

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

魏书荣, 周海林, 符 杨, 黄玲玲, 葛晓琳

(海上风电技术教育部工程研究中心(上海电力大学), 上海 200090)

摘要: 自海上风电进入平价时代, 迫切需要更加精准的故障预警提高风力发电机的可靠运行水平, 减少发电损失, 但仅依靠物理或数据模型进行风力发电机早期故障预警受限于模型准确性的问题影响预警精度。为此, 提出一种模型-数据融合的建模方法, 基于等效热网络模型和 Stacking 集成算法融合驱动实现海上双馈风力发电机早期故障预警。首先, 利用等效热网络法构建风力发电机温度的热平衡矩阵, 求解得到各节点稳态温度值, 采用一阶 RC 热网络模型描述温度随时间变化的趋势; 然后, 将热模型计算得到的定子绕组温度和其他相关变量作为 Stacking 集成算法的输入特征, 对定子绕组温度值进行校正; 最后, 利用 K-S(Kolmogorov-Smirnov)检验原理确定自适应阈值, 根据残差的变化趋势进行早期故障预警。以国内某海上风电场 SCADA 数据为例进行分析, 验证融合模型的有效性。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方法具有通用性, 为海上风电高质量发展提供技术支撑。

关键词: 海上风电; 故障预警; 模型-数据融合; Stacking 集成算法

Early Warning of Offshore Wind Turbine Failures Based on Temperature-thermal Modeling and Data Fusion Drive

WEI Shurong, ZHOU Hailin, FU Yang, HUANG Lingling, GE Xiaolin

(Engineering Research Center of Offshore Wind Power Technology, Ministry of Education (Shanghai University of Electric Power), Shanghai 200090, China)

Abstract: After offshore wind power enters the parity era, there is a more urgent need for accurate fault prediction to improve the reliable operation of wind turbines and reduce power generation losses; however, relying only on physical or data models for wind turbine early fault warning often faces the problem of model accuracy. Based on the idea of model-data fusion modeling, a fusion-driven offshore doubly-fed wind turbine early-fault-warning method based on equivalent thermal network model and Stacking integration algorithm is proposed. Firstly, the equivalent thermal network method is used to construct the wind turbine thermal balance matrix, the matrix is solved to obtain the steady state temperature values of each node, and the first-order RC thermal network model is adopted to describe the temperature trend over time. Then, the stator winding temperature and other related variables calculated by the thermal model are used as input features of the Stacking integration algorithm to correct the stator winding temperature values. Finally, the K-S (Kolmogorov-Smirnov) test principle is utilized to determine the adaptive threshold value, and early fault warning is carried out according to the trend of the residuals. The SCADA data of a domestic offshore wind farm are analyzed as an example to verify the effectiveness of the fusion model. The early warning method for offshore wind turbine faults based on the temperature-heat model driven with data fusion is generalizable and provides a technical support for the healthy and sustainable development of offshore wind power in the era of parity.

Key words: offshore wind power; fault warning; model-data fusion; Stacking integration algorithm

基金资助项目: 国家自然科学基金(52377063); 上海市教委自然科学基金重大项目(2021-01-07-00-07-E00122); 上海科技创新行动计划(22dz1206100); 上海高校特聘教授(东方学者)(TP2020066)。

Project supported by National Natural Science Foundation of China (52377063), Major Natural Science Project of Shanghai Municipal Education Commission (2021-01-07-00-07-E00122), Shanghai Science and Technology Innovation Action Plan (22dz1206100), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (TP2020066).

0 引言

海上风电高质量发展是深入贯彻落实“四个革命、一个合作”能源安全新战略、保障沿海地区能源供应安全的重要举措^[1-2]。我国海上可开发风能资源约 30 亿 kW, 截至 2023 年海上风电累计装机容量约 0.37 亿 kW, 发展空间巨大。然而, 受高盐高湿海洋环境影响, 海上风力发电机故障率高, 大约是陆上故障率的 8 倍^[3]。发电机作为风力发电能量转换的关键设备, 一旦损坏, 停机维护时间长, 停运损失巨大, 国内曾有 80 台风机因故障无法及时维护停运 107 天, 损失约 2 亿元。风力发电系统的安全运行和维护离不开对发电机的有效管理^[4-6]。作为海上风电的主流机型, 双馈风力发电机在平价时代面临更高的故障预警要求。及时准确的故障预警不仅能够提高风电系统的可靠性, 避免故障快速恶化及次生故障发生, 还能通过有效应对措施减少维护成本和停机时间, 为海上风电的规模化、可持续发展提供理论基础与技术支撑, 具有重要的理论意义与工程实用价值。

目前海上风力发电机常用的故障预警方法主要分为两类: 基于机理模型的解析方法和基于数据驱动的智能方法。机理模型通常利用物理规律来分析系统正常和故障运行状态。例如, 文献[7]采用有限元和等效热网络法计算不同条件下的轴承温度。然而, 有限元法模型复杂, 计算结果容易受到网络划分、边界条件和求解算法等因素影响, 难以满足实时性要求; 而等效热网络法虽简化了模型, 但计算精度有限。文献[8]提出基于定子电流分析和磁场分析理论结合的方法, 用于检测无刷双馈发电机偏心故障。然而, 由于发电机在实际工程应用中受工艺、材料、结构等非理想因素影响, 使得实际系统具有复杂性和非线性特征, 单纯依靠机理模型难以精准建模。

基于数据模型的智能方法则依赖于历史数据和统计方法进行故障预警。随着人工智能技术的发展, 各类智能算法在风力发电机故障预警领域广泛应用, 由于深度学习模型能够自动提取特征、处理复杂数据和适应工况变化, 正逐步成为主流的研究手段, 如深度置信网络(deep belief network, DBN)^[9-10], 卷积神经网络(convolutional neural network, CNN)^[11-13], 循环神经网络(recurrent neural network, RNN)^[14]等。文献[15]提出了一种基于

