

DOI: 10.19650/j.cnki.cjsi.12413415

基于回声状态网络的风电机组运行状态监测*

金晓航, 喻轩昂, 关汉林

(浙江工业大学机械工程学院 杭州 310023)

摘 要:“双碳”目标引领下,风能作为一种清洁可再生能源得到了广泛的利用。风电机组作为将风能转化为电能的重要装备,累计装机容量持续增长。然而,风电机组工作环境恶劣,运行工况多变,故障频发。为保障风电机组安全高效运行,亟需故障诊断与智能运维技术。针对风电机组运行工况复杂多变以及循环神经网络在时间序列学习过程中存在梯度消失和爆炸等问题,提出一种融合运行工况识别与回声状态网络的风电机组运行状态监测方法。首先,采用最大互信息系数对数据采集与监控系统(SCADA)采集的多维数据进行特征选择,筛选出与风电机组运行状态相关性高的特征。其次,利用K-means聚类算法构建机组的工况识别模型,对不同运行工况进行有效划分。然后,利用差分进化算法优化不同工况下的回声状态网络模型,增强其对复杂运行工况的适应能力,以此开展不同工况下风电机组有功功率预测。继而,结合功率预测残差分析确定相应的健康阈值,用于评判机组运行状态。最后,通过两个实际风电机组的案例研究表明,所提方法可有效监测机组的运行状态,当故障发生时,比SCADA系统提前发现机组运行状态的异常,可实现故障的早期预警。

关键词: 风电机组; 状态监测; 回声状态网络

中图分类号: TH17 **文献标识码:** A **国家标准学科分类代码:** 460.99

Condition monitoring of wind turbine based on echo state networks

Jin Xiaohang, Yu Xuan'ang, Guan Hanlin

(College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: Under the guidance of the “dual-carbon” goals, wind energy, as a clean and renewable energy source, has been widely harnessed. Wind turbines (WTs), which are crucial for converting wind energy into electrical energy, have seen a growing cumulative installed capacity. However, WTs operate in harsh environments with highly variable conditions, leading to frequent failures. To ensure their safe and efficient operation, fault diagnosis and intelligent maintenance technologies are urgently needed. Aiming to address the complex and variable operating conditions of WTs and the issues of gradient disappearance and explosion in recurrent neural networks during time-series learning, this paper proposes a condition monitoring method for WTs that integrates operating condition recognition with echo state networks (ESNs). First, the maximum mutual information coefficient is employed to select features from the supervisory control and data acquisition (SCADA) system data, prioritizing those with high relevance to the operational status of WTs. Second, the K-means clustering algorithm is utilized to construct a model for effective classification of different operating conditions. Subsequently, ESN models are optimized under various conditions using the differential evolution algorithm to enhance their adaptability to complex operating conditions, enabling active power prediction of WTs under different conditions. Then, by analyzing the residuals of power prediction, corresponding health thresholds are determined to assess the operating conditions of WTs. Finally, case studies of two actual WTs demonstrate that the proposed method can effectively monitor the operating status of WTs. It can detect abnormal operating conditions earlier than the SCADA system when faults occur, thus realizing early fault warning.

Keywords: wind turbine; condition monitoring; echo state networks

0 引 言

面对日益紧迫的能源危机,促进新能源领域的可持

续发展,风力发电机组(简称风电机组)作为将风能转换为电能的大型复杂装备,发挥着重要的作用^[1-2]。受外部环境与机组内部复杂机电特性的共同作用,风电机组的运行状态呈现出高度的多变性和不确定性,容易导致机

收稿日期:2024-10-27 Received Date: 2024-10-27

* 基金项目:国家重点研发计划(2022YFE0198900)、国家自然科学基金(62473336)、浙江省自然科学基金(LZ25F030004)项目资助

组发生故障。此外,风电机组通常被安装在风能丰富但环境恶劣的地区,这无疑增加了机组运维工作的难度和成本。据统计,大型风电场的运维成本高达其总收入的15%~35%^[1,3]。为有效降低运维成本,提升风电机组的运行稳定性和可靠性,积极开展状态监测等智能运维工作具有重要的意义和工程价值。

风电机组状态监测方法通常可分为3种类型:基于信号、物理模型和数据驱动的方法^[4]。其中,数据驱动方法无需增设额外传感器,同时在一定程度上摆脱了复杂物理模型的束缚,仅依靠历史数据便可实现风电机组状态监测。鉴于上述优势,该方法得到广泛的关注与研究。

随着人工智能技术的发展,深度学习方法在风电机组状态监测方面得到广泛应用^[5]。在深度学习众多算法中,循环神经网络(recurrent neural network, RNN)凭借其特有的循环连接架构,在时间序列数据分析上表现出色,常用于风速、功率等的预测^[6]。但是,RNN在训练时存在梯度消失和梯度爆炸等问题。为此,Wu等^[7-8]提出了一种基于长短期记忆网络(long short-term memory, LSTM)的风电机组状态监测方法。虽然LSTM能够有效捕捉长期依赖关系并解决梯度消失问题,但其计算复杂度较高^[9]。为了解决RNN训练速度慢和记忆消失方面的问题,Jaeger等^[10]提出了具有强大非线性时间序列建模能力和高效训练方式的回声状态网络(echo state network, ESN)。尽管ESN在功率^[11]和风速^[12]预测等方面已取得了较为显著的效果,但在风电机组状态监测方面,其应用尚未被有效拓展。考虑到风电机组的运行状态数据往往呈现出高度的非线性特征和时间相关性,将ESN引入并充分利用其优势,以提高风电机组状态监测的准确性和有效性。

风电机组服役过程中,环境因素的影响使得其运行状态复杂多变,且在不同的运行工况下,机组的运行参数存在较大差异。目前,大多数风电机组状态监测方法尚未充分考虑机组运行状态,仅依赖于单一的监测模型进行状态监测,这在一定程度上限制了状态监测的准确性和有效性。Jin等^[13]提出了一种基于马氏距离的新型集成方法,用于检测风力发电机中的异常情况;Zhan等^[14]提出了一种基于多向时空特征聚合网络的风电机组状态监测方法;Zhang等^[15]提出了一种基于双层集成(K-nearest neighbor, KNN)的风电机组齿轮箱的状态监测方法。此外,一些学者在研究风电机组状态监测时也关注到了机组动态变工况的运行方式。齐咏生等^[16]提出了一种基于风速区间划分运行工况和核主成分分析算法的风电机组状态监测与异常辨识模型;陈结等^[17]提出了一种基于工况辨识的风电机组主传动系统运行状态监测方法。虽然上述学者对风电机组不同的运行工况进行划分,但未对运行工况划分结

果进行充分的讨论与剖析。

综上所述,提出了一种基于回声状态网络的风电机组运行状态监测方法。主要贡献为:

1) 建立了一个多维数据采集与监控系统(supervisory control and data acquisition system, SCADA)特征输入的K-means聚类模型,该模型能够有效识别风电机组不同的运行工况,确保了机组状态监测的准确。

2) 提出了一种基于多工况识别和回声状态网络的风电机组运行状态监测方法。通过建立不同工况下的机组特有的正常行为模型,并确定相应的监测阈值,这一策略有效解决了单一模型在风电机组多变工况下难以适应的问题,通过实际案例的剖析验证了所提方法的有效性。

