

基于 CNN-BiLSTM-Attention 的风电机组故障预警<sup>\*</sup>

倪炳阳 何 青

(长沙理工大学电气与信息工程学院 长沙 410114)

**摘 要:** 在研究基于深度学习的风电机组故障预警问题时,针对模型的预测精度与故障预警的准确性问题,提出了一种融合 CNN、BiLSTM、注意力机制 Attention 的组合模型预警方法。首先,针对 SCADA 原始数据质量较低的问题,使用参数优化的 DBSCAN 算法结合风电机组控制原理完成数据清洗,并使用 GRA 分析法筛选原始特征以降低特征间冗余;针对模型预测精度问题,为提高 BiLSTM 网络的特征提取能力以及对于关键特征的聚焦能力,分别引入 CNN 和注意力机制,搭建出组合网络模型;最后使用指数加权方法对功率残差进行平滑处理,从而确定预警阈值,实现风电机组的故障预警。通过某风电场的 SCADA 数据验证了方法的有效性。实验结果表明:本文模型相比于 BiLSTM 模型误差指标 RMSE、MAE 分别降低了 29.8%、30.7%,拟合度  $R^2$  提高了 4.8%。预警时间比 SCADA 报警日志提前 2~6 h。

**关键词:** SCADA 系统;双向长短期记忆网络;故障预警;数据清洗

**中图分类号:** TM614;TN03 **文献标识码:** A **国家标准学科分类代码:** 470.4047

## Early fault warning of wind turbines based on CNN-BiLSTM-Attention

Ni Bingyang He Qing

(School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China)

**Abstract:** In the study of wind turbine fault early warning based on deep learning, aiming at the prediction accuracy of the model and the accuracy of fault early warning, a combined model early warning method combining CNN, BiLSTM and attention mechanism Attention is proposed. Firstly, aiming at the problem of low quality of SCADA raw data, the parameter-optimized DBSCAN algorithm is combined with the control principle of wind turbine to complete data cleaning, and the GRA analysis method is used to screen the original features to reduce the redundancy between features. Aiming at the problem of model prediction accuracy, in order to improve the feature extraction ability of BiLSTM network and the focusing ability of key features, CNN and attention mechanism are introduced respectively to build a combined network model. Finally, the exponential weighting method is used to smooth the power residual, so as to determine the early warning threshold and realize the fault early warning of wind turbines. The effectiveness of the method is verified by the SCADA data of a wind farm. The experimental results show that compared with the BiLSTM model, the error indexes RMSE and MAE of the proposed model are reduced by 29.8% and 30.7% respectively, and the fitting degree  $R^2$  is increased by 4.8%. The warning time is 2~6 hours earlier than the SCADA alarm log.

**Keywords:** SCADA system; bidirectional long short-term memory network; fault warning; data cleaning

## 0 引 言

随着风力发电的广泛应用,风电机组的健康监测和故障预警成为确保其高效运行的关键因素。风电机组的故障不仅会影响发电效率,造成设备损坏,还可能会导致长时间停机,甚至引发灾难性故障,造成巨大的经济损失。因此,

利用已有的风电机组运行数据预测潜在故障的发展趋势,在故障发展初期进行预警并制定预防性措施,已成为当下的研究热点<sup>[1]</sup>。

风电机组数据采集与监控 SCADA 系统提供了海量的风机运行数据,这些数据具有非线性、维度多、规模大、复杂度高等特点,为传统的机器学习方法带来了挑战<sup>[2]</sup>。然而,

相比浅层结构的机器学习模型,深度学习具有强大的特征学习能力,能够构建复杂的深层次模型,从而有效地处理这类数据。周伟等<sup>[3]</sup>提出了一种改进的 MFO-LSTM 网络的风电机组齿轮箱故障预警方法。该方法结合了飞蛾火焰优化算法(moth-flame optimization, MFO)与长短期记忆网络(long short-term memory, LSTM),通过优化 LSTM 的超参数来提高预测精度。但该方法在多次迭代优化过程中的计算量仍然庞大。且单一的 LSTM 模型在处理长时间依赖关系和高维特征时存在一定局限性,难以充分挖掘数据中的复杂模式。赵洪山等<sup>[4]</sup>提出了一种将卷积自编码(contractive autoencoder, CAE)与双向长短期记忆网络相结合的风电机组齿轮箱故障预警方法。通过卷积自编码算法对齿轮箱进行建模和仿真,获取其运行状态数据。尽管该方法在特征提取方面引入了 CAE 的优势,但其对于数据中的局部特征的捕捉仍然依赖于手工设计的卷积自编码器结构,这限制了模型的自适应能力。朱彦民等<sup>[5]</sup>提出了一种改进的 Transformer 模型,通过引入线性解码器结构来简化传统 Transformer 模型,提高了模型的稳定性,但未对模型精度进行深入研究。Fu 等<sup>[6]</sup>提出了一种基于深度学习模型 CNN-LSTM 的风电机组齿轮箱轴承状态监测方法,通过卷积神经网络进行特征提取和降维,结合长短时记忆网络对时间序列数据进行预测,有效解决了传统方法在处理大规模风电数据时的精度低和过拟合问题。但该模型在长时间序列数据上缺少对关键信息的关注,没有显示地为不同特征分配权重,从而降低了重要特征的影响。可见,上述模型在风电机组故障预测方面存在着迭代训练时间过长、特征提取能力不足、难以自适应分配特征权重等问题。相较于以上深度学习模型, CNN-BiLSTM-Attention 模型具有多方面的优势。其中,卷积神经网络(convolutional neural networks, CNN)通过卷积操作提取数据中的特征,并通过池化操作减少数据维度,从而减低计算复杂度,提高计算效率。通过(bidirectional long short-term memory, BiLSTM)处理提取的特征,解决单向 LSTM 在处理长时间序列的信息丢失问题,最后引入注意力机制,动态地聚焦于数据中的关键特征,自动调整对不同特征的关注权重,增强模型对重要特征的感知能力。综上, CNN-BiLSTM-Attention 模型通过多种模型融合,充分发挥各自优势,能够充分挖掘长时间数据序列中的重要特征以及各个要素之间的相互关联,从而提高模型对于风电机组故障预警的精度。

本文使用的数据集来源于某风电场的一台 2 050 kW 变速恒频双馈型风力发电机,在 CNN-LSTM-Attention 模型的基础上,采用(density-based spatial clustering of applications with noise, DBSCAN)算法进行数据清洗,以提高建模数据的质量,并引入多种优化算法寻找模型最优参数,构建风电机组正常工况模型,同时对比多种深度学习模型在该台风电机组上的预测精度,并通过指数加权法平

滑处理功率残差,从而提高故障预警信号的可靠性,降低误报率。以此实现风电机组整机故障预警。

## 1 结合风机控制原理的原始数据处理

SCADA 原始数据采样时间为 2019 年 1 月~2021 年 4 月,总计覆盖 28 个月,采样间隔为 10 min,数据数量约 12 万条,该数据集的风速—有功功率分布情况如图 1 所示。