CNN-CVAE 的方法, 用于检测风电机组齿轮箱的早期故障, 提前了 68.17 h 实现故障预警。文献[16]提出了一种基于 LSTM-RNN 的风电机组预警算法, 仿真结果表明该算法能在故障发生前 35 s 内发出预警(故障发生时间设定为 150~300 s, 仿真范围为 0~300 s)。这些算法通常直接从数据中提取和学习复杂系统的行为特性, 尽管具备时效性优势, 但其预警结果容易受到故障样本有限性、数据质量波动和算法局限性的影响。若仅依赖数据驱动模型预测故障, 虽然可以一定程度上捕捉系统行为的变化趋势, 但往往缺乏物理意义上的解释, 在故障样本有限时预警准确度也有待提高。

鉴于单独使用机理模型或数据驱动模型在故障预警中的局限性, 模型-数据驱动融合的故障预警方法受到了学者关注。文献[17]通过分析故障机理, 研究了齿轮箱中各关键参数在重要子部件发生故障时的变化规律, 并基于此建立故障与参数的定性映射关系, 用于故障分类。文献[18]提出基于机理的 ELM 薄壁叶片加工变形预测模型, 取得了较好的预测精度, 并对比了不同模型的预测结果。然而, 在实际风电系统中, 各部件联系紧密, 定性映射关系可能忽略一些重要的热效应, 难以构建准确的热传递模型。此外, 机理-数据模型中数据模型若仅采取单一智能算法, 数据学习层次较浅, 在处理复杂、多变环境时表现出一定局限性, 预测精度仍有待提升。

针对上述问题, 本文提出一种基于温度热模型和数据驱动融合的海上风力发电机早期故障预警分析方法。首先, 利用等效热网络法对发电机内部温度场进行建模, 得到各节点的稳态温度值。以定子绕组温度为例, 采用一阶 RC 热网络模型描述其随时间的变化, 得到高熵的因果性数据-机理定子绕组温度。然后, 将其作为关键变量, 结合其他影响定子绕组温度的变量作为集成数据驱动算法的输入特征, 对机理定子绕组温度进行校正, 提高实际定子绕组温度的预测精度。最后, 根据 K-S 检验原理确定自适应故障预警阈值, 并采用海上风电场实际运行数据验证该融合方法的有效性, 为海上风力发电机的早期故障预警提供一种新思路。

1 模型-数据融合架构

我国海上风电已投运 15 年, 机组配备的数据采集与监视控制系统(supervisory control and data acquisition, SCADA)系统中积累了大量历史数据。

通过分析和挖掘这些数据,可以帮助识别故障模式、预测故障发生,并提供实时监测和反馈,提高风力发电机的可靠性、安全性和运行效率。当发电机故障时,能量转换效率下降,未能有效转化为电能的部分能量会以热量的形式释放,导致元件温度升高。同时机组持续运行在短路或过流状态时,会造成绕组发热温度升高。此外,SCADA系统中布置多达25个温度传感器,用于实时监测风力发电机各元件的温度变化。温度作为监测机组早期故障的重要变量,可以用来评判发电机的健康状态。因此,本文选取温度状态量进行深入分析。

结合海上风力发电机的实际运行情况,基于机理引导数据的融合方式进行故障预警,本文所提模型具体融合架构如图1所示。其中,主要影响因素是指热网络模型构建过程中的直接影响变量风速、电流、环境温度等。通过等效热网络法构建温度热模型,求解热平衡矩阵得到具有高熵特征的温度信息,为数据模型提供先验知识,并以此作为数据模型输入特征之一,该过程的数学表示如下:

$$\begin{cases} \mathbf{y}'_t = F(\mathbf{X}_t) \\ \mathbf{y}_t = H(\mathbf{y}'_t, \mathbf{Y}_t) + \delta \end{cases} \quad (1)$$

式中: t 为时间状态标签; F 和 H 分别为机理模型和数据模型的映射函数; \mathbf{y}'_t 为 t 时刻机理模型求解的目标向量; \mathbf{y}_t 为 t 时刻数据驱动模型预测的目标向量; δ 为随机误差向量; \mathbf{X}_t 和 \mathbf{Y}_t 为 SCADA 系统量测数据组成的向量。而单独的数据模型表达如下:

$$\mathbf{y}_t = G(\mathbf{Z}_t) + \mathbf{u} \quad (2)$$

式中: G 为数据模型的映射函数; \mathbf{Z}_t 为 SCADA 系统量测数据组成的向量; \mathbf{u} 为随机误差向量。两者表达式的差异性表明,当机理模型提供高熵的输入特征来补充数据时,数据模型能够更精确地反映待解决的机理问题特性,进而建立更优化的预测模型。

2 风力发电机热网络模型

2.1 发电机热机理模型构建

本文以 3 MW 水冷双馈风力发电机为例,采用等效热网络法对其传热过程建模。发电机结构如图 2 所示,主要由机壳、轴承、冷却器、定子、转子、气隙和绕组等组成。

双馈异步发电机在工作状态下会产生绕组损耗、铁芯损耗、机械损耗和附加损耗,但主要热源来自绕组损耗和铁芯铁耗,机械损耗和附加损耗比重较低,可忽略不计。当电流流经时铜绕组时,其热效应会产生相应的热量,可表示如下:

$$P_{cu} = \sum I_m^2 r_m \quad (3)$$

式中: I_m 为流经 m 相绕组的电流; r_m 为 m 相绕组电阻。

电机定转子齿部与轭部中,由交变磁通引起的损耗为基本铁损耗。单位质量导磁材料的基本铁耗^[19]为:

$$p_{Fe} = p_{10/50} (f / 50)^2 B_m^2 \quad (4)$$

式中: $p_{10/50}$ 为单位质量导磁材料在 1 T、50 Hz 时产生的损耗,不同材料可查表; f 为磁通交变频率; B_m 为磁通密度,计算 B_m 时,铁芯齿部取齿部磁通