1 基于回声状态网络的风电机组运行状态监测方法

基于多工况识别和回声状态网络的风电机组运行状态监测方法主要包括以下4个步骤:特征工程、工况识别、功率预测和状态监测,其整体框架如图1所示。其中:1) 特征工程模块负责数据清洗、特征选择和归一化,旨在构建高质量的数据集,为后续机组运行状态监测模型的构建提供了数据基础;2) 工况识别模块利用K-means聚类算法对风电机组复杂多变的运行工况进行划分,并结合 t -分布随机邻域嵌入(t -distributed stochastic neighbor embedding, t -SNE)算法对聚类结果进行降维可视化,确保工况划分的合理性,工况识别模块解决了机组复杂运行工况难以识别的问题;3) 功率预测模块基于功能强大的回声状态网络,构建不同工况特有的功率预测模型,并结合预测功率残差分析确定各工况的监测阈值,动态阈值的设定能够有效避免复杂运行工况对状态监测结果的影响。4) 依托上一阶段构建的多工况状态监测模型,实现对风电机组运行状态的监测。

2 特征工程

风电SCADA数据通常具有数据量大、特征维度高的特点,并且常常伴随着噪声大、数据缺失等问题。为了改善数据质量、提升模型训练速度和预测性能,特征工程模块的作用显得尤为关键。该模块的核心环节主要包括数据清洗、特征选择以及归一化处理等多个方面。其中,特征选择环节能够有效选择出与风电机组运行状态相关性高的特征,从而降低特征数据的维度,减少噪声的干扰,进而提升模型训练的效率和预测的准确性。

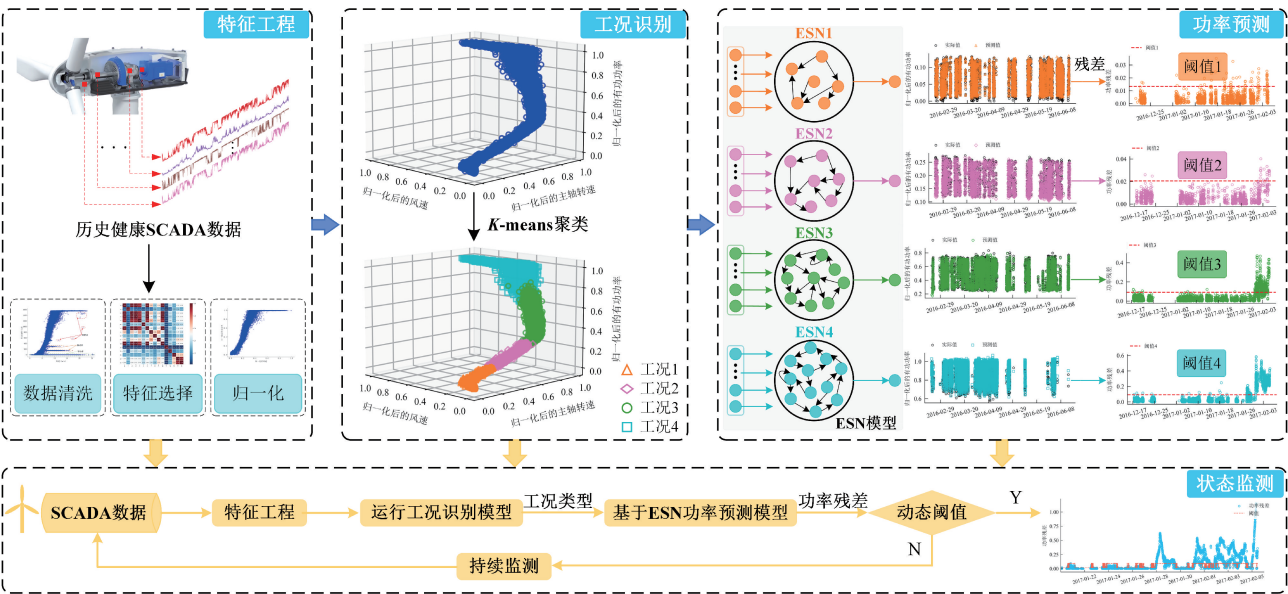


图1 基于多工况识别和回声状态网络的风电机组运行状态监测方法

Fig.1 Overall framework of wind turbine condition monitoring based on multi conditions identification and echo state networks

2.1 风电数据清洗

数据清洗^[18]主要针对风电 SCADA 数据中存在的以下两类问题进行处理:1) 对于缺失值、零功率、限功率以及位于切入风速与切出风速之外的数据,采取直接删除的策略;2) 针对离群点数据,采用局部异常因子 (local outlier factor, LOF) 算法进行检测与剔除。LOF 算法的核心在于,它依据数据点的局部密度来评估其相对于周围邻域的异常程度,从而有效识别出离群点^[19]。风电数据清洗前后的对比如图 2 所示。

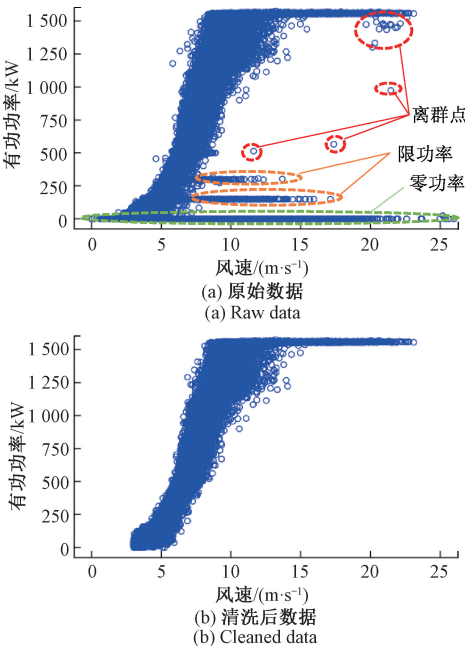


图2 风电数据清洗前后对比

Fig.2 Comparison of wind data before and after data cleaning

2.2 风电特征选择

SCADA 数据特征维度高,其中包含了一些冗余和无关的特征。通过特征选择,能够有效地从高维数据中选择与机组运行状态密切相关的特征,提升模型训练速度和预测准确性。鉴于最大互信息系数 (maximal information coefficient, MIC) 具有较低的计算复杂度和较高的鲁棒性,适用于 SCADA 非线性数据之间的相关性计算^[17,20],选择 MIC 作为分析工具,度量两个变量 X 、 Y 之间非线性关系, MIC 的计算公式为:

$$I(X;Y)=\iint p(x,y)\log_2\left(\frac{p(x,y)}{p(x)p(y)}\right)dx dy \tag{1}$$

$$R_{MIC}=\max_{a \times b \leq n} \frac{I^*(X;Y)}{\log_2(\min(a,b))} \tag{2}$$

式中: $p(x)$ 、 $p(y)$ 为概率密度函数; $p(x,y)$ 为联合概率密度函数; $I^*(X;Y)$ 为所有网格划分中最大互信息值; a 、 b 分别为 x 、 y 方向上的网格数; n 为样本数量; k 通常设为 0.6。 R_{MIC} 的取值范围在 $[0,1]$, 越接近 1, 表示变量 X 与 Y 间的相关性越高, 反之亦然。