从图 1 中可以看出,因受到外部环境因素与内部设备老化、损伤等因素的影响,或由于计划停机、故障停机、弃风限电等情况,SCADA 原始数据中存在着大量偏离风电机组正常运行特性的异常数据点。这些异常值对于风电机组正常运行特性的建模精度会造成显著影响,因此在模型训练前必须对异常值进行清洗和处理,以确保模型的准确性和可靠性。

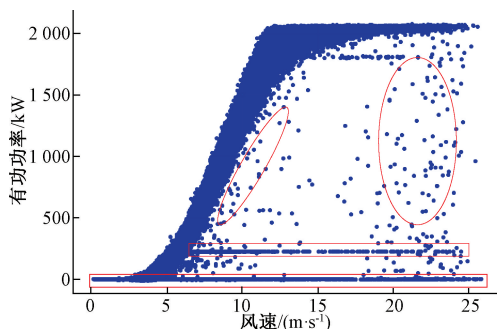


图 1 风速-有功功率分布图

Fig. 1 Wind speed-active power distribution diagram

根据异常数据的特点,本文将其划分为以下 3 类<sup>[7]</sup>:

A 类(停机待机类功率点):这类异常值是由环境停机、故障停机、用户停机、技术待机等原因导致,表现为功率值在零附近的集中的一簇点。

B 类(弃风限电类功率点):这类异常值通常由机组参与并网型能检测或自动发电控制(automatic generation control, AGC)调试引起,包括限功率点及弃风限电功率点。

C 类(异常功率点):此类数据主要是由传感器故障或机组故障导致,呈现为无规律、低密度分散在风功率曲线附近的散点。

### 1.1 异常数据清洗

针对上述提出的 3 种异常数据点,本文采用了一种风电机组控制原理结合贝叶斯优化(bayesian optimization, BO)的基于密度的含噪声空间聚类方法<sup>[8]</sup>的数据清洗方法。

#### 1) 风电机组控制原理完成一次清洗

变速恒频风电机组的控制原理进行清洗操作过程及其依据如下:

操作一:剔除发电机转速  $Gen\ Speed \leq 1\ 050\ r/min$  的数据。在 Z1 阶段风电机组未并网,发电机转速约在 1 050 r/min 以下,此阶段功率点主要是由未达到切入风速

(通常为 3 m/s)或风电机组处于停机、待机状态导致,属于图 1 中的 A 类异常点。

操作二:剔除桨叶角度  $\text{Blade Angle} \geq 5^\circ$  且功率 < 额定值的数据。风电机组的桨叶角度在  $0^\circ$  附近时可以达到最佳风能利用率,只有在实施变桨控制时桨叶角度才会显著增大。弃风限电即通过改变桨叶角度达到限功率的目的,这类数据点的功率往往达不到额定功率值,即图 1 中的 B 类异常点。考虑到叶片风阻差异以及测量误差的,本文将桨叶角度大于  $5^\circ$  且功率被限定低于额定功率的数据视为异常点并予以剔除。

经过以上的数据清洗操作,A 类异常数据会被完全剔除,绝大部分的 B 类异常数据和部分 C 类异常数据会被清除。

2) 使用 BO-DBSCAN 方法进行二次清洗

为清除剩余的异常数据点,采用 DBSCAN 方法进行二次清洗。DBSCAN 是一种基于密度的空间聚类算法<sup>[9]</sup>,其核心思想是通过数据点的密度分布来识别聚类,从而有效处理噪声点。该算法的核心参数为邻域半径  $\text{eps}$  和最小点数  $\text{Minpts}$ 。其中  $\text{eps}$  确定一个点的搜索范围, $\text{Minpts}$  指定一个点的领域内需包含的最少点数,才能够被视为“高密度区域”。由于  $\text{eps}$  和  $\text{Minpts}$  的取值会影响聚类的效率和效果。为解决此问题,引入贝叶斯优化算法 BO 对参数  $\text{eps}$  和  $\text{Minpts}$  进行自寻优。贝叶斯算法可通过少量目标函数评估获得较为理想的全局优化结果,非常适合优化 DBSCAN 的关键参数。经过两步数据处理的效果如图 2 所示。

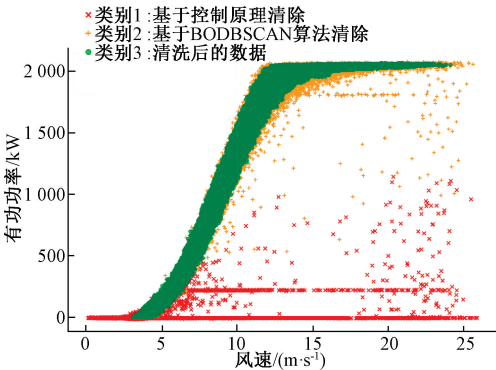


图 2 数据处理效果  
Fig. 2 Data processing effect

由图 2 可知,经过两步清洗操作后,风速-功率图中已经不存在底部堆积数据点和离散数据点,噪声点几乎被全部去除。在剩余数据方面,该方法处理近 12 万条数据用时 40 s 左右,数据的保存量达到 85% 以上,保留了核心特征数据,使数据更加集中分布于功率曲线附近。

1.2 模型输入输出参数的确立

风电机组的有功功率是衡量其运行状态的重要指标之一,能够直观反映机组的工作状况,同时它对多种故障表现出高度敏感性。例如风轮故障通过影响风能捕获效率从而

使有功功率下降;发电机内部绕组过热、机械磨损等会引发功率输出异常;控制系统故障可能导致功率输出不稳定或者将功率限制在特定值。因此选取有功功率作为模型的输出参数。SCADA 系统记录的数据通常包含数十种特征变量。为最终确立模型的输入参数,首先结合风电机组的运行原理,选取风速、桨距角、扭矩设定值、风向以及偏航角度这 5 个对有功功率影响较大的参数作为输入变量。对于其余特征变量,通过采用灰色关联分析 (grey relation analysis, GRA)<sup>[10]</sup> 挑选出与有功功率关联性大于 0.6 的数据,作为补充输入变量。具体计算过程如下:首先,将对二次清洗后的数据进行归一化处理,消除不同量纲的影响;其次,计算风电机组有功功率与各个状态变量的灰色关联系数;然后,使用熵权理论计算各状态变量的权重大小,重新计算各状态变量的关联度得分;最后,对灰色关联度得分进行排序,选择相关度大于 0.6 的状态变量作为下一过程深度学习网络的输入。最终得到的状态变量和补充状态变量按关联度由高低排序,结果分别如表 1 和 2 所示。