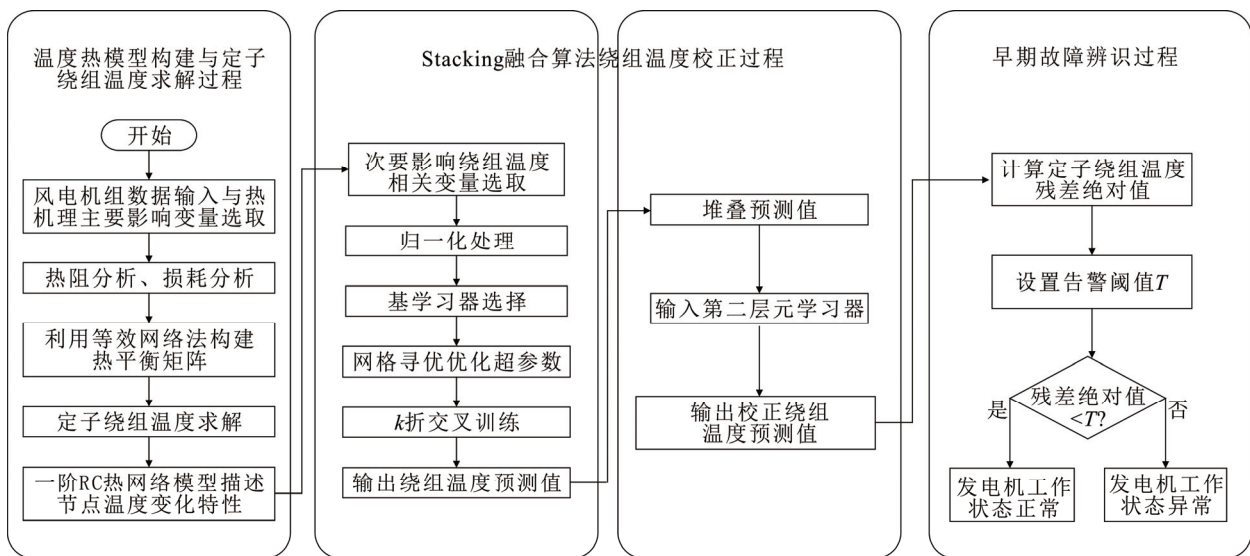


图 1 基于模型-数据融合模型的故障预警流程图

Fig.1 Flowchart of fault warning based on model-data fusion model

密度平均值,铁芯轭部取轭部区域磁通密度最大值。则发电机工作时基本铁损耗可由下式计算:

$$P_{Fe} = K_a p_{Fe} G \quad (5)$$

式中: K_a 为经验系数; G 为导磁材料质量。双馈发电机工作时,由于转差率较小,转子的相对速度很低,转子铁损较小,可忽略不计^[20]。

由于辐射散热对发电机整体热量传递影响较小,因此在本文中忽略了辐射散热。发电机内部热量主要通过热传导和对流换热传递,热传导包括平板热传导和圆管形热传导。平板导热热阻计算如下:

$$R_{plate} = \frac{L}{kA_p} \quad (6)$$

式中: L 为平板厚度; k 为材料导热系数; A_p 为平板截面积。圆管壁的径向导热热阻可表示为:

$$R_{cylinder} = \frac{1}{2\pi kL_p} \ln \frac{r_0}{r_i} \quad (7)$$

式中: L_p 为管道长度; r_0 和 r_i 分别为管道外径和管道半径。

热对流换热时对流换热热阻可计算如下:

$$R_{conv} = \frac{1}{hA_c} \quad (8)$$

式中: h 为对流换热系数; A_c 是传热表面面积。实际计算中,流体特性、流动状态和表面特性等因素会影响对流换热系数,发电机不同位置处于不同流体物态下,因此对流换热系数取值也会不同。

风力发电机内部温度场建模过程中,本文做出以下假设:忽略辐射散热;将绕组材料视为绝缘长方体;假定发电机温度沿圆周方向均匀分布;忽略温度变化对材料物性的影响^[21];忽略绕组表面趋肤效应。此外,由于水冷结构内部复杂,热模型中不考虑水冷散热,在数据驱动部分利用发电机冷却水温度表征。同时,由于端部空间有限且空气导热系数微小,忽略端部空气冷却作用^[22]。

因此,热传递路径主要分为两部分:首先,一部分热量通过发热体内部传导至表面,再由表面传递到外部环境;其次,另一部分热量通过各元件表面与气隙中空气进行对流换热,并散失到外界环境中。根据图 2,将发电机划分为轴、转子轭部、转子齿部、转子绕组、机壳、气隙、定子轭部、定子齿部、定子绕组和定子绕组端部等不同区域。区域中心温度设定为待求温度。利用热电类比的方法,将传热过程中的各部分表示为热阻,温度差类比为电压,从而将双馈异步风力发电机稳态下的复杂温

度场简化为等效热网络模型,如图 3 所示。

本文构建的双馈异步风力发电机等效热网络模型的部分相关参数如附录 A 表 A1 所示。

2.2 热阻分析和计算

2.2.1 气隙

发电机转子旋转时会改变气隙内空气的流动状态,需要考虑转子旋转对气隙空气流动产生的影响。在不同转速下,气隙对流换热系数受转子转速变化影响如附录 B 图 B1 所示。

气隙内的空气流动状态可以采用泰勒数 Ta 进行描述,计算如下^[23]:

$$Ta = \frac{\omega^2 R_i (R_o - R_i)^3}{\nu^2} \cdot \frac{1}{F_G} \quad (9)$$

式中: ν 为空气的运动黏度; F_G 为固定系数,与气隙几何尺寸有关,本文研究的风力发电机气隙径向宽度远小于定子和转子半径,因此可将 F_G 取值为 1; R_i 、 R_o 分别为气隙内、外圆半径; ω 为转子转速。

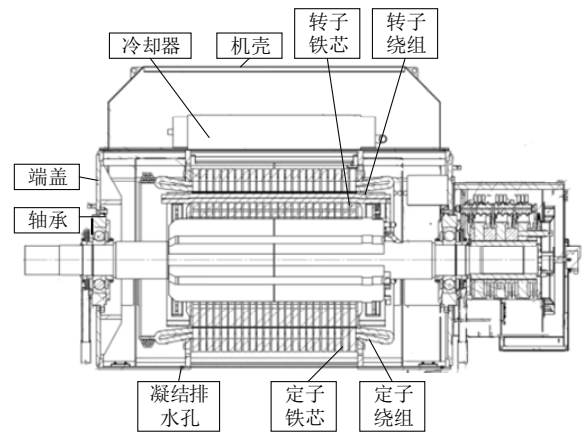


图 2 双馈风机发电机结构图

Fig.2 Structure of doubly-fed wind turbine generator

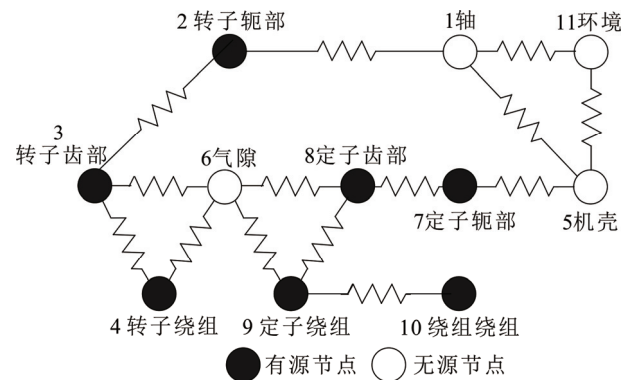


图 3 双馈异步发电机等效热网络简化模型

Fig.3 Simplified model of equivalent thermal network of doubly-fed induction generator