风电机组的有功功率是反映其运行状态的重要指标,通过计算 SCADA 特征数据与有功功率之间的 MIC 值,分析各特征与有功功率之间的相关性,如图 3 所示。电网电压 1、2、3、发电机驱动端轴承温度以及发电机的横向和纵向振动实时值与有功功率的相关性较低。电网电流 1、2、3 之间呈现出高度的相关性,表明它们所包含的信息存在冗余。因此,为了简化分析,仅选择电网电流 1 相作为代表性参数进行后续研究。表 1 详细列出了各 SCADA 特征与有功功率之间的 MIC 值,最终筛选出 MIC 值 >0.5 的特征作为模型的输入特征。这一环

节确保了模型能够聚焦于与有功功率密切相关、对运行状态具有显著影响的特征,进而提高模型预测的准确性。

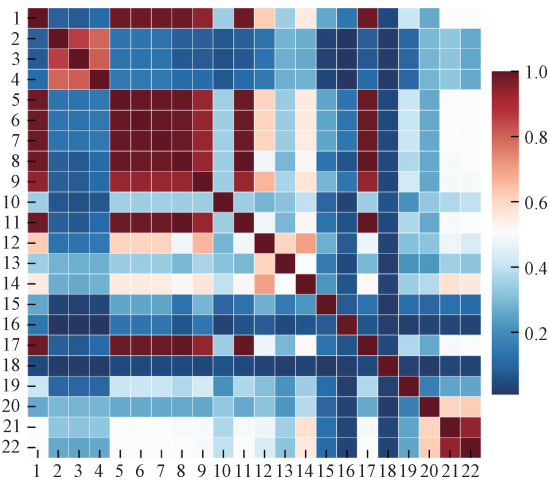


图 3 SCADA 特征之间的 MIC 值热力图

Fig. 3 Heatmap of MIC values between SCADA features

表 1 SCADA 特征之间的 MIC 值列表

Table 1 MIC values between SCADA features

编号	特征	MIC 值
1	有功功率	1.00
2	电网电压 1	0.09
3	电网电压 2	0.08
4	电网电压 3	0.12
5	电网电流 1	0.98
6	电网电流 2	0.98
7	电网电流 3	0.98
8	变流器电机转速	0.98
9	平均风速 30 s	0.93
10	平均风向 30 s	0.35
11	发电机转速	0.98
12	发电机绕组 1 温度	0.62
13	发电机驱动端轴承温度	0.35
14	发电机非驱动端轴承温度	0.55
15	横向振动实时值	0.26
16	纵向振动实时值	0.14
17	主轴转速	0.98
18	偏航夹角	0.04
19	齿箱进口压力	0.41
20	齿轮箱油温	0.26
21	齿箱前端轴承温度	0.50
22	齿箱后端轴承温度	0.51

3 风电机组运行工况识别

由于风电机组运行工况复杂多变,单一模型难以准确捕捉机组的运行特性。因此,为提高状态监测模型的准确性,需要对机组的不同运行工况进行有效识别。本节采用 K -means 聚类算法对机组运行数据进行工况划分,并通过原始高维数据的可视化结果,从侧面验证了聚类分析的合理性和有效性。

K -means 聚类算法^[21]是一种迭代求解的聚类分析方法,具有原理简单、计算效率高、收敛速度快等特点,该算法的主要思路是将数据集中的样本点划分为 K 个不同的簇,使得簇内的样本点之间的距离尽可能小,而不同簇之间的距离尽可能大。 K -means 聚类通过手肘法确定最佳聚类中心数目 K 。

手肘法的核心评价指标是误差平方和 (sum of squared errors, SSE), 计算公式为:

$$SSE = \sum_{i=1}^K \sum_{j=1}^{C_i} \|x_j - m_i\|^2$$

(3)

式中: K 为聚类中心数目; C_i 为第 i 个簇的样本数目; m_i 为第 i 个簇的簇心; $\|x_j - m_i\|^2$ 为第 i 个簇中各样本点到其簇心的欧式距离。

SSE 通常会随着聚类中心数目 K 的增加而减少,因为更多的聚类中心意味着每一类中的样本点更接近该类的中心,但存在一个拐点。在拐点之前,增加聚类数目会显著减少 SSE ,但在拐点之后,继续增加聚类数目对于减少 SSE 的贡献就显得不那么突出。 SSE 与 K 的关系图呈现一个手肘的形状,而这个肘部所对应的 K 值即为数据的最佳聚类中心数目。同时随着聚类数目的增加, K -means 聚类算法的收敛计算耗时也将显著增加。

将特征工程环节中选择的多维 SCADA 数据输入到 K -means 聚类模型中,对机组不同运行工况进行识别与划分。通过绘制 SSE 与 K 值的关系,确定机组数据的最佳聚类中心数目 K 为 4,聚类效果到达一个拐点,如图 4 所示。因此,将风电机组的运行状态划分为 4 种工况,不同工况数据的分布情况,如图 5 所示。利用 K -means 聚类模型进行风电机组运行工况识别,一方面综合考虑了风电机组运行特性,另一方面为后续多工况模型的构建以及运行状态监测提供有效支持。

为了进一步挖掘风电机组 SCADA 数据的内在结构特征,简化数据的理解和分析流程,引入了 t -SNE 算法^[22]。该算法能够将高维输入特征有效地映射到低维空间中,进而在二维平面上清晰地呈现出不同运行工况之间的边界,如图 6 所示。这一可视化结果验证了 K -means 聚类算法将风电机组运行数据划分为 4 类的合理性。

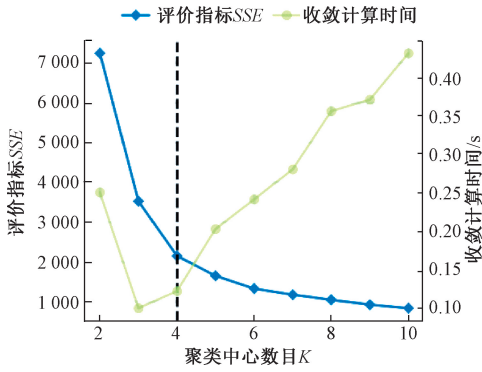


图4 SSE与K值的关系

Fig. 4 Relationship between SSE and K value

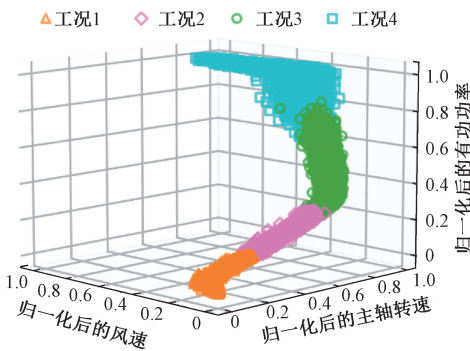


图5 案例1风电机组数据的聚类结果

Fig. 5 Clustering results of wind turbine data in case 1

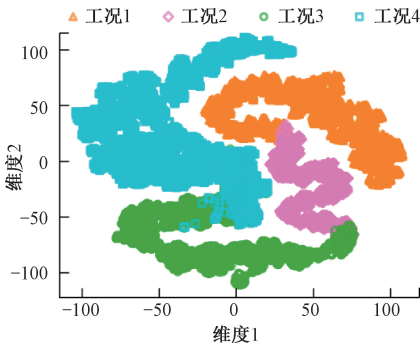


图6 案例1风电机组数据的降维可视化结果

Fig. 6 Visualization results of wind turbine data in case 1

4 风电机组有功功率预测

利用工况识别模块,划分风电数据集;针对不同运行工况构建相应的ESN模型,预测风电机组有功功率,结合功率残差分析确定其相应的监测阈值;根据不同工况下的阈值分析并判断风电机组的实时运行状态。