表 1 状态变量关联度得分排序  
Table 1 Sorting of correlation scores of state variables

状态变量	关联度得分
风速/( $\text{m} \cdot \text{s}^{-1}$ )	0.798 5
桨距角/( $^\circ$ )	0.790 8
扭矩设定值/( $\text{kN} \cdot \text{m}$ )	0.728 8
风向/( $^\circ$ )	0.640 4
偏航方位角/( $^\circ$ )	0.638 1

表 2 补充变量  
Table 2 Supplementary variables

补充变量	关联度得分
机舱温度/ $^\circ\text{C}$	0.720 3
轮毂温度/ $^\circ\text{C}$	0.667 8
发电机后轴承温度/ $^\circ\text{C}$	0.638 7

2 组合模型的设计与算法流程

2.1 卷积神经网络

CNN 是一种具备出色特征提取能力的前馈神经网络<sup>[11]</sup>,主要由卷积层、池化层和全连接层组成,其经典结构如图 3 所示。

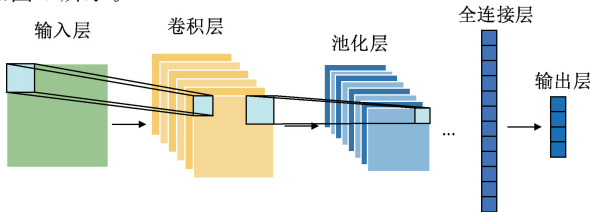


图 3 经典卷积神经网络结构  
Fig. 3 Classical convolutional neural network structure

其中卷积层的作用在于通过设置不同大小的卷积核来提取输入数据的局部特征;池化层则通过既定规则舍弃部分网络参数,既保留原始特征,又降低计算复杂度,从而降低过拟合的风险;全连接层对提取的局部特征进行整合和重组,以形成全局特征。

## 2.2 双向长短期记忆网络

BiLSTM<sup>[12]</sup>是一种改进的循环神经网络(recurrent neural network, RNN)结构,有效缓解了梯度消失和梯度爆炸问题,它通过同时处理序列的正向和反向信息,不仅关注当前数据与历史数据之间的依赖关系,还结合未来数据对当前状态的影响,使得 BiLSTM 能够充分挖掘时间序列中早期和后续信息的关联性,从而显著提高模型对时间序列的整体利用效率和预测精度。图 4 展示的是 BiLSTM 的结构,其中 LSTM 层的细胞单元的结构如图 5 所示。

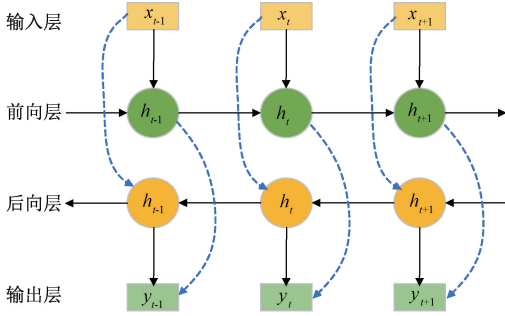


图 4 双向长短期网络结构

Fig. 4 Bidirectional long-term and short-term network structure

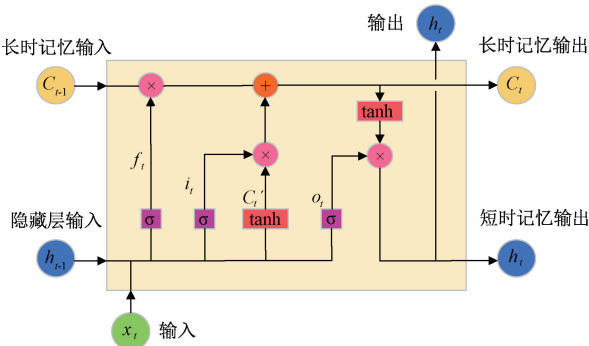


图 5 LSTM 结构单元

Fig. 5 LSTM structure unit

LSTM 细胞单元的数据传递过程如下:

遗忘门的输出  $f_t$  决定 LSTM 中的哪些信息将被删除, 计算公式为:

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

式中:参数中  $\sigma$  为激活函数,  $w_f$  和  $b_f$  分别是遗忘门的权值和偏置量。

输入门中  $i_t$  决定将要被更新到记忆单元  $C_t$  中的新信息。计算公式为:

$$\begin{cases} C_t = f_t \cdot C_{t-1} + i_t \cdot C'_t \\ i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \\ C'_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \end{cases} \quad (2)$$

式中:  $C'_t$  为当前记忆单元候选值,  $\sigma$  为激活函数,  $w_i$ 、 $w_c$  和  $b_i$ 、 $b_c$  分别是输入门的权重和偏置。

输出门  $o_t$  和  $h_t$  的计算公式如下:

$$\begin{cases} o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t \cdot \tanh(C_t) \end{cases} \quad (3)$$

式中:  $\sigma$  为激活函数,  $w_o$  和  $b_o$  分别是遗忘门的权重和偏置。

## 2.3 注意力机制

注意力机制<sup>[13]</sup>源于人脑在特定情境下关注与集中注意力的行为,即在获取特定情况信息时忽略其他不重要信息的特性。在多维数据输入的时间序列处理方面,这个机制使得模型能够有效地学习序列中不同部分的依赖关系和重要性,并根据需要调整其关注的焦点,从而提高模型性能。

在 CNN-BiLSTM-Attention 组合模型中,注意力机制能够有效地捕捉 LSTM 层输出的信息特征,根据重要程度划分权重比例,提取权重信息,从而提高模型的预测精度。具体注意力机制的计算步骤如下:

计算注意力分数,计算公式为:

$$S_i = v^T \cdot \tanh(W_q Q + W_k K_i + b) \quad (4)$$

式中:计算查询向量  $Q$  与键向量  $K_i$  之间的相关性分数  $S_i$ , 式中  $v$  作为可训练的权重向量,用于将  $\tanh$  的输出映射为标量分数,  $W_q$  和  $W_k$  与  $v$  作用相似,分别用于线性变换  $Q$  和  $K_i$ ,使他们映射到相同的维度。

归一化注意力权重,将注意力分数  $S_i$  转化为注意力权重  $\alpha_i$ ,即时间步的重要性权重,计算公式如下:

$$\alpha_i = \text{softmax}(S_i) = \frac{\exp(S_i)}{\sum_{j=1}^T \exp(S_j)} \quad (5)$$

式中:  $\text{softmax}$  是归一化函数用于使所有  $\alpha_i$  的和为 1,  $T$  是输入序列的时间步数。

加权求和,根据注意力权重  $\alpha_i$  对值向量  $V_i$  进行加权求和,得到特征的加权组合  $C$ ,其中值向量  $V_i$  的取值通常与键向量  $K_i$  相同。计算公式如下:

$$C = \sum_{i=1}^T \alpha_i \times V_i \quad (6)$$

## 2.4 PSO 优化 LSTM 层超参数

深度神经网络在处理时间序列预测问题有很大优势,但所需设置的参数数量也更加庞大,且模型参数设置往往根据研究人员的经验决定。但是模型参数的选取会在很大程度上影响模型的性能,因此引入粒子群优化算法 PSO 对 CNN-BiLSTM-Attention 的部分参数进行优化,以期模型能够达到最优性能。