当泰勒数低于临界值 41.19 时, 气隙内气体呈层流状态, 热量通过导热传递, 努赛尔数为 2; 当泰勒数超过临界值 41.19 时, 气隙中形成泰勒涡, 涡流增强对流散热, 此时的努赛尔数为^[24]:

$$Nu = 0.128Ta^{0.734} \quad (10)$$

当泰勒数超过 100 时, 气隙内气体发展为湍流状态, 此时努赛尔数为:

$$Nu = 0.409Ta^{0.482} \quad (11)$$

则不同转速下气隙的对流换热系数 c_{gap} 计算如下:

$$c_{\text{gap}} = \frac{Nu\lambda_{\text{air}}}{D_{\text{gap}}} \quad (12)$$

式中: λ_{air} 为空气的导热系数; D_{gap} 为气隙的水力直径, 计算式如下:

$$D_{\text{gap}} = \frac{4\pi(R_o^2 - R_i^2)}{2\pi(R_o + R_i)} \quad (13)$$

2.2.2 定子

1) 定子绕组主要在气隙、绕组端部、定子齿部和轭部之间进行热量传递。然而, 由于技术和工艺等问题, 定子绕组与定子轭部的接触面积远小于与齿部的接触面积, 因此可忽略两者间的热传递量。定子绕组节点 9 通过绝缘层与定子齿部节点 8 进行热传导, 热阻包括 3 部分, 表示为:

$$R_{8,9} = \left(\frac{t_{\text{eq}}/2}{\lambda_{\text{cu}}} + \frac{t_{\text{st}}/2}{\lambda_{\text{Fe}}} + \frac{d_{\text{ins}}}{\lambda_{\text{ins}}} \right) \frac{1}{h_{\text{eq}}L_{\text{ax}}N_{\text{slots}}} \quad (14)$$

式中: t_{eq} 为定子槽内绕组的等效宽度; t_{st} 为定子齿宽; d_{ins} 为绝缘层厚度; λ_{cu} 为铜绕组导热系数; λ_{Fe} 为铁芯导热系数; λ_{ins} 为绝缘层导热系数; h_{eq} 为定子槽内绕组的等效高度; L_{ax} 为定子轴向长度; N_{slots} 为定子槽数。

定子绕组节点 9 与绕组端部节点 10 在轴向上通过热传导交换热量, 热阻可描述为:

$$R_{9,10} = \frac{L_{\text{ax}}/2 + L_{\text{end}}/2}{\lambda_{\text{cu}}t_{\text{eq}}h_{\text{eq}}N_{\text{slots}}} \quad (15)$$

式中: L_{end} 为绕组端部长度。

定子绕组节点 9 与气隙节点 6 通过对流换热传递热量, 热阻可描述为:

$$R_{9,6} = \left(\frac{h_{\text{eq}}/2}{\lambda_{\text{cu}}} + \frac{d_{\text{ins}}}{\lambda_{\text{ins}}} + \frac{1}{c_{\text{gap}}} \right) \frac{1}{t_{\text{eq}}L_{\text{ax}}N_{\text{slots}}} \quad (16)$$

2) 定子铁芯齿部节点 8 通过热传导与定子绕组和定子轭部交换热量, 通过对流方式与气隙交换热量。定子齿部节点 8 与定子轭部节点 7 之间的热阻可描述为:

$$R_{8,7} = \frac{(r_{\text{so}} - r_{\text{si}})/2}{\lambda_{\text{Fe}}t_{\text{st}}L_{\text{ax}}N_{\text{slots}}} \quad (17)$$

式中: r_{si} 和 r_{so} 分别为定子内圆和外圆半径。

定子铁芯齿部节点 8 与气隙节点 6 之间热阻由铁芯内部和对流换热两部分组成, 可描述为:

$$R_{8,6} = \frac{h_{\text{st}}/2}{\lambda_{\text{Fe}}t_{\text{st}}L_{\text{ax}}N_{\text{slots}}} + \frac{1}{c_{\text{gap}}t_{\text{st}}L_{\text{ax}}N_{\text{slots}}} \quad (18)$$

式中: h_{st} 为定子齿高。

3) 定子轭部通过热传导与定子齿部和机壳交换热量, 则定子轭部节点 7 与机壳节点 5 的热阻可描述为:

$$R_{7,5} = \frac{1}{2\pi\lambda_{\text{Fe}}L_{\text{ax}}} \ln\left(\frac{r_{\text{so}}}{r_{\text{so}} - h_{\text{sy}}/2}\right) + \frac{1}{2\pi\lambda_{\text{c}}L_{\text{ax}}} \ln\left(\frac{r_{\text{c}} - d_{\text{c}}/2}{r_{\text{so}}}\right) \quad (19)$$

式中: h_{sy} 为定子轭部高度; r_{c} 为机壳外圆半径; d_{c} 为机壳厚度; λ_{c} 为机壳导热系数。

2.2.3 转子

转子铁芯轭部与发电机轴部节点通过热传导进行热量传递, 热阻由轭部和轴两部分组成, 可描述为:

$$R_{2,1} = \frac{1}{\lambda_{\text{Fe}}2\pi L_{\text{ax}}} \ln\left(\frac{h_{\text{ry}}/2 + r_{\text{g}}}{r_{\text{g}}}\right) + \frac{1}{\lambda_{\text{g}}2\pi L_{\text{ax}}} \ln\left(\frac{r_{\text{g}}}{r_{\text{g}}/2}\right) \quad (20)$$

