA W A, BARBARA G V, et al. Diagnostics of early faults in wind generator bearings using Hjorth parameters [J]. *Sustainability*, 2023, 15(20):14673.
- [2] 胡姚刚,刘怀盛,时萍萍,等. 风电机组偏航系统故障诊断与寿命预测综述[J]. *中国电机工程学报*, 2022, 42(13):4871-4883.
HU Yaogang, LIU Huaisheng, SHI Pingping, et al. Overview of fault diagnosis and life prediction for wind turbine yaw system [J]. *Proceedings of the CSEE*, 2022, 42(13):4871-4883. (In Chinese)
- [3] 沈小军,杜万里. 大型风力发电机偏航系统控制策略研究现状及展望[J]. *电工技术学报*, 2015, 30(10):196-203.
SHEN Xiaojun, DU Wanli. Expectation and review of control strategy of large wind turbines yaw system[J]. *Transactions of China Electrotechnical Society*, 2015, 30(10):196-203. (In Chinese)
- [4] 李辉,胡姚刚,李洋,等. 大功率并网风电机组状态监测与故障诊断研究综述[J]. *电力自动化设备*, 2016, 36(1):6-16.
LI Hui, HU Yaogang, LI Yang, et al. Overview of condition monitoring and fault diagnosis for grid-connected high-power wind turbine unit[J]. *Electric Power Automation Equipment*, 2016, 36(1):6-16. (In Chinese)
- [5] 杜浩飞,张超,李建军. 基于SENet-ResNext-LSTM的风机轴承故障诊断[J]. *机械强度*, 2023, 45(6):1271-1279.
DU Haofei, ZHANG Chao, LI Jianjun. Fault diagnosis of wind turbine bearing based on SENet-ResNext-LSTM[J]. *Journal of Mechanical Strength*, 2023, 45(6):1271-1279. (In Chinese)
- [6] BADIHI H, ZHANG Y M, JIANG B, et al. A comprehensive review on signal-based and model-based condition monitoring of wind turbines: fault diagnosis and lifetime prognosis[J]. *Proceedings of the IEEE*, 2022, 110(6):754-806.
- [7] ATTALLAH O, IBRAHIM R A, ZAKZOUK N E. CAD system for inter-turn fault diagnosis of offshore wind turbines via multi-CNNs & feature selection[J]. *Renewable Energy: An International Journal*, 2023, 203(Suppl. C):870-880.
- [8] 张海涛,高锦宏,吴国新,等. 蚁群优化算法在风电偏航故障检测中的应用[J]. *可再生能源*, 2013, 31(11):48-50.
ZHANG Haitao, GAO Jinhong, WU Guoxin, et al. Ant colony optimization applied in the fault detection of wind yaw[J]. *Renewable Energy Resources*, 2013, 31(11):48-50. (In Chinese)
- [9] 冯俊恒,刘晓辉,许波峰,等. 偏航偏差角对风力机轮毂载荷的影响[J]. *可再生能源*, 2023, 41(2):221-226.
FENG Junheng, LIU Xiaohui, XU Bofeng, et al. The studies of the influence of yaw deviation angle on hub loads of wind turbine[J]. *Renewable Energy Resources*, 2023, 41(2):221-226. (In Chinese)
- [10] 邓子豪,李录平,刘瑞,等. 基于SCADA数据特征提取的风电机组偏航齿轮箱故障诊断方法研究[J]. *动力工程学报*, 2021, 41(1):43-50.
DENG Zihao, LI Luping, LIU Rui, et al. Research on diagnosis method of wind turbine yaw gearbox based on SCADA data feature extraction[J]. *Journal of Chinese Society of Power Engineering*, 2021, 41(1):43-50. (In Chinese)
- [11] ZHAO H, ZHOU L W, ZHANG S W, et al. XE112-2000 wind turbine yaw strategy with adaptive yaw speed using DEL look-up table [J]. *IEEE Access*, 2021, 9:125724-125738.
- [12] ZHAO G L. Research on fault diagnosis and evaluation method of wind turbine gearbox[C]// *Proceedings of the IEEE 3rd International Conference on Electronic Technology*, 2023:976-980.
- [13] 宁文钢,姜宏伟,王岳峰. 风力发电机组偏航系统常见故障分析[J]. *机械管理开发*, 2018, 33(11):67-68.
NING Wengang, JIANG Hongwei, WANG Yuefeng. Common faults analysis of wind turbine yaw system[J]. *Mechanical Management and Development*, 2018, 33(11):67-68. (In Chinese)
- [14] 金晓航,孙毅,单继宏,等. 风力发电机组故障诊断与预测技术研究综述[J]. *仪器仪表学报*, 2017, 38(5):1041-1053.
JIN Xiaohang, SUN Yi, SHAN Jihong, et al. Fault diagnosis and prognosis for wind turbines: an overview[J]. *Chinese Journal of Scientific Instrument*, 2017, 38(5):1041-1053. (In Chinese)
- [15] WU C C, GUPTA J N D, CHENG S R, et al. Robust scheduling for a two-stage assembly shop with scenario-dependent processing times [J]. *International Journal of Production Research*, 2021, 59(17):5372-5387.
- [16] 李雨沛,王新利. 改进的ReliefF-BPNN分类模型[J]. *计算机时代*, 2023(6):20-24.
LI Yupei, WANG Xinli. Improved ReliefF-BPNN classification model[J]. *Computer Era*, 2023(6):20-24. (In Chinese)
- [17] WANG L X, JIANG S Y. A feature selection method via analysis of relevance, redundancy, and interaction [J]. *Expert Systems with Applications*, 2021, 183(Suppl. C):115365.
- [18] 赵耀,虞莉娟,苏义鑫,等. 基于聚类分析和Pearson相关系数法的电网负荷数据清洗与去重[J]. *船电技术*, 2023, 43(6):69-75.
ZHAO Yao, YU Lijuan, SU Yixin, et al. Grid load data cleaning