粒子群优化(particle swarm optimization, PSO)是一种基于群体智能的优化算法<sup>[14]</sup>,其算法思想为:



在  $d$  维空间中由  $m$  个粒子构成的一个种群  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ , 这个种群中的任意一个粒子  $\mathbf{x}_i (i \in (1, m))$  在  $t$  时刻的位置为  $\mathbf{X}_i(t) = [\mathbf{x}_{i1}(t), \mathbf{x}_{i2}(t), \dots, \mathbf{x}_{id}(t)]^T$ , 速度为  $\mathbf{V}_i(t) = [\mathbf{v}_{i1}(t), \mathbf{v}_{i2}(t), \dots, \mathbf{v}_{id}(t)]^T$ , 个体最优位置为  $\mathbf{P}_i(t) = [\mathbf{p}_{i1}(t), \mathbf{p}_{i2}(t), \dots, \mathbf{p}_{id}(t)]^T$ , 全局最优位置  $\mathbf{G}_d(t) = [\mathbf{g}_1(t), \mathbf{g}_2(t), \dots, \mathbf{g}_d(t)]^T$ , 粒子  $\mathbf{x}_i$  的形式为  $\mathbf{x}_i = \{\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{id}\}$ , 其速度更新公式和位置更新公式为:

$$\begin{cases} \mathbf{v}_{id}(t+1) = w \cdot \mathbf{v}_{id}(t) + c_1 r_1 (\mathbf{p}_{id}(t) - \mathbf{x}_{id}(t)) + \\ c_2 r_2 (\mathbf{g}_d(t) - \mathbf{x}_{id}(t)) \\ \mathbf{x}_{id}(t+1) = \mathbf{x}_{id}(t) + \mathbf{v}_{id}(t) \end{cases} \quad (7)$$

式中:  $\mathbf{v}_{id}$  表示粒子  $i$  的速度在第  $d$  维的分量;  $\mathbf{x}_{id}$  表示粒子  $i$  的位置在第  $d$  维的分量;  $w$  为惯性权重, 可调节粒子对解空间的搜索范围;  $c_1$  和  $c_2$  代表加速度因子;  $r_1$  和  $r_2$  是两个随机数, 用以增加搜索的随机性, 二者的可取值范围是  $[0, 1]$ , 且二者的取值不会互相影响。

本文采用 PSO 算法对组合模型中的 LSTM 层数, 各层卷积神经元数进行寻优。数据处理以及 PSO 优化组合模型的流程如图 6 所示。

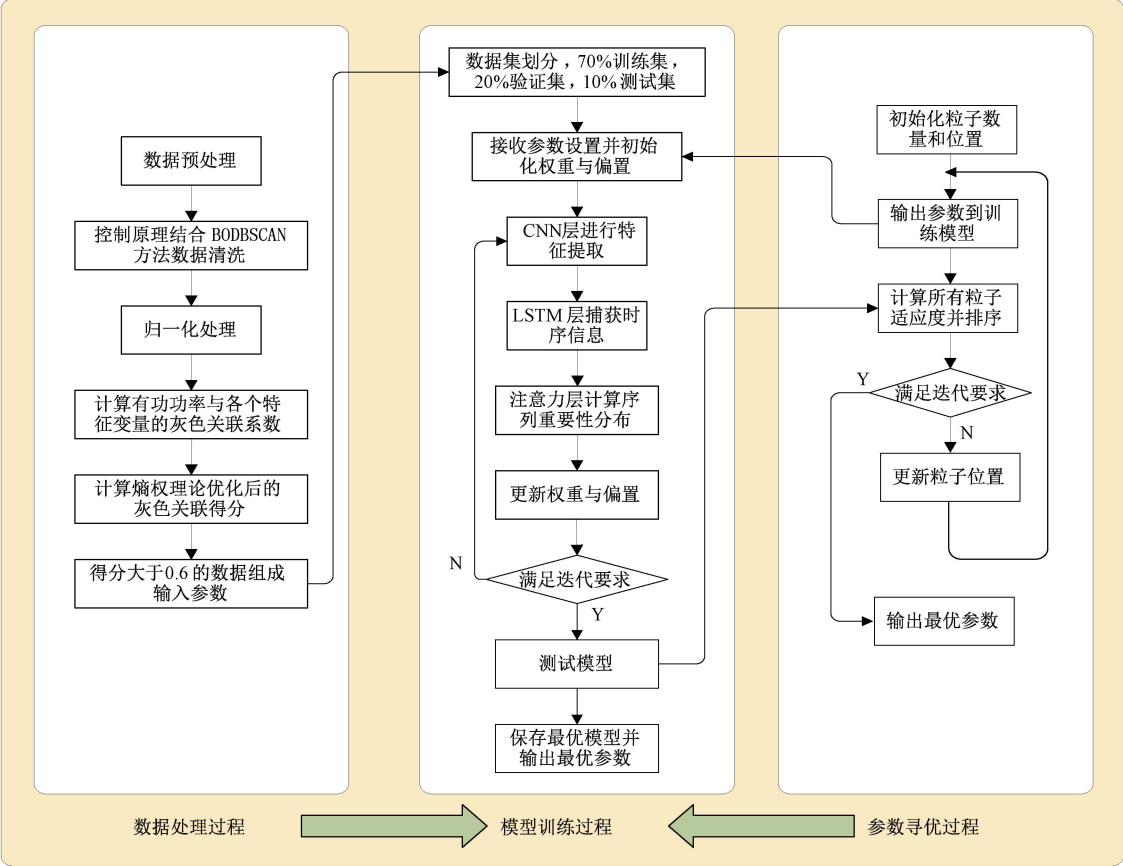


图 6 组合模型的建立流程

Fig. 6 The establishment process of the combined model

3 组合模型的效果验证

3.1 组合模型的参数选择

通过 GRA 关联度分析, 按照从高到低排序得到的 8 个特征变量, 组成 8 维数据作为输入, 有功功率作为输出, 采取 7:2:1 的比例划分训练集、验证集、测试集。

采用 PSO 算法对 LSTM 层的超参数进行优化, 为了确保优化过程的有效性, 事先对 PSO 算法的粒子搜索范围进行了前设置, 其中粒子群的规模为 30, 最大迭代次数为 50, 惯性权重设置为 0.5, 认知因子和社会因子  $c_1$ 、 $c_2$  均设

置为 1.5。LSTM 超参数寻优范围如表 3 所示。

表 3 超参数寻优范围	
Table 3 Super parameter optimization range	
超参数	取值范围(均取整数)
隐藏层神经元数量	[16, 256]
LSTM 层数	[1, 4]
批量大小(batch_size)	[32, 256]