式中: h_{ry} 为转子铁芯轭部高度; r_{g} 为发电机轴半径; λ_{g} 为轴导热系数。转子其余部件的传热机理分析和计算过程与定子类似, 这里不再赘述。

2.2.4 其他节点热阻计算

机壳通过热传导和热对流与定子轭部和外界环境之间交换热量。机壳端盖部分通过热传导和热对流与轴和外界环境交换热量。为简化计算流程, 将机壳及端盖视为材料一致的整体, 则机壳与机壳端盖部分和环境节点 11 之间热阻分别如式(21)和式(22)所示。

$$\begin{cases} R_{5,11,1} = \frac{1}{2\pi\lambda_{\text{c}}L_{\text{c}}} \ln\left(\frac{r_{\text{c}}}{r_{\text{c}} - d_{\text{c}}/2}\right) + \frac{1}{c_{\text{air}}A_{\text{t}}} \\ A_{\text{t}} = 2\pi r_{\text{c}}L_{\text{c}} \end{cases} \quad (21)$$

$$\begin{cases} R_{5,11,2} = \left(\frac{d_1}{2\lambda_{\text{c}}} + \frac{1}{c_{\text{air}}}\right) \left(\frac{1}{A_{\text{c}}}\right) \\ A_{\text{c}} = \pi(r_{\text{c}}^2 - r_{\text{g}}^2) \end{cases} \quad (22)$$

式中: L_{c} 为机壳轴向长度; d_1 为端盖厚度。

发电机机壳对流换热系数经验公式见式(23)^[25], 式中 V 为风速。

$$c_{\text{air}} = 9.73 + 14V^{0.62} \quad (23)$$

机壳端盖部分与轴之间的热阻可描述为:

$$R_{1,5} = \frac{1}{\lambda_c 2\pi L_{cg}} \ln\left(\frac{r_1}{r_g}\right) + \frac{1}{\lambda_g 2\pi L_{cg}} \ln\left(\frac{r_g}{r_g/2}\right) \quad (24)$$

式中: L_{cg} 为端盖与轴接触面的轴向长度; r_g 为端盖半径。

2.3 等效热网络模型温度分析

双馈异步风力发电机稳态条件下的等效热网络模型中, 节点内的热量流入等于流出, 即节点内部热量平衡。具体地, 两个节点之间的热量传递 Q_{ij} 可表示为:

$$Q_{ij} = \frac{T_i - T_j}{R_{ij}} \quad (25)$$

式中: T_i 和 T_j 分别为两个节点温度; R_{ij} 为两节点间热阻。因此, 各节点的热平衡方程可写成矩阵形式:

$$\mathbf{GT} = \mathbf{P} \quad (26)$$

式中: \mathbf{G} 是由各热阻倒数构成的矩阵; \mathbf{T} 为各节点温度矩阵; \mathbf{P} 则为热源矩阵, 由电机各部件的损耗计算得出。通过求解该矩阵, 可得到双馈异步风力发电机各节点的稳态温度值。

发电机各元件温度监测中, 定子绕组温度变化是一个重要且敏感的指标。因此, 本文以定子绕组节点温度作为最终预测量。考虑到定子的热容特性, 在 SCADA 数据的记录周期内, 定子绕组温度可能尚未达到稳定状态。因此, 为准确描述定子绕组温度的时变特性, 本文采用一阶 RC 热网络模型来模拟其动态过程, 则定子绕组温度在 t 时刻可表示为:

$$T_2(t) = T_1 + (T_{\text{initial}} - T_1)e^{-\frac{t}{\tau}} \quad (27)$$

$$\tau = R_{\text{sw}} C_{\text{sw}} \quad (28)$$

式中: T_1 为当前周期矩阵定子绕组的稳态温度值; T_{initial} 为当前周期定子绕组的初始温度, 在计算过程中初始化为 $T_{\text{initial}} = T_1$; τ 为时间常数; R_{sw} 和 C_{sw} 分别为定子绕组的热阻和热容。

为了方便观察机理模型的拟合效果, 将冷却水持续作用影响等效为一个定值。需要说明的是, 冷却水的散热效果并非固定不变。本文将冷却水进出口温度作为融合模型的输入变量, 以反映其在不同工况下对发电机温度的影响。结果如图 4 所示, 考虑一阶 RC 热网络影响后, 定子绕组温度变化趋势与实际运行情况更加相符。

3 模型-数据融合应用

3.1 Stacking 集成算法介绍

Stacking 算法是一种集成学习技术^[26], 通过结合多个不同机器学习模型, 平衡各模型的优劣, 提高预测性能。其基本原理是利用多个基模型的预测结果作为输入, 再通过一个次级模型进行综合学习和预测, 从而提升整体的预测精度, 学习框架如图 5 所示。

在第 1 层中, 多个基本模型并行训练, 预测输出通过训练一个元模型来组合。为了避免过拟合, 通常对每个基本模型采用 K 折交叉验证^[27]。具体来说, 训练数据被划分为 K 份, 其中 $K-1$ 份数据用于训练, 剩下的 1 份用于验证模型的预测结果。这个过程重复 K 次, 每次用不同的验证集来测试模型。

在每次交叉验证中, 基本模型会对验证集生成预测值, 并将这些预测值保存下来, 作为第 2 层元模型的训练数据的新特征。同时, 基本模型还会对原始测试集进行预测, 所有的测试集预测结果会堆叠在一起, 作为元模型的测试数据输入。