- and de-duplication based on cluster analysis and Pearson correlation coefficient method [J]. *Marine Electric & Electronic Engineering*, 2023, 43(6):69-75. (In Chinese)
- [19] 朱翔,谢峰. 基于核主成分分析BP-AdaBoost算法的数控铣床故障诊断[J]. *机械强度*, 2019, 41(6):1292-1297.
ZHU Xiang, XIE Feng. Tool wear state monitoring based on wavelet packet BP-AdaBoost algorithm [J]. *Journal of Mechanical Strength*, 2019, 41(6):1292-1297. (In Chinese)
- [20] 李国庆,刘钊,金国彬,等. 基于随机分布式嵌入框架及BP神经网络的超短期电力负荷预测[J]. *电网技术*, 2020, 44(2):437-445.
LI Guoqing, LIU Zhao, JIN Guobin, et al. Ultra short-term power load forecasting based on randomly distributive embedded framework and BP neural network[J]. *Power System Technology*, 2020, 44(2):437-445. (In Chinese)
- [21] 刘昌杰,段斌,张潇丹. 基于BPNN-NCT的风电机组主轴异常辨识方法[J]. *电力系统保护与控制*, 2022, 50(14):114-122.
LIU Changjie, DUAN Bin, ZHANG Xiaodan. An abnormal identification method for the main bearing of wind turbines based on BPNN-NCT [J]. *Power System Protection and Control*, 2022, 50(14):114-122. (In Chinese)
- [22] 王伟,王海云,黄晓芳. 基于Informer的风电机组叶根载荷预测[J]. *水力发电*, 2023, 49(9):85-89.
WANG Wei, WANG Haiyun, HUANG Xiaofang. Blade root load prediction of wind turbine based on Informer [J]. *Water Power*, 2023, 49(9):85-89. (In Chinese)
- [23] 徐硕,邓艾东,杨宏强,等. 基于改进残差网络的旋转机械故障诊断[J]. *太阳能学报*, 2023, 44(7):409-418.
XU Shuo, DENG Aidong, YANG Hongqiang, et al. Rotating machinery fault diagnosis method based on improved residual neural network [J]. *Acta Energetica Sinica*, 2023, 44(7):409-418. (In Chinese)

Research on data-driven abnormal warning methods for wind turbine yaw positions

SHEN Xu¹ WANG Haiyun¹ HUANG Xiaofang²

(1. Engineering Research Center of Renewable Energy Power Generation and Grid Connection Technology, Ministry of Education, Xinjiang University, Urumqi 830047, China)

(2. Beijing Goldwind Science & Creation Windpower Equipment Co., Ltd., Beijing 100176, China)

Abstract: Abnormal yaw positioning during yaw operations induces progressive deviation in yaw alignment accuracy, thereby compromising wind-tracking precision and risking excessive cable twisting that threatens operational safety. Concurrently, frequent position oscillations or repetitive short-duration position holding generate transient control errors, destabilizing the yaw control system. These coupled mechanisms collectively escalate yaw system failure frequency and operational maintenance costs. To proactively mitigate these risks, a data-driven fault diagnosis methodology is proposed for early detection of anomalous yaw positioning in wind turbines. Firstly, a large amount of data in a supervisory control and data acquisition (SCADA) system was processed using a standardized interaction gain and Relief-F (SIG-Relief-F) feature selection algorithm to identify multiple feature parameters with the strongest correlation with the target variable (which in this case may be yaw system failure). The advantage of this method lied in its ability to consider effectively the correlation between features, thus maximizing the retention of relevant features related to yaw system failures and interaction features. Secondly, a back propagation neural network (BPNN) yaw position prediction model was established, and the distribution of residuals was statistically analyzed using a sliding window method to determine the fault threshold. Finally, through empirical verification, the effectiveness and accuracy of the proposed method were demonstrated, and compared with multivariate state estimation technique (MSET) and support vector machine (SVM) algorithms, it was shown to have superior abnormal warning performance. The conclusions drawn can serve as a reference for the fault diagnosis of a practical yaw system.

Key words: Wind turbine; Yaw position; Interactive information; Relief-F; BPNN; Abnormal warning

Corresponding author: WANG Haiyun, E-mail: why@xju.edu.cn

Fund: Tianshan Talent Program (2022TSYCJC0030); Key Research and Development Program of the Autonomous Region (2022B03031); Science and Technology Project of Hami High-tech Zone (HGX2023KJXM008)

Received: 2024-02-21 Revised: 2024-03-20