基于此, 构建的组合网络模型参数如下。输入数据为

时间序列,形状为 $[\text{batch\_size}, \text{window}, \text{dim}]$ ,三项参数的对应取值设置为 $[128, 30, 9]$ ,考虑到时序序列之间信息的相关性,加载数据时保留原始顺序。通过维度调整适配 CNN 的一维卷积操作。CNN 模块使用的卷积核大小为 5,激活函数为 Sigmoid,卷积后添加 Dropout 层(丢弃率为 30%);随后引入 LSTM,超参数选取如上;LSTM 层输出的上下文向量经注意力层进行加权,通过展平层转化为一维向量;接着通往一个全连接层,隐藏神经元数为 64,激活函数为 ReLU,经过 Dropout(丢弃率 30%)层以进一步正则化;输出层是一个全连接层,激活函数为 linear,输出单节点,用于输出预测值。

最终,得到完整的风电机组有功功率预测模型架构如图 7 所示。

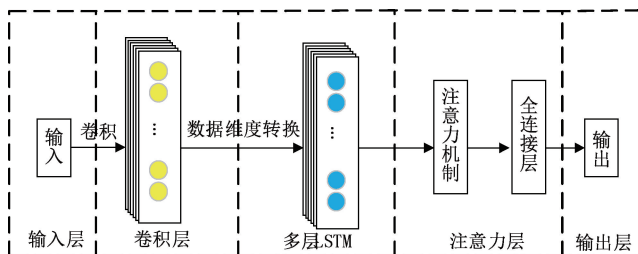


图 7 组合模型结构

Fig. 7 Combination model structure

### 3.2 模型预测结果

在完成原始数据处理和模型结构搭建后,设置好组合模型的所有参数,即可开始进行序列预测任务。将已划分好的 1 万组测试数据输入模型,用于评估模型的性能。图 8 绘制的是有功功率预测值与真实值的对比效果,为确保对比效果清晰明显,图中仅展示了 300 组数据的预测值与实际值的对比曲线,可以看到模型输出的预测值比较准确,曲线拟合效果良好,在功率发生较大波动时反应灵敏。

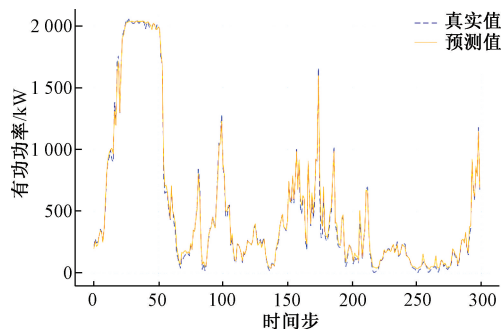


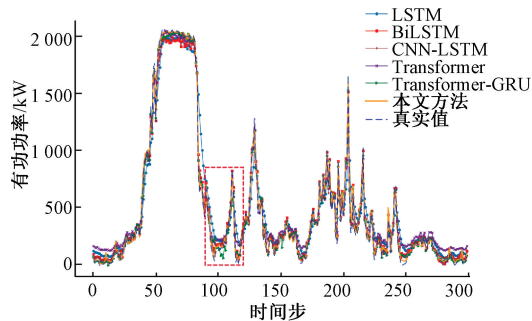
图 8 预测效果图

Fig. 8 Prediction effect diagram

### 3.3 模型对比

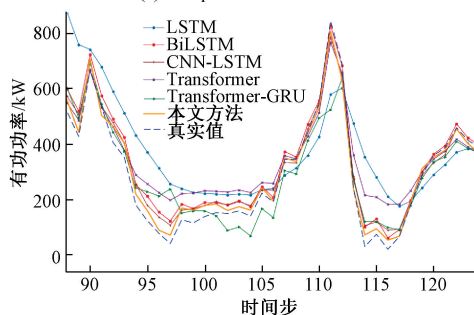
为论证分析本文提出模型的性能,本文同时引进了 5 种深度学习的单模型 LSTM、BiLSTM、Transformer,组合模型 CNN-LSTM、Transformer-GRU<sup>[15]</sup>,和本文提出的模型进行对照测试。为避免人工设置超参数对模型预测性

能的影响,各模型的超参数均使用 PSO 方法进行寻优。6 种模型的功率预测结果和真实值对比情况如图 9 所示。



(a) 整体效果对比

(a) Comparison of overall effects



(b) 红框部分放大效果

(b) Enlarging effect in the red box

图 9 多模型预测效果对比图

Fig. 9 Comparison chart of multi-model prediction effect

由图 9(a)可以看出经 PSO 进行参数优化后的模型均有一定的预测效果,其中 3 种组合模型相较于 3 种单模型效果更佳,特别是对功率波动大的时序序列更为敏感,说明了组合策略可以发挥出各模型的优势,能够挖掘更深层次的信息。观察图 9(b)可以看出本文提出的 CNN-BiLSTM-Attention 模型得到的预测结果最接近实际曲线,拟合度高,特别是在拐点处的跟随能力最强,验证了注意力机制的引入能更加准确的捕捉序列的动态变化,聚焦于重要特征从而提升预测模型的精确度。证明本文所使用的模型在风电机组有功功率预测方面具有良好的适用性和精确度。

为了进一步评估模型的预测性能,基于上述各个深度学习模型,进行 1~3 步的功率预测对比试验,采用 RMSE(均方根误差)、MAE(平均绝对误差)和  $R^2$ (决定系数)3 个指标对测试集的输出的结果进行准确性评估,结果如表 4 所示,各个模型的多步预测指标分布情况如图 10 所示。

结合表 4 和图 10 可知,随着预测步长的增加,需要预测模型提取更深层次的特征,各个模型都表现出一定的预测精度下降,且各模型预测精度和稳定性方面的差距逐步增大。在 1~3 步预测中,单一模型的误差增加幅度相较于组合模型的误差增加幅度更大,说明组合模型的自适应能力和鲁棒性较差。在 3 种组合模型中,

Transformer-GRU 模型相较于 CNN-LSTM 的误差指标更低,而 CNN-LSTM-Attention 模型在测试集中的精确度评价指标 RMSE 分别为 32.264 5、34.213 5、39.643 1,是所有模型中最低的,再次证明了该模型的关键特征捕捉

能力和长时间序列的处理性能的优势。证明本文提出的组合模型产生的预测结果更加可靠,在时序序列的特征提取方面、应对数据剧烈波动方面的能力更强,模型预测效果良好。

表 4 多步预测误差评价指标对比  
Table 4 Comparison of multi-step prediction error evaluation indexes