4.1 ESN模型

ESN模型是一种新型的RNN改进模型^[23],解决了传统RNN模型中存在的梯度消失和爆炸问题,该模型中包含一个输入层、一个动态储备池和一个输出层,其结构如图7所示。其中输入层用于接收输入信号;动态储备池替代了传统RNN的隐含层,储备池中包含了多个随机稀疏连接的神经元,用于处理输入层的信号和储备池自身上一时刻的信号;输出层用于产生输出信号。ESN模型中,假设输入层有 M 个神经元,储备池有 N 个神经元,输出层有 L 个神经元。在 t 时刻,输入层状态、储备池状态和输出层状态分别如式(4)~(6)所示。

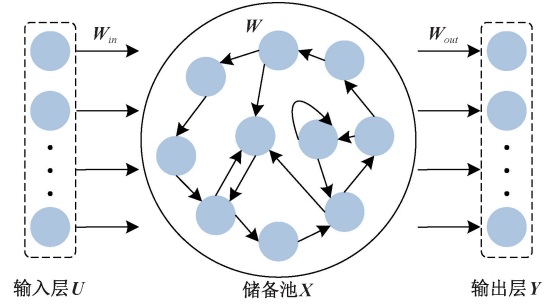


图7 ESN网络结构

Fig. 7 Structure of ESN network

$$\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_M(t)]^T \quad (4)$$

$$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T \quad (5)$$

$$\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_L(t)]^T \quad (6)$$

在第 $t+1$ 时刻,储备池状态按照式(7)进行更新,输出层状态按照式(8)进行更新。其中,输入层到储备池的连接权值为 $\mathbf{W}_{in}$,储备池内部连接权值为 $\mathbf{W}$,储备池到输出层的连接权值为 $\mathbf{W}_{out}$ 。

$$\mathbf{x}(t+1) = f(\mathbf{W}_{in} \cdot \mathbf{u}(t+1) + \mathbf{W} \cdot \mathbf{x}(t)) \quad (7)$$

$$\mathbf{y}(t+1) = g(\mathbf{W}_{out} \cdot [\mathbf{u}(t+1); \mathbf{x}(t+1); \mathbf{y}(t)]) \quad (8)$$

式中: f 和 g 分别是储备池内部和输出层的神经元激活函数。

ESN训练过程中输入层连接权值 $\mathbf{W}_{in}$ 和储备池内部连接权值 $\mathbf{W}$ 在随机初始化后保持不变,只有输出层连接权值 $\mathbf{W}_{out}$ 需要通过训练得到。将训练数据依次输入到网络中,根据式(7)和(8),计算每一时刻的储备池状态和输出层状态。在训练迭代阶段,收集储备池状态矩阵 $\mathbf{X}$ 和期望输出矩阵 $\mathbf{Y}$,使ESN的实际输出逼近期望输出。将上述问题转换成求解最小二乘法问题,即可得到输出层连接权值 $\mathbf{W}_{out}$ 计算公式如式(9)所示。

$$\mathbf{W}_{out} = [(\mathbf{X}^T \mathbf{X})^{-1} \times (\mathbf{X}^T \mathbf{Y})]^T \quad (9)$$

4.2 ESN 模型中的关键参数

ESN 模型预测性能主要受储备池 3 个参数包括储备池神经元个数 N 、稀疏度 α 和谱半径 ρ 的影响。

1) 储备池神经元个数 N 通常设定为训练样本数量的 $1/10\sim 1/2$,与输入样本数量及问题复杂性相关。一般而言,较大的 N 能更好地表征复杂动态系统,但过大可能导致过拟合。

2) 稀疏度 α 表示储备池中相互连接的神经元占总神经元数的比例,反映神经元之间的连接情况。稀疏连接有助于降低计算量,提高 ESN 的运行效率, α 通常取值为 $1\sim 10\%$ 。

3) 谱半径 ρ 代表储备池内部连接权值矩阵特征值绝对值的最大值,决定网络的记忆能力。过小会削弱记忆能力,过大会影响状态稳定性。通常 $\rho<1$ 以确保回声状态特性,即储备池状态经过迭代后仅与历史输入和输出相关,通常取值在 $0.1\sim 0.99$ 。

5 风电机组案例分析与讨论

为验证所提方法的有效性,将其应用于两台实际风电机组的状态监测中。案例 1 是一台 1.5 MW 双馈式风电机组,机组 SCADA 系统于 2017 年 2 月 5 日 11:20 发出发电机轴承温度过高的异常报警。经运维工作人员检修后,发现该报警是由于发电机轴承发生跑圈故障所致。案例 2 是另一台 1.5 MW 双馈式风电机组,该机组在 2017 年 6 月至 8 月期间正常运行。上述两台机组的 SCADA 系统每隔 5 min 进行一次数据的采样。

5.1 基于 DE 优化 ESN 的风电机组有功功率预测

储备池参数对回声状态网络的预测性能影响显著,因此,需进行合理的参数设置。差分进化算法(differential evolution,DE)作为一种高效的连续型优化工具,通过模拟自然界的进化机制,如种群初始化、变异、交叉和选择等步骤,不断迭代进化,以搜寻全局最优解。本文采用 DE 来优化设置这些关键参数^[23]。

DE-ESN 参数优化流程为:

1) 解码方案。DE 种群中每个个体的基因维度 D 为 3。 N 的搜索空间是根据训练样本的数量和机组各工况的复杂程度设定。为减少搜索空间的规模,降低算法的计算复杂度,以 100 为间隔离散化 N 的搜索空间; α 的搜索空间是 $[0.01, 0.1]$; ρ 的搜索空间是 $[0.1, 0.99]$ 。

2) 适应度函数。DE 种群中的每个个体,以 ESN 模型的预测值 $\hat{y}_i$ ($i=1,2,\cdots,k$; k 表示输出样本的数量)和实际值 y_i 之间的差作为适应度值,选择均方根误差(root mean squared error, RMSE)为适应度函数,第 i 个个体的适应度值如式(10)所示。

$$f_i = \sqrt{\frac{1}{k} \sum_{i=1}^k (y_i - \hat{y}_i)^2}$$

(10)

3) 优化流程。DE-ESN 模型的优化流程具体步骤为:

(1) 种群初始化。设置种群规模 NP 、基因维度 D 、基因取值范围 $[U_{\min}, U_{\max}]$ 、变异算子 F 、交叉概率 CR 、预测精度 μ 和最大迭代次数 m 。生成初始种群后,设置迭代次数 $G=0$ 。