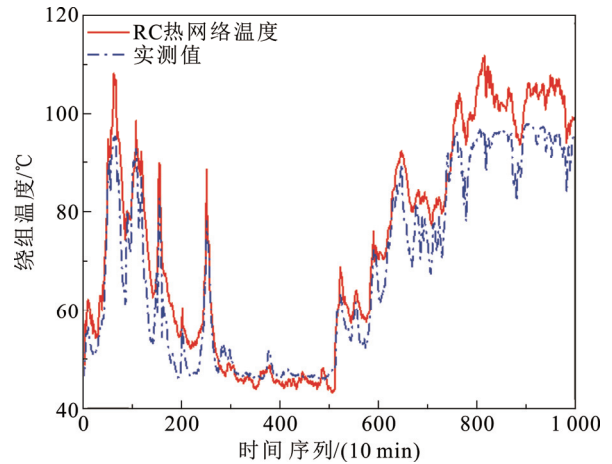


图 4 RC 热网络温度和温度实测值对比图

Fig.4 Comparison of RC thermal network temperature and temperature measurements

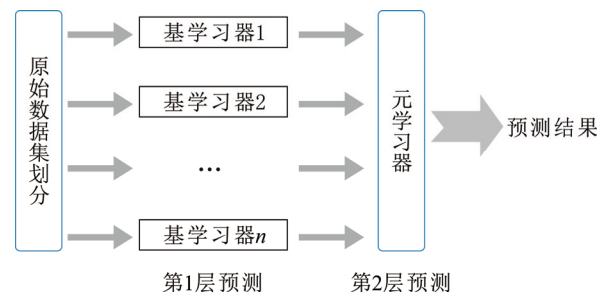


图 5 Stacking 基本框架

Fig.5 The basic framework of Stacking

最后, 使用 K 折交叉验证过程中产生的第 2 层训练数据来训练元模型, 并通过元模型对第 2 层测试数据进行预测, 输出最终的预测结果。图 6 展示了 5 折交叉验证的训练过程。

3.2 学习器选择与相关性分析

选择最佳基学习器时, 为确保融合模型的预测性能, 在满足一定预测效果的前提下, 基学习器还应具备一定差异性。因此, 本文采用皮尔逊(Pearson)相关系数作为评价基学习器差异性的指标。二维向量的 Pearson 相关系数计算式如下:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (29)$$

式中: \bar{X} 、 \bar{Y} 表示两者的平均值; $r=1$ 表明完全正相关; $r=-1$ 表明完全负相关; $r=0$ 表明线性无关。

在选择第 2 层元学习器时, 需要综合考虑多个基学习器的预测结果, 还需考虑算法的稳定性和泛化能力。由于岭(Ridge)回归具有良好的泛化能力和较低的计算复杂度, 可以有效地集成多个基学习器的预测结果。因此本文选择 Ridge 回归模型作为集成模型的第 2 层学习器, 可以在保持模型性能和稳定性的同时, 有效管理模型复杂度和提升泛化能力。

3.3 预测模型评价指标

为检验模型性能, 本文使用模型的决定系数(coefficient of determination, R^2)、均方根误差(root mean square error, RMSE)、平均绝对误差(mean absolute error, MAE)3 个评价指标, 指标计算公式见附录 C。

3.4 告警阈值设定

海上风电场静态告警阈值设定中, 为平滑短期

波动, 一般采用移动平均法来计算每个窗口内的预测误差, 设定固定阈值, 实现异常状态预警。具体计算过程如附录 D 所示。

海上风电场运行环境通常复杂且严酷, 如高盐雾、高湿度和强风等。在这种环境下, 风力发电机的 SCADA 参数可能出现偏差, 即使这些偏差在可接受范围内, 但在残差中仍可能出现极值点。若故障报警系统使用固定阈值, 一旦残差极值超出设定值, 易引发误报。为此, 本文基于自适应原理设计了一种动态阈值设置方法。具体步骤如下:

步骤 1: 确定滑动窗口。K-S 检验认为, 如果两个数据集的 p 值大于 0.05, 则认为它们具有相同的分布特性^[28]。基于此, 本文选取能够反映原始数据集特征的最小数据子集的长度作为滑动窗口的大小。如图 7 所示(图中 avg 表示取平均值), 首先从数据开始选取一定范围的数据作为子数据集, 与原始父数据集进行 K-S 检验, 评估其相似性。然后逐步扩大子集范围, 直到 p 值大于 0.05, 并记录此时的子集长度, 即窗口大小 N 。

步骤 2: 计算阈值。根据上一步确定的滑动窗口大小, 选择 $\{\varepsilon_{i-N}, \dots, \varepsilon_i\}$ 范围内的数据计算阈值。由于风机正常状态下温度变化缓慢, 残值的波动也较小。因此, 本文在设置动态阈值时, 充分考虑了窗口内前一时段的残差变化趋势。计算式如下:

$$C_{dt} = \frac{\sum_{t=i-N}^i \sigma_{\text{Ref}}}{N} + 3S \quad (30)$$

式中: σ_{Ref} 为时刻 t 的残差值; S 为正常情况下残差的标准差。

步骤 3: 基于步骤 2, 对各时刻设定新的阈值, 并连接所有时刻的阈值, 形成自适应阈值曲线, 以

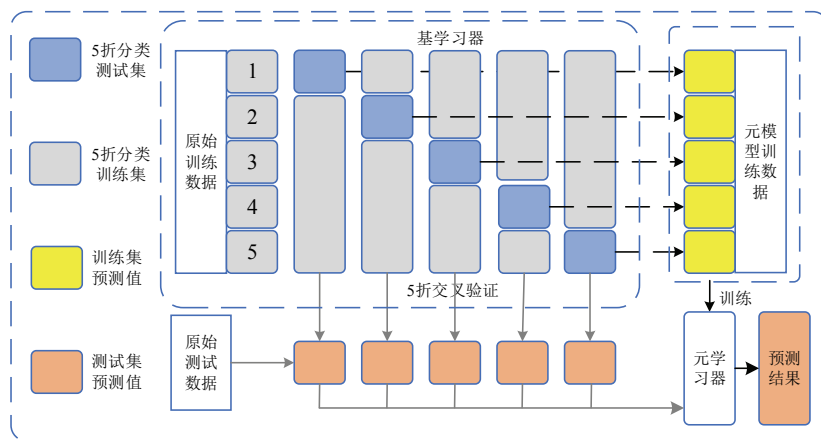


图 6 5 折交叉验证过程结构

Fig.6 Structure of the 5-fold cross-validation process

拟合残差趋势。

4 算例分析

4.1 数据来源与输入状态变量选择

本文选取国内某海上风电场实际运行数据进行分析。该海上风电场有 34 台 3 MW 双馈风力发电机，每台机组 SCADA 系统以每 10 min 1 次的频率采样，监测的连续变量和离散变量共计 106 项。本文选择 3 号风力发电机 2016 年全年实际运行数据为例，经过数据筛选和归一化处理，最终得到 3 号机 2016 全年的正常数据共 30 364 组。其中，24 291 组用作模型训练集，6 073 组用作测试集。