预测模型	RMSE			MAE			R <sup>2</sup>		
	1 步	2 步	3 步	1 步	2 步	3 步	1 步	2 步	3 步
LSTM	53.578 3	57.156 7	64.758 0	43.238 9	45.645 9	53.128 1	0.872 3	0.863 8	0.858 3
BiLSTM	45.982 3	49.304 8	56.384 9	36.759 3	38.055 7	43.454 6	0.934 9	0.925 1	0.917 8
CNN-LSTM	44.225 6	48.691 2	53.602 1	30.723 4	33.224 0	37.446 2	0.953 2	0.934 8	0.927 2
Transformer	50.472 1	54.528 3	60.437 8	40.375 1	43.376 2	48.578 2	0.907 3	0.890 4	0.886 5
Transformer-GRU	39.839 5	43.947 8	48.290 3	32.482 6	35.224 0	40.517 3	0.953 5	0.932 8	0.923 3
本文模型	32.264 5	34.213 5	39.643 1	25.468 1	28.025 1	32.446 2	0.982 8	0.966 3	0.956 1

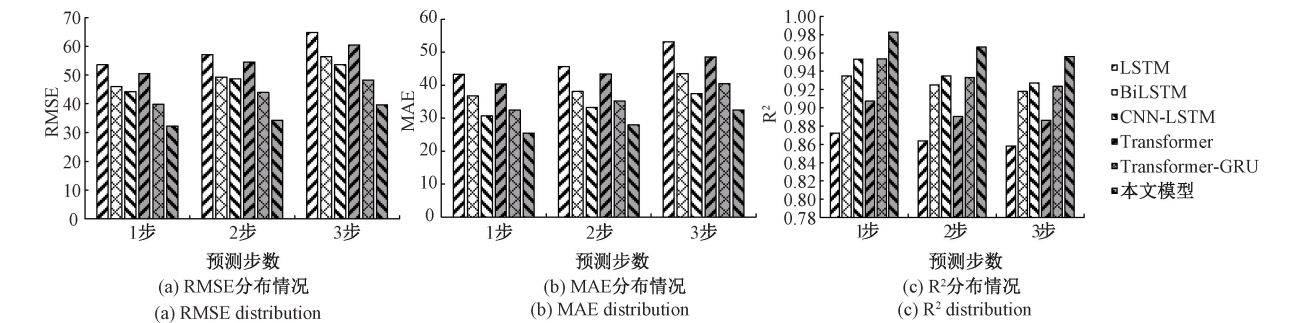


图 10 各模型多步预测效果  
Fig. 10 The multi-step prediction effect of each model

4 故障报警的实现与案例分析

4.1 故障阈值的设置

在完成训练集的数据预测后,通过预测值和真实值的计算就可以得到数据的残差,对残差进行分析可以得出故障预警的阈值。其中最常用的方法是统计历史残差,将误差的分布作为正态分布进行计算从而设定阈值。此方法虽能发现故障,但由于风电机组所处环境恶劣,风速、风向外部因素随机性较大,并且模型也有出现误预测的可能,多种因素导致的传统阈值设置方法可能会导致误报率较高的问题。此外,考虑到输入模型的数据具有时序性,通过预测值和实际值计算所得到的残差也并不是完全相互独立的序列,受到噪声影响。针对以上问题,本文采用一种指数加权移动平均值法(exponentially weighted moving-average, EWMA)对残差进行处理<sup>[16]</sup>,并根据平滑后的残差值设定阈值。图 11 和 12 是测试集数据的残差分布情况。

1) 指数加权法对残差进行平滑处理

指数加权移动平均值原理可以监视整个输出过程的历史数据,赋予最近的观测值更高的权重,远离当前时刻

的历史数据权重递减,因此能减弱噪声的影响。其基本计算公式为:

$$S_t = \alpha \cdot e_t + (1 - \alpha) \cdot S_{t-1}$$
 (8)

式中:  $e_t$  为  $t$  时刻的残差,  $S_t$  为当前时刻的统计指标,  $S_{t-1}$  为上一时刻的统计指标,是由历史残差得到的指数加权平均值,  $\alpha$  是平滑系数( $0 < \alpha \leq 1$ ),  $\alpha$  越大,平滑效果越弱,更接近原始数据;  $\alpha$  越小,平滑效果更强,对短期波动的影响较慢。经过参数调试后,曲线平滑的效果如图 13(a)所示。

2) 确定故障预警阈值

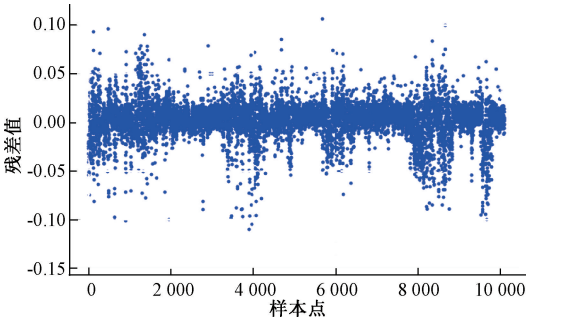


图 11 测试集残差分布散点图  
Fig. 11 Test set residual distribution scatter plot

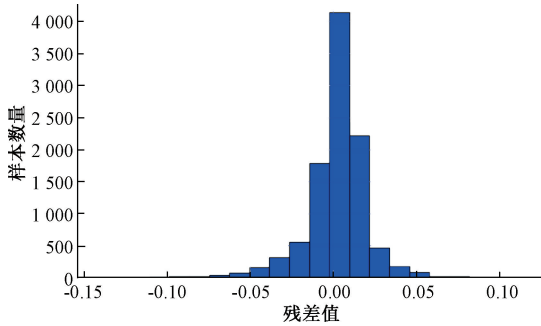


图 12 测试集残差分布直方图

Fig. 12 Residual distribution histogram of test set

机组的整体工作性能可能由于外部环境的改变或工作年限的增加而发生变化,所以定期更新预警阈值有助于提升故障报警的准确度<sup>[17]</sup>。残差曲线经过平滑处理后,已经得到平滑后的残差指标  $S_t$ ,基于此,结合已有的残差数据  $e_t$ ,设计动态报警阈值  $R$  的计算公式为:

$$R = \mu_{S_t} + c \cdot \sigma_{S_t} \quad (9)$$

式中:  $R$  为当前窗口的预警阈值;  $c$  为构造触发预警阈值控制极限的宽度系数,一般取值为 3;  $\sigma_{S_t}$  为  $S_t$  的均方根,可由计算得到:

$$\sigma_{S_t}^2 = \sigma_{e_t}^2 \cdot \left( \frac{\alpha}{2-\alpha} \right) \cdot [1 - (1-\alpha)^{2n}] \quad (10)$$

式中:  $n$  为采样长度,  $R$  的取值会随着选取数据的长度不同而发生改变,并且随着每次采样的样本点所对应的  $e_t$  和  $S_t$  的变化而变化,可以根据新的正常工况不断进行历史数据的更新。