(2) 迭代终止条件判定。若全局最优值达到预测精度 μ 或者迭代次数达到最大迭代次数 m ,终止 DE 迭代,转至(7);否则,执行(3)。

(3) 执行变异、交叉和选择,生成下一代个体。

(4) 重复(3),直至生成下一代种群。

(5) 计算下一代种群的适应度值,确定当前全局最优值及其对应的全局最优个体。

(6) 使 $G=G+1$,然后返回(2)。

(7) 将 DE 种群中的最优个体解码,映射为 ESN 的 3 个关键参数值。

(8) 利用训练集训练 ESN,获得训练好的网络。

(9) 输入测试集进行预测。

各工况下 DE-ESN 模型的参数设置,如表 2 所示。

表 2 各工况下 DE-ESN 模型的参数设置
Table 2 Parameters of the DE-ESN model under different operating conditions

	工况 1	工况 2	工况 3	工况 4
NP	30	30	30	30
m	50	50	50	50
DE μ	0.0001	0.000 1	0.000 5	0.000 5
F	0.8	0.8	0.8	0.8
CR	0.2	0.2	0.2	0.2
N	[1 000,2 000]	[600,1 000]	[1 200,2 000]	[1 500,2 200]
ESN α	[0.01,0.1]	[0.01,0.1]	[0.01,0.1]	[0.01,0.1]
ρ	[0.1,0.99]	[0.1,0.99]	[0.1,0.99]	[0.1,0.99]

为验证 DE-ESN 模型功率预测的有效性,采用 ESN、LSTM^[24]和门控循环单元(gated recurrent unit, GRU)^[25]等模型进行对比实验。并以均方根误差、平均百分比误差(mean absolute percentage error,MAPE)和决定系数(r -square, R^2)作为评价预测性能的指标。该实验所用电脑配置:CPU 为 Intel Core i5-7 400,运行内存 8 G,操作系统为 window10。表 3 为风电机组不同运行工况下 DE-ESN 模型的最优参数及其对应的评价指标。表 4 为案例 1 机组运行在工况 4 下不同模型评价指标的对比,可以看到 DE-ESN 模型的预测性能优于其他 3 个模型,ESN 模型的训练时间明显小于 LSTM 和 GRU 模型。图 8 为案例 1 机组不同工况下,风电机组有功功率的预测结果。通过上述结果可以看到 DE-ESN 模型在风电机组功率预测方面具有一定优势。

表 3 案例 1 风电机组不同工况下基于 DE-ESN 模型的最优参数及其评价指标

Table 3 Optimal parameters and evaluation metrics of DE-ESN model for wind turbine under different operating conditions in case 1

工况类型	储备池最优参数	RMSE	MAPE	R ²
1	[1 400, 0.06, 0.73]	0.003 8	4.838	0.983
2	[800, 0.02,0.8]	0.005 5	2.358	0.981
3	[1 500, 0.05,0.62]	0.024 2	3.642	0.968
4	[1 800, 0.03, 0.64]	0.025 4	2.041	0.938

表 4 案例 1 风电机组工况 4 下 DE-ESN 与其他模型评价指标的对比分析

Table 4 Comparative results of different models under operating condition 4 of wind turbine in case 1

模型	训练时间/s	RMSE	MAPE	R ²
DE-ESN	5.3	0.025 4	2.041	0.938
ESN	3.6	0.027 8	2.294 7	0.936
LSTM	63.6	0.034 1	3.092 1	0.904
GRU	36.8	0.031 6	2.798 1	0.918

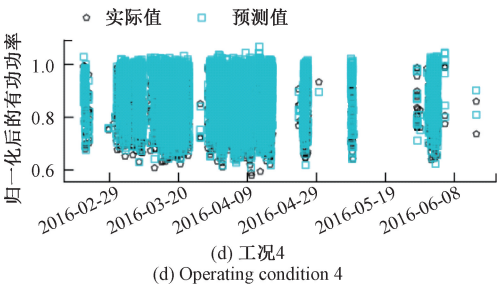
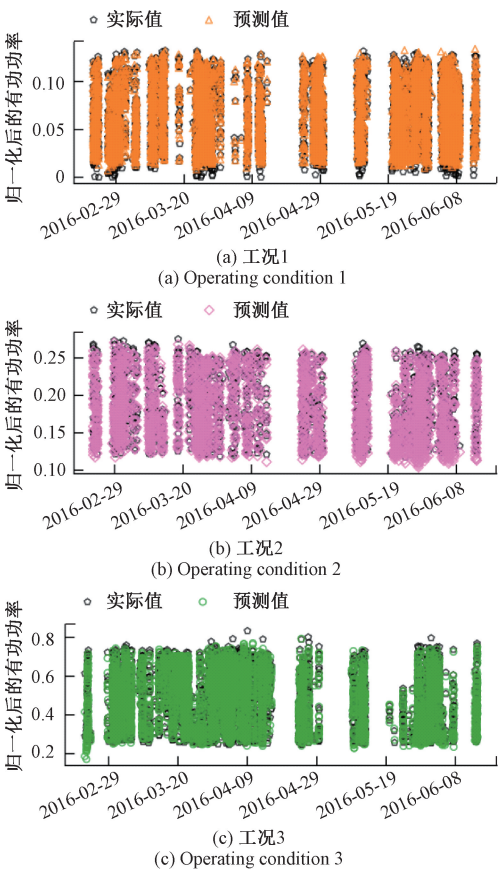


图 8 案例 1 不同工况下基于 DE-ESN 的功率预测结果
Fig. 8 Power prediction results based on DE-ESN model under different operating conditions in case 1

5.2 案例 1-故障风电机组

选取该机组 2016 年 2 月~6 月正常运行时段的数据作为训练集,故障发生前 1 个半月(2016 年 12 月 15 日~2017 年 2 月 5 日)的数据作为测试集。以预测功率与实际功率的残差作为反映机组运行状态的健康指标。

鉴于风电机组运行工况复杂多变,在状态监测任务中采用了动态阈值方法。该方法基于不同工况分别构建风电机组的正常行为模型,并为每个工况确定相应的监测阈值,进而有效避免复杂工况对监测结果的影响。根据功率预测结果,结合统计学方法^[13,26],确定不同运行工况下对应的监测阈值,如表 5 所示。

表 5 案例风电机组不同工况下的动态监测阈值

Table 5 Dynamic monitoring thresholds for wind turbines under different operating conditions in case studies

工况类型	案例 1	案例 2
1	0.013 1	0.019 1
2	0.016 8	0.014 4
3	0.088 4	0.062 9
4	0.091 0	0.078 4

为验证运行工况划分和动态阈值设定对机组状态监测结果的影响,将所提方法与未进行工况划分的 ESN 状态监测模型进行了对比分析,结果如图 9 所示。对比结果表明,工况划分能够有效减少状态监测中的误报警的次数,并且能更早地识别机组的异常状态,从而提高监测的准确性。从图 9(b)中可以看到,故障发生的早期阶段出现了少量误报警。