在选择基学习器的输入特征时，根据温度热模型分析，考虑了以下几个关键因素：1) 发电机的水冷散热：水冷系统能有效带走发电机产生的大部分热量；2) 变风速条件下的动态响应：风速变化会导致风力发电机各项参数产生动态响应，发电机输出功率增加时，损耗和温升也随之增加；3) 机械损耗与转速的关系：在高速电机中，机械损耗占总损耗比例较高；4) 湍流对换热的影响：湍流会影响对流换热系数，从而影响发电机的温度分布。

基于以上分析，本文选取了以下 7 个变量作为数据驱动模型的输入特征，用于预测定子绕组温度：发电机冷却水入水口温度、发电机冷却水出水口温度、发电机转速、风速、功率、湍流以及温度热模型求解的定子绕组温度。

4.2 机理温度有效性验证

为验证热模型计算所得的定子绕组温度的有效性，对比分析 GRA-LSTM 模型^[29]及其基于机理引导的改进模型。首先，通过 GRA 灰色关联度分析方法，计算出绕组温度与各状态变量间的灰色关联系数，评估它们的相关性。表 1 列出了与绕组温度高度相关的前 7 个状态变量。

基于上述分析，选取表 1 中 7 个变量作为 LSTM 模型的输入，并通过多次迭代训练和超参数优化对测试集进行预测。此外，将 GRA 分析中相关性最低的主轴承温度替换为热模型计算的机理绕组温度，并采用相同训练方式优化模型。图 8 显示，机理改进模型相比原模型拟合度更高。

由表 2 可以看出，改进 GRA-LSTM 模型评价指标 RMSE 和 MAE 比原 GRA-LSTM 模型分别降低 0.53、0.25； R^2 提高了 0.007，验证了机理部分推导的有效性。

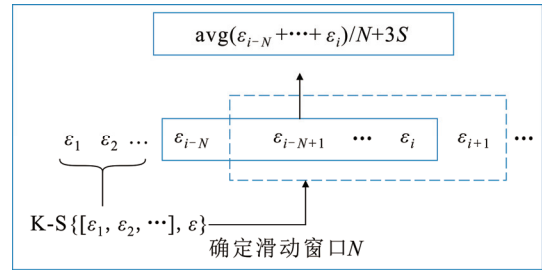


图 7 动态阈值设置示意图

Fig.7 Schematic diagram of dynamic threshold setting

表 1 定子绕组温度与各状态变量灰色关联度

Table 1 Gray correlation between winding temperature and each state variable

状态变量	关联度	状态变量	关联度
发电机冷却出水口温度	0.832 8	功率	0.742 1
发电机温度	0.810 5	发电机冷却入水口温度	0.709 1
发电机网测电压	0.780 7	主轴承温度	0.706 2
发电机电流	0.745 3		

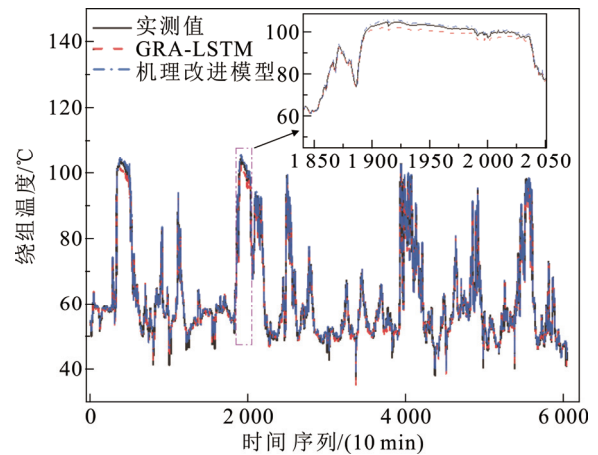


图 8 GRA-LSTM 和其改进模型的结果对比

Fig.8 Comparison of the results of GRA-LSTM and its improved models

表 2 GRA-LSTM 和其改进模型的预测性能对比

Table 2 Comparison of the prediction performance of GRA-LSTM and its improved models

预测模型	评价指标		
	RMSE	MAE	R^2
GRA-LSTM	1.765 6	1.030 5	0.985 8
改进 GRA-LSTM	1.232 7	0.780 5	0.993 1

4.3 机理引导的 Stacking 集成算法温度预测

本文选择极端梯度提升算法(eXtreme gradient boosting, XGBoost)、梯度提升决策树(gradient

boosted decision tree, GBDT)、随机森林(random forest, RF)、反向传播神经网络(back propagation, BP)、LASSO(least absolute shrinkage and selection operator)最小绝对收缩与选择算子回归、支持向量机(support vector machine, SVM)、K 近邻(K nearest neighbors, KNN)、卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)等多个模型作为初步的基学习器。涵盖了决策树、线性回归、神经网络、支持向量机及距离度量等不同类型的机器学习方法, 能够从多方面提升模型的预测能力和稳健性。通过网络搜索对超参数调优, 确保各模型在不同数据集上的最佳表现。各基模型参数如附录 E 表 E1 所示, 第二层元模型选择 Ridge 回归模型, 并使用其默认参数。

如 4.1 节所述, 本文算例预测模型选取的是某风电场 3 号机组 2016 年全年的正常运行数据进行训练, 训练数据量庞大。表 3 为各模型的拟合时间和 RMSE 指标比较, 拟合时间为各模型确定超参数后最优模型拟合时间。

在工程应用中, 预测模型仅需依赖于已训练好的模型。在预测阶段, 各模型均可在秒级快速得出预测结果, 因此这里主要比较的是模型的拟合效率。

同时, 为了减少过拟合风险, 对 Stacking 集成算法基学习器的选取进行皮尔逊相关性分析, 利用式(29)计算各基学习器预测的残差值的皮尔逊相关系数, 结果如图 9 所示。由图 9 可以看出, RF、GBDT、XGBoost 3 种算法间相关性较高, 这是因为 3 种算法本质上都是基于决策树的集成算法。相比之下, SVM、LASSO 聚焦于对预测中最重要的特征, 因此二者相关性也较高。BP、CNN、KNN 都是基于神经网络的算法, KNN 由于计算机制不同, 与其他算法表现出较大的差异性。基于上述分析, 3 种决策树算法中 XGBoost 拟合时间和评价指标均优于另外两种, SVM 拟合时间虽然比 LASSO 长, 但评价指标较优, 3 种神经网络算法 CNN 在拟合时间和评价指标均优于其他两类。因此, 本文选择 SVM、CNN、KNN、XGBoost 4 种算法作为 Stacking 集成算法的基学习器。