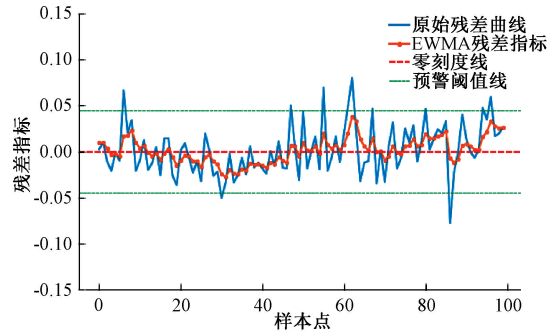
为了展示动态预警阈值的有效性,图 13(a)展示的是参数取值为  $c=3, a=0.6, n=100$  的预警效果,可看到通常不会出现残差指标值超过预警阈值的情况。图 13(b)是通过正态分布公式计算得阈值为 0.054,可明显看出会有个别数据点超过阈值曲线,这些采样点可能是由于测量误差、噪声信号或者模型预测不准确等多种因素造成,可能会引发错误报警。

#### 4.2 案例分析

SCADA 系统除了记录历史数据,其报警日志还记录着故障的发生时刻和引发故障的原因。将本风电机组 2021 年 8 月 ~ 2021 年 12 月的数据输入模型。根据 SCADA 报警日志记录的两个故障案例分析来验证本文所使用方法的有效性。

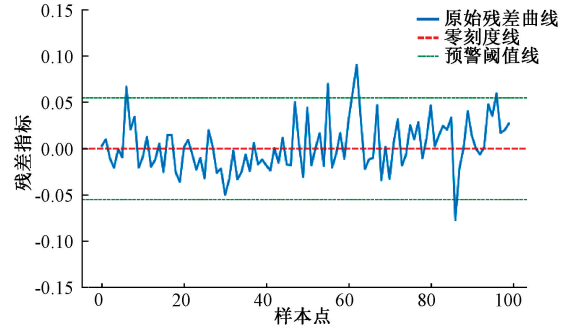
##### 1) 案例一分析

SCADA 报警日志显示 2021 年 9 月 9 日 3 点 50 分发生故障“主控与变桨系统通信失败”,选取该故障发生前的 200 个采样点的数据进行分析,得到的残差指标曲线和报警阈值线如图 14(a)所示。从图 14 中可以看出在第 162 个采样点处指标超出阈值,预警时间比 SCADA 报警日志第 200 采样点提前 6 时 20 分,证明该平滑后的曲线故障识别能力仍然敏感,具有足够的预警能力。



(a) 本文方法计算阈值效果

(a) The method used in this paper calculates the threshold effect



(b) 概率分布方法计算阈值效果

(b) The probability distribution method calculates the threshold effect

图 13 平滑曲线阈值与概率分布阈值对比

Fig. 13 Comparison of smooth curve threshold and probability distribution threshold

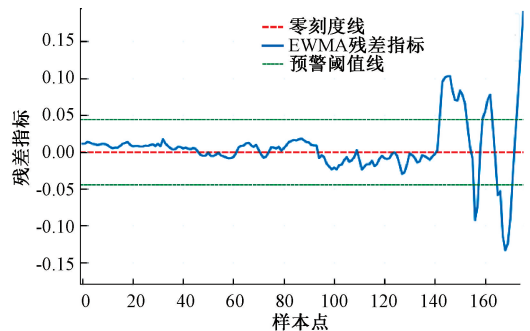
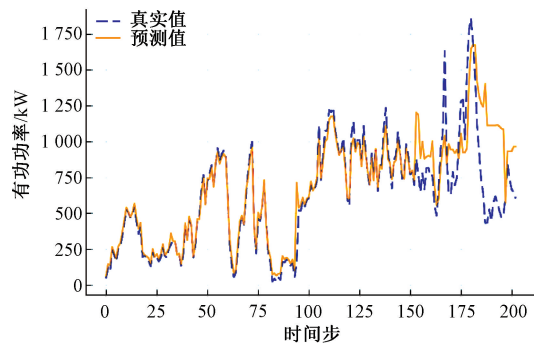
(a) 残差指标  
(a) Residual indicator(b) 功率曲线对比  
(b) Power curve comparison

图 14 案例一残差指标与功率对比图

Fig. 14 Case 1 residual index and power comparison diagram



为了更清晰地观测风机的变化,绘制预测功率和实测值之间的变化趋势如图 14(b)所示。

图 14 中第 162 采样点附近预测功率和实际功率出现明显差异,查找 SCADA 原始数据,发现桨距角变化出现明显波峰,结合报警日志可以分析出由于主控未收到变桨系统的反馈,导致记录的桨距角与实际不符,因此导致模型预测功率与实际功率不匹配。

## 2) 案例二分析

SCADA 报警日志于 2021 年 11 月 16 日 3 点 49 分记录的故障为“PHC 震动导致安全链触发”。选取该故障发生前 200 个采样点的数据进行分析,得出的指标曲线如图 15(a)所示。从图 15(a)中可以看出,该机组在第 187 采样点开始,连续 3 个采样点数据超出阈值,相比报警系统记录的第 200 采样点提前约 2 h。后面指标在预警线附近变动,反复发出预警信号。

绘制预测功率和实测功率之间的变化趋势如图 15(b)所示,可发现预测值和实际值不匹配,同时机组功率曲线表现为能够达到满发但难以维持,存在严重的波动现象。可推断出此时发电机或传动系统内部可能出现了部件老化或润滑不足的情况,亦可能是出现了机械不平衡的情况。

性。然而,SCADA 系统记录的多维特征数据中包含了大量风电机组运行工况的关键信息。为将这些数据用于风电机组的故障预警,本文提出了一种基于注意力机制的组网模型,并对传统的故障识别方法进行改进,形成了更加高效的风电机组故障预警方法。

首先,结合风电机组的控制原理,采用 DBSCAN 算法对 SCADA 原始数据进行清洗。通过有效去除噪声点、限功率点等异常工况数据,提取了更符合风功率曲线的核心特征数据。随后,使用粒子群优化算法对 LSTM 层的超参数进行调优,构建了基于 CNN-BiLSTM-Attention 的注意力机制组网模型,并与其他深度学习模型进行了对比实验。在模型精度和拟合度方面,所提出的模型均表现出显著优势。在多步时间步长预测对比中,模型所表现出的稳定性更优。通过指数加权平均值法对功率残差进行平滑处理,设置平滑后的报警阈值,从而有效降低了故障误报的概率。最终,通过真实案例分析验证,该基于注意力机制的混合神经网络模型能够准确监测风电机组的运行状态,从而实现对风电机组的故障预警。