为进一步分析风电机组状态监测结果,图 10 显示了不同工况下的监测结果。图 10(a)和(b)是机组在低功率运行状态下的状态监测结果,可以看出,故障发生时超出阈值的数据点较少。图 10(c)和(d)是机组在高功率运行状态下的状态监测结果,明显出现了大量超出阈值的数据点,机组故障的异常状态表征也更加明显,并且

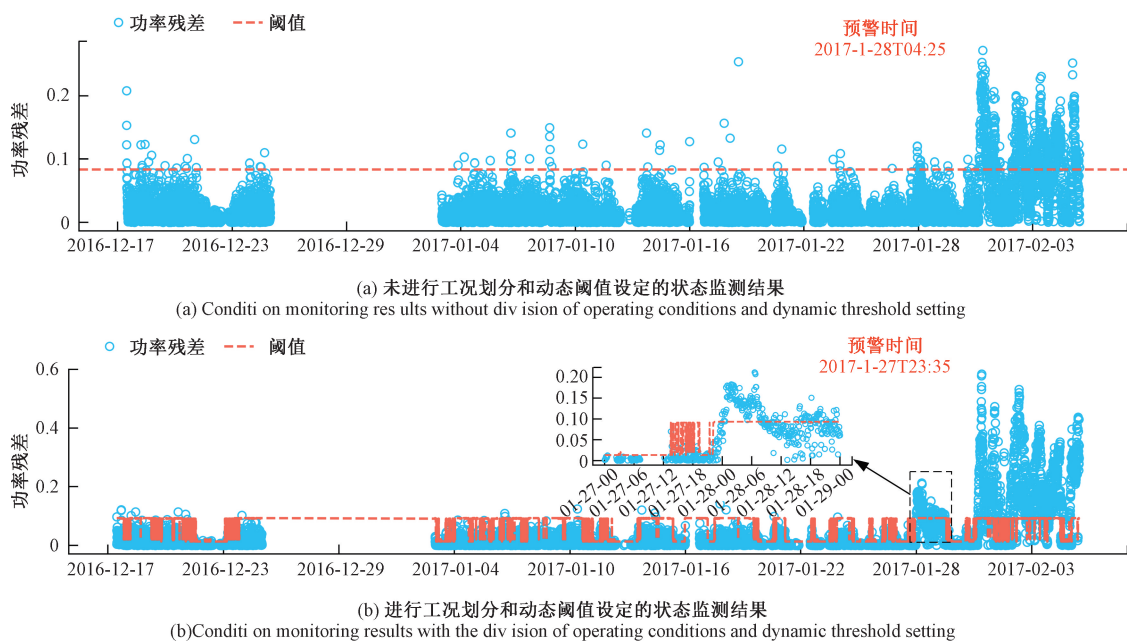


图9 案例1风电机组运行状态的监测结果

Fig. 9 Condition monitoring results of wind turbine in case 1

监测阈值明显高于低工况状态的阈值。造成这一现象的原因是高功率运行时机组面临更大的负载和应力,导致机械部件间的摩擦加剧,温度升高,从而增加了超阈值监测数据出现的概率。

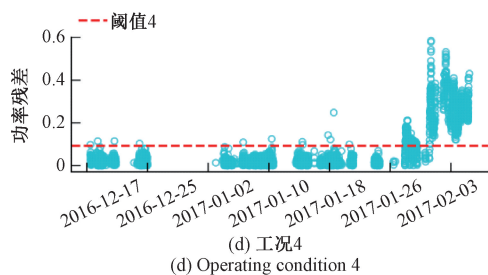
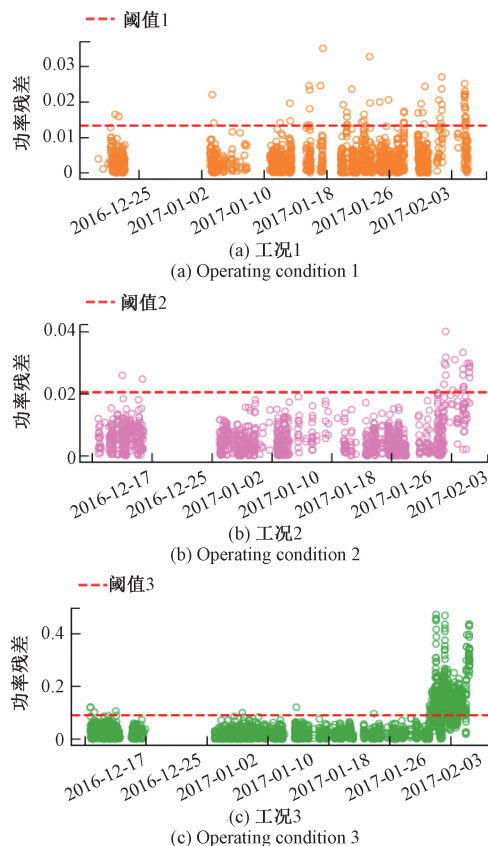


图10 案例1风电机组不同工况下状态监测结果

Fig. 10 Condition monitoring results of wind turbine under different operating conditions in case 1

针对机组故障发生早期阶段出现的少量超阈值现象。本研究对部分超出阈值的数据点进行了统计,深入分析了这些时刻的原始 SCADA 数据,如表 6 所示。例如,机组在 2017 年 1 月 17 日 12 点 15 分时预测功率残差超出监测阈值,该时刻风速为 4.048 m/s、风向为 178.289°,而其前 5 min 的风速为 6.032 m/s、风向为 235.130°。通过对比原始数据可见,机组在该时刻所处的环境(如风速、风向)发生了显著变化。通过分析发现,大多数超阈值点通常出现在风速或风向急剧变化的时刻。这种现象由机组的动态响应特性与控制系统复杂性等内外因素共同作用引起,导致功率预测偏差较大,从而引发误报警。因此,需要设置合适的预警策略,以降低这种异常对风电机组状态监测结果的影响。根据风电机组的运维经验,将连续 3 次超阈值设定为预警条件^[27]。本研究所提方法在 2017 年 1 月 27 日 23 点 35 分时首次

出现了 3 个连续超出阈值的异常点,满足了异常预警条件。该时刻较 SCADA 系统异常报警时间提前了 8 天,较好的验证了该方法的有效性。

表 6 案例 1 风电机组各工况下部分超阈值及其前一时刻的风速、风向数据

Table 6 Some original wind speed and wind direction data of wind turbine under different operating conditions in Case 1

工况类型	时间	风速/($\text{m}\cdot\text{s}^{-1}$)	风向/($^{\circ}$)
1	2017-01-17T12:10	6.032	235.130
	2017-01-17T12:15	4.048	178.289
	2017-01-20T13:20	4.196	173.248
	2017-01-20T13:25	4.772	218.401
2	2016-12-23T16:50	8.925	225.095
	2016-12-23T16:55	5.772	227.639
	2017-01-17T11:55	2.73	183.591
	2017-01-17T12:00	6.686	225.604
3	2016-12-18T11:40	9.785	264.113
	2016-12-18T11:45	7.861	274.886
	2017-01-08T15:30	8.197	288.431
	2017-01-08T15:35	6.805	250.697
4	2017-01-14T9:50	9.027	271.944
	2017-01-14T9:55	7.602	278.091
	2017-01-18T2:45	10.045	263.775
	2017-01-18T2:50	8.627	265.615

5.3 案例 2-健康风电机组

选取该机组 2017 年 6 月~7 月的 SCADA 数据作为训练集,7 月~8 月 SCADA 数据作为测试集。按照图 1 所提方法开展风电机组运行状态监测。首先,对训练集数据进行特征工程处理,构建高质量的机组正常运行数据集。其次,将预处理后的 SCADA 数据输入到 K -means 聚类模型中,实现机组运行工况的划分,结果如图 11 所示。为了更直观地理解数据分布,采用 t -SNE 算法对高维 SCADA 数据进行降维可视化,如图 12 所示。图 12 结果表明,机组的 4 类运行工况数据在二维空间中也呈现出较好的连续性和区分度,进一步验证了运行工况划分的合理性。然后,针对不同运行工况的数据,分别构建了相应的 DE-ESN 功率预测模型,结合功率预测残差分析确定各工况的监测阈值,如表 5 所示。模型训练阶段,机组各工况下风电机组功率预测结果,如图 13 所示。最

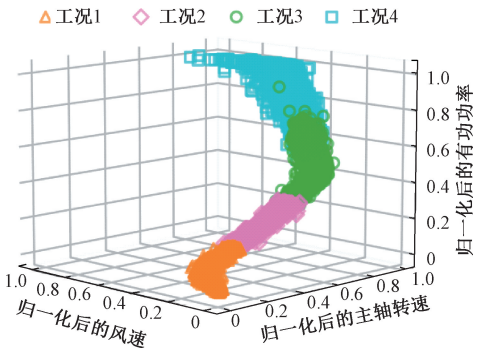


图 11 案例 2 风电机组数据的聚类结果

Fig. 11 Clustering results of wind turbine data in case 2

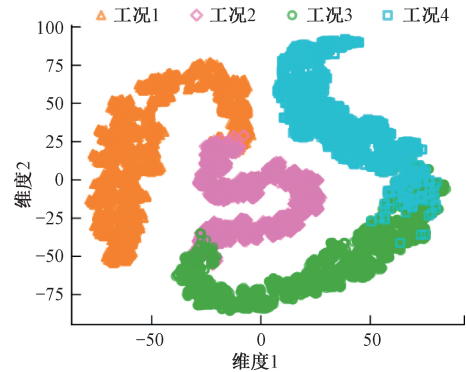
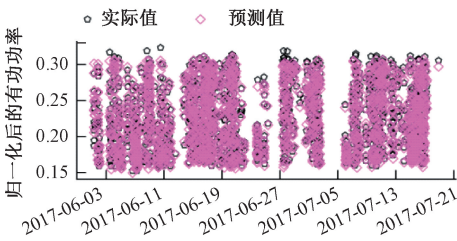


图 12 案例 2 风电机组数据的降维可视化结果

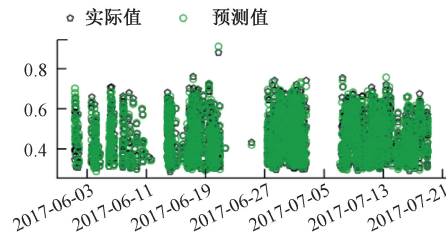
Fig. 12 Visualization results of wind turbine data in case 2



(a) Operating condition 1



(b) Operating condition 2



(c) Operating condition 3

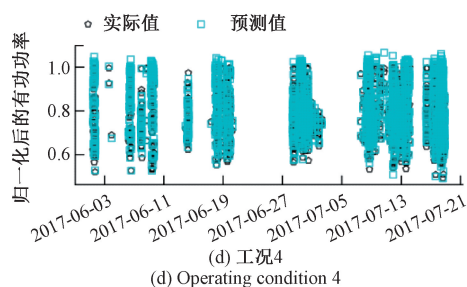


图 13 案例 2 不同工况下基于 DE-ESN 的功率预测结果

Fig. 13 Power prediction results based on DE-ESN model under different operating conditions in case 2

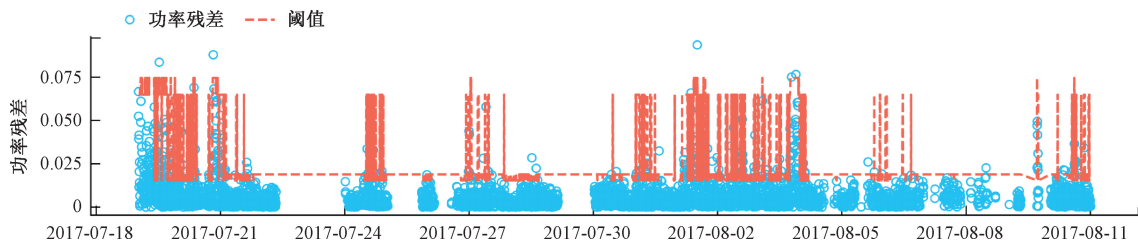


图 14 案例 2 风电机组运行状态的监测结果

Fig. 14 Condition monitoring results of wind turbine in case 2

6 结 论

针对风电机组运行状态难识别以及 RNN 在时间序列预测中的局限性,提出了一种融合运行工况识别与回声状态网络的风电机组运行状态监测方法。该方法通过 K-means 聚类算法构建工况识别模型,实现了风电机组不同运行工况的有效划分。针对 RNN 训练中梯度消失和梯度爆炸的问题,引入 ESN 进行时间序列预测,构建了基于 DE-ESN 的多工况状态监测模型,并结合动态阈值的设置,提高了模型在复杂工况下的适应能力。

实际应用结果表明,该方法能够准确识别机组运行工况并对其运行状态进行有效的监测,当机组故障发生时相比传统预警方法可更早发现异常。

参考文献

- [1] 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 X H, SUN Y, SHAN J H, et al. Fault diagnosis and prognosis for wind turbines: An overview[J]. Chinese Journal of Scientific Instrument, 2017, 38(5): 1041-1053.
- [2] 刘杰, 曹静, 赵昕. 基于 OOB-GWO-SVR 的风电机组齿轮箱故障预警[J]. 电子测量与仪器学报, 2022,

后,基于已训练的 DE-ESN 模型对测试集数据进行功率预测,并结合残差分析机组的运行状态,案例 2 的风电机组运行状态监测结果,如图 14 所示。

从图 14 中可以看出,风电机组在状态监测过程中出现了少量超阈值情况。通过统计分析发现,这些异常点并未连续 3 次超过阈值,因此未触发异常预警。进一步分析原始 SCADA 数据后发现,大多数异常源于风速或风向的急剧变化,导致机组功率预测出现波动,与案例 1 中的情况类似。案例 2 的分析结果进一步验证了所提方法的有效性。

36 (12): 97-105.

LIU J, CAO J, ZHAO X. Wind turbine gearbox fault warning based on OOB-GWO-SVR [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36 (12): 97-105.