## 参考文献

- [1] 龙寰,杨婷,徐劭辉,等.基于数据驱动的风电机组状态监测与故障诊断技术综述[J].电力系统自动化,2023,47(23):55-69.  
LONG H, YANG T, XU SH H, et al. Overview of wind-turbine condition monitoring and fault diagnosis technology based on data drive[J]. Automation of Electric Power Systems, 2023, 47(23): 55-69.
- [2] 魏书荣,张鑫,符杨,等.基于 GRA-LSTM-Stacking 模型的海上双馈风力发电机早期故障预警与诊断[J].中国电机工程学报,2021,41(7):2373-2383.  
WEI SH R, ZHANG X, FU Y, et al. Early fault warning and diagnosis of offshore doubly-fed wind turbines based on GRA-LSTM-Stacking model[J]. Proceedings of the CSEE, 2021, 41(7): 2373-2383.
- [3] 周伟,魏鑫,李西兴.改进 MFO-LSTM 网络的风电机组齿轮箱故障预警研究[J].机床与液压,2024,52(4):185-194.  
ZHOU W, WEI X, LI X X. Research on gearbox fault warning of wind turbine based on improved MFO-LSTM Network [J]. Machine Tools & Hydraulics, 2024, 52(4): 185-194.
- [4] 赵洪山,刘辉海,刘宏杨,等.基于堆叠自编码网络的风电机组发电机状态监测与故障诊断[J].电力系统自动化,2018,42(11):102-108.  
ZHAO H SH, LIU H H, LIU H Y, et al. Condition monitoring and fault diagnosis of wind turbine generator based on stacked self-coding network[J]. Automation of Electric Power Systems, 2018, 42

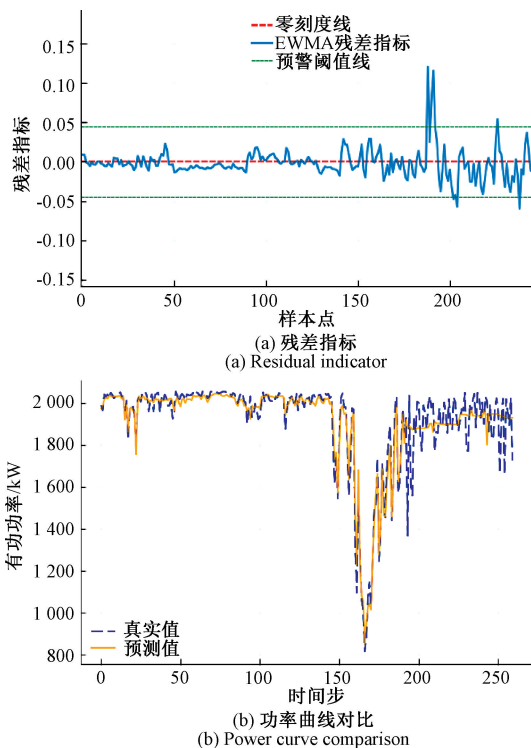


图 15 案例二残差指标与功率对比图

Fig. 15 Case 2 residual index and power comparison diagram

## 5 结 论

SCADA 系统的故障报警信号通常存在一定的滞后

- (11): 102-108.
- [5] 朱彦民,李忠虎,王金明,等. 基于 Transformer 的风电机组故障预测[J]. 电子测量技术, 2024, 47(13): 45-52.
- ZHU Y M, LI ZH H, WANG J M, et al. Fault prediction of wind turbine based on Transformer[J]. Electronic Measurement Technology, 2024, 47(13): 45-52.
- [6] FU J, CHU J CH, GUO P, et al. Condition monitoring of wind turbine gearbox bearing based on deep learning model[J]. IEEE Access, 2019, 7: 57078-57087.
- [7] 刘晨,刘建虎. 基于风电机组控制原理与 QM-BODSCAN 的风功率异常数据清洗方法[J]. 风能, 2023(10):78-86.
- LIU CH, LIU J H. Wind power abnormal data cleaning method based on wind turbine control principle and QM-BODSCAN[J]. Wind Energy, 2023(10): 78-86.
- [8] 马文涛. 基于区间划分的风电机组偏航对风误差在线预警方法[J]. 青海电力, 2024, 43(S1):81-85.
- MA W T. Online warning method of wind turbine yaw to wind error based on interval division[J]. Qinghai Electric Power, 2024, 43(S1): 81-85.
- [9] 付雪娇. 风电机组状态参数预测与运行状态评估方法研究[D]. 沈阳:沈阳工业大学, 2022.
- FU X J. Research on state parameter prediction and operation state evaluation method of wind turbine[D]. Shenyang: Shenyang University of Technology, 2022.
- [10] 陈涛,孙伟,张旭. 基于灰色关联度的风电齿轮箱传动系统故障树分析[J]. 太阳能学报, 2012, 33(10): 1655-1660.
- CHEN T, SUN W, ZHANG X. Fault tree analysis of wind power gearbox transmission system based on grey correlation degree[J]. Acta Solar Energy Sinica, 2012, 33(10): 1655-1660.
- [11] WANG L, ZHANG Z J, LONG H, et al. Wind turbine gearbox failure identification with deep neural networks [J]. IEEE Transactions on Industrial Informatics, 2017, 13(3):1360-1368.
- [12] 董绍江,李洋,梁天,等. 基于 CNN-BiLSTM 的滚动轴承变工况故障诊断方法[J]. 振动测试与诊断, 2022, 42(5):1009-1016.
- DONG SH J, LI Y, LIANG T, et al. Fault diagnosis method of rolling bearing with variable working condition based on CNN-BiLSTM[J]. Vibration Measurement and Diagnosis, 2022, 42(5): 1009-1016.
- [13] WANG AN Q, PEI Y, ZHU Y Y, et al. Wind turbine fault detection and identification through self-attention-based mechanism embedded with a multivariable query pattern[J]. Renewable Energy, 2023, 211: 918-937.
- [14] 孙燕成,陈富安. 基于 PSO 优化 LSTM 神经网络的机械臂逆运动学求解研究[J]. 电子测量技术, 2022, 45(13):40-45.
- SUN Y CH, CHEN F AN. Research on inverse kinematics of manipulator based on PSO optimization LSTM neural network[J]. Electronic Measurement Technology, 2022, 45(13): 40-45.
- [15] 蒲晓云,杨靖,杨兴,等. 基于分解技术的 IZOA-Transformer-BiGRU 短期风电功率预测[J]. 电子测量技术, 2025, 48(2):39-48.
- PU X Y, YANG J, YANG X, et al. Transform-BiGRU short-term wind power prediction based on parameter optimization and decomposition technology[J]. Electronic Measurement Technology, 2025, 48(2): 39-48.
- [16] 丁宸宇,岳瑞华,李远冬. 修正 EWMA 控制图在 MAP 中的应用[J]. 火力与指挥控制, 2019, 44(9):78-82.
- DING CH Y, YUE R H, LI Y D. Application of modified EWMA control chart in MAP[J]. Fire Control & Command Control, 2019, 44(9): 78-82.
- [17] WU H, DAI Y, WANG CH, et al. Identification and forewarning of GNSS deformation information based on a modified EWMA control chart[J]. Measurement, 2020, 160: 107854.

## 作者简介

倪炳阳,硕士研究生,主要研究方向为风电机组故障预警、控制算法优化与 PLC 编程。

E-mail:1191799319@qq.com

何青(通信作者),博士,副教授,主要研究方向为智能控制、现代控制理论以及嵌入式系统。

E-mail:254739027@qq.com