- [3] JIN X H, LYU SH Y, KONG Z Q, et al. Graph spatio-temporal networks for condition monitoring of wind turbine[J]. IEEE Transactions on Sustainable Energy, 2024, 15(4): 2276-2286.
- [4] HAN H Y, YANG D SH. Correlation analysis based relevant variable selection for wind turbine condition monitoring and fault diagnosis [J]. Sustainable Energy Technologies and Assessments, 2023, 60: 103439.
- [5] 王中行, 周元贵, 张学广. 基于人工智能算法的风电机组状态监测和故障诊断技术研究综述[J]. 东北电力大学学报, 2024, 44(1): 42-51.
WANG ZH X, ZHOU Y G, ZHANG X G. Review of artificial intelligence algorithms-based wind turbine condition monitoring and fault diagnosis techniques[J]. Journal of Northeast Electric Power University, 2024, 44(1): 42-51.
- [6] WANG Y, ZOU R M, LIU F, et al. A review of wind speed and wind power forecasting with deep neural

- networks [J]. *Applied Energy*, 2021, 304 (15): 117766.
- [7] WU Y Q, MA X D. A hybrid LSTM-KLD approach to condition monitoring of operational wind turbines [J]. *Renewable Energy*, 2022, 181: 554-566.
- [8] 马良玉, 梁书源, 程东炎, 等. 基于 QM-DBSCAN 与 BiLSTM 的风电机组异常工况预警研究[J]. *计量学报*, 2024, 45(9): 1384-1393.
- MA L Y, LIANG SH Y, CHENG D Y, et al. Research on early warning of abnormal working conditions of wind turbine based on QM-DBSCAN and BiLSTM [J]. *Acta Metrologica Sinica*, 2024, 45(9): 1384-1393.
- [9] 陈进, 靳佳澍, 陈跃鹏, 等. 基于贝叶斯优化 ESN 的 PEMFC 性能退化预测[J/OL]. *中国电机工程学报*, 1-12[2024-10-12].
- CHEN J, JIN J SH, CHEN Y P, et al. PEMFC performance degradation prediction based on Bayesian optimized ESN model[J/OL]. *Proceedings of the CSEE*, 1-12[2024-10-12].
- [10] JAEGER H, HAAS H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication[J]. *Science*, 2004, 304 (5667): 78-80.
- [11] GHIMIRE S, NGUYEN-HUY T, AL-MUSAYLH M S, et al. A novel approach based on integration of convolutional neural networks and echo state network for daily electricity demand prediction[J]. *Energy*, 2023, 275(15): 127430.
- [12] HU H L, WANG L, TAO R. Wind speed forecasting based on variational mode decomposition and improved echo state network[J]. *Renewable Energy*, 2021, 164: 729-751.
- [13] JIN X H, XU ZH W, QIAO W. Condition monitoring of wind turbine generators using SCADA data analysis[J]. *IEEE Transactions on Sustainable Energy*, 2021, 12(1): 202-210.
- [14] ZHAN J, WU CH K, YANG C Q, et al. Condition monitoring of wind turbines based on spatial-temporal feature aggregation networks [J]. *Renewable Energy*, 2022, 200: 751-766.
- [15] ZHANG SH Y, HUANG K ZH, LIU CH L, et al. Wind turbine condition monitoring based on double-layer ensemble KNN method[J]. *Energy Science & Engineering*, 2024, 12(1): 136-148.
- [16] 齐咏生, 景彤梅, 高学金, 等. 基于多维度 SCADA 参数的风电机组异常辨识[J]. *控制工程*, 2021, 28(12): 2393-2401.
- QI Y SH, JING T M, GAO X J, et al. Wind turbine anomaly identification based on multi-dimensional SCADA parameters [J]. *Control Engineering of China*, 2021, 28(12): 2393-2401.
- [17] 陈结, 陈换过, 肖志奇, 等. 基于工况辨识的风电机组主传动系统运行状态监测[J]. *太阳能学报*, 2024, 45(2): 77-85.
- CHEN J, CHEN H G, XIAO ZH Q, et al. Operational state monitoring of wind turbine main transmission system based on working condition recognition [J]. *Acta Energetica Sinica*, 2024, 45(2): 77-85.
- [18] 金晓航, 王宇, ZHANG Bin. 工业大数据驱动的故障预测与健康监测[J]. *计算机集成制造系统*, 2022, 28(5): 1314-1336.
- JIN X H, WANG Y, ZHANG B. Industrial big data-driven fault prognostics and health management [J]. *Computer Integrated Manufacturing Systems*, 2022, 28(5): 1314-1336.
- [19] 董泽, 贾昊. 基于 EWT-LOF 的热工过程数据异常值检测方法[J]. *仪器仪表学报*, 2020, 41(2): 126-134.
- DONG Z, JIA H. Outlier detection method for thermal process data based on EWT-LOF[J]. *Chinese Journal of Scientific Instrument*, 2020, 41(2): 126-134.
- [20] WEI N, YIN CH, YIN L H, et al. Short-term load forecasting based on WM algorithm and transfer learning model[J]. *Applied Energy*, 2024, 353: 122087.
- [21] IKOTUN A M, EZUGWU A E, ABUALIGAH L, et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data[J]. *Information Sciences*, 2023, 622: 178-210.
- [22] 于重重, 宁亚倩, 秦勇, 等. 基于 T-SNE 样本熵和 TCN 的滚动轴承状态退化趋势预测[J]. *仪器仪表学报*, 2019, 40(8): 39-46.
- YU CH CH, NING Y Q, QIN Y, et al. Prediction of rolling bearing state degradation trend based on T-SNE sample entropy and TCN [J]. *Chinese Journal of Scientific Instrument*, 2019, 40(8): 39-46.
- [23] 胡焕玲. 基于改进回声状态网络的能源预测问题研

究[D]. 武汉: 华中科技大学, 2021.

HU H L. Research on energy forecasting problems based on improved echo state network[D]. Wuhan: Huazhong University of Science & Technology, 2021.

[24] LIU ZH H, WANG CH T, WEI H L, et al. A wavelet-LSTM model for short-term wind power forecasting using wind farm SCADA data [J]. Expert Systems with Applications, 2024, 247: 123237.

[25] KONG Z Q, TANG B P, DENG L, et al. Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units[J]. Renewable Energy, 2020, 146: 760-768.

[26] 安睿, 缪书唯. 基于互 Box-Cox 变换和 Markov 链风速云模型的发电系统充裕度评估[J]. 电力自动化设备, 2024, 44(3): 113-119,141.

AN R, MIAO SH W. Adequacy assessment of power generation system based on wind speed cloud model of mutual Box-Cox transform and Markov chain[J]. Electric

Power Automation Equipment, 2024, 44(3): 113-119, 141.

[27] SAARI J, STRÖMBERGSSON D, LUNDBERG J, et al. Detection and identification of windmill bearing faults using a one-class support vector machine (SVM) [J]. Measurement, 2019, 137: 287-301.

作者简介



金晓航(通信作者),2014 年于香港城市大学获得博士学位,现为浙江工业大学教授,主要研究方向为工业大数据分析,复杂装备智能运维。

E-mail:xhjin@zjut.edu.cn

Jin Xiaohang (Corresponding author) received the Ph. D. degree from City University of Hong Kong in 2014. He is a Professor with the College of Mechanical Engineering, Zhejiang University of Technology. His research interests include industrial big data analysis, condition monitoring, fault diagnosis and prognosis, and intelligent maintenance for complex electromechanical equipment.