将训练完成的集成模型用于测试集预测, 如图 10 所示, Stacking 集成算法的预测曲线相比单一模型更为准确和稳定, 这表明通过结合多个基础学习器的预测结果, 集成模型提高了预测精度。

为更好地比较, 表 4 记录了各模型的评价指标。

表 3 各基模型拟合时间及评价指标比较

Table 3 Comparison of fitting times and evaluation indicators for each base model

基学习器	拟合时间/min	RMSE	R ²
XGBoost	1.49	1.098 5	0.994 4
GBDT	41.12	1.177 5	0.993 6
RF	24.17	1.150 2	0.993 9
BP	39.58	1.312 9	0.992 1
LASSO	1.64	1.760 7	0.985 8
SVM	6.83	1.726 2	0.986 4
KNN	0.48	1.350 1	0.991 7
CNN	9.09	1.294 9	0.992 8
RNN	12.56	1.524 4	0.984 5

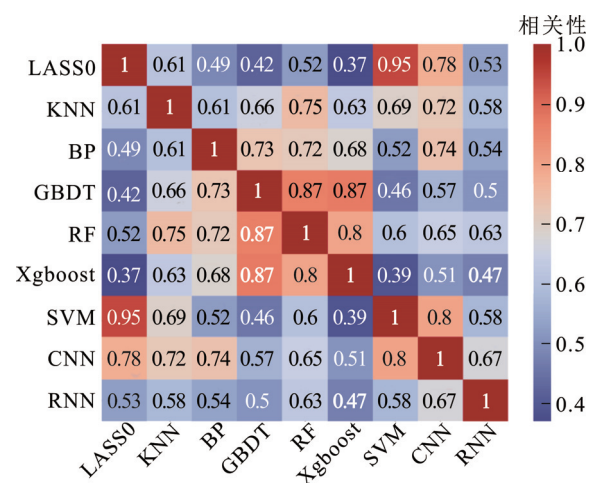


图 9 基学习器预测误差相关性分析

Fig.9 Correlation analysis of prediction errors of graph-based learners

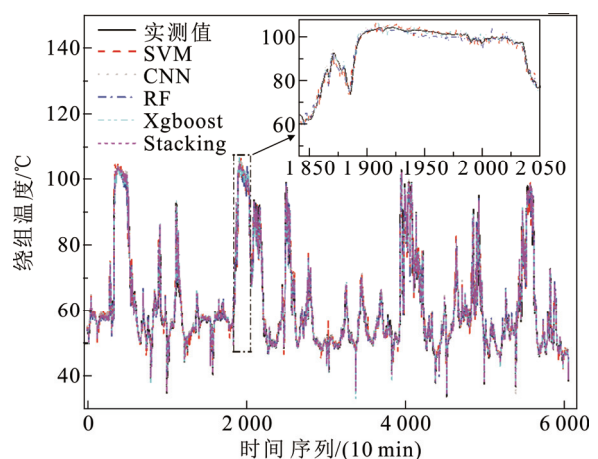


图 10 基学习器和 Stacking 集成算法预测对比

Fig.10 Comparison of base learner and Stacking integration algorithm predictions

由表 4 可知, Stacking 集成算法的 RMSE 和 MAE 的值分别比其他模型平均降低了 0.60 和 0.35; R²

表 4 Stacking 集成算法与基模型预测性能对比

Table 4 Comparison of Stacking integration algorithm and base model prediction performance

预测模型	评价指标		
	RMSE	MAE	R^2
SVM	1.726 2	1.100 0	0.986 4
CNN	1.294 9	0.875 4	0.992 8
KNN	1.350 1	0.884 3	0.991 7
XGB	1.098 5	0.727 3	0.994 4
Stacking	0.767 1	0.543 6	0.997 1

值平均提高了 0.006。因此，本文所提融合模型的预测性能和精度方面均有较好的表现。

4.4 早期故障预警

为验证本文所提模型预警效果，将异常数据清理后得到故障预警数据集。利用 K-S 检验原理，确定自适应窗口大小为 32 个采样点，利用式(30)设置自适应阈值。为方便比较，设定窗宽 w 为 10，窗口移动增量为 1，移动平均法计算得到固定告警阈值为 3.244。

案例 1: 3 号机于 2018 年 1 月 23 日 09:00 机组发现水冷循环系统故障。根据图 11 中状态指标残差绝对值随时间的变化曲线，状态指标在 2018 年 1 月 20 日 17:00 首次越过动态阈值，比静态阈值提前了 7.7 h，比故障列表中的记录时间提前了 64 h。实际预警信号发生后，由于大风无法出海维修，通过采取措施让机组采用限功率运行方式，状态指标值明显下降，但仍有部分样本点超过告警阈值，验证了模型的有效性，可有效避免重大故障发生。

案例 2: 7 号机组 2018 年 11 月 07 日 19:20 机组发现发电机轴承温度高故障。根据图 12 中状态指标残差绝对值随时间的变化曲线，状态指在 2018 年 11 月 7 日 06:10 首次越过动态阈值，且残差值出现抖动上升趋势，比静态阈值提前了 2.5 h，比故障列表中的记录时间提前了 13.2 h。实际故障发生后，台账故障记录显示 7 号机组于 2018 年 11 月 8 日 09:20 再次出现发电机绕组故障，再次验证了模型的有效性。同时，不同机组的预警结果表明，该模型具有较强的泛化能力。

案例 3: 图 13 展示了 18 号机组在 2017 年 7 月 8 日 12:00 至 9 日 12:00 的 24 h 期间静动态阈值预测残差对比。在该数据集中，海上风电场的平均气温超过 39 °C。从图中可以看出，在初始阶段，动态阈值短时间内低于实际预测残差，而静态阈值则在后续较长时间内始终低于残差值。这表明，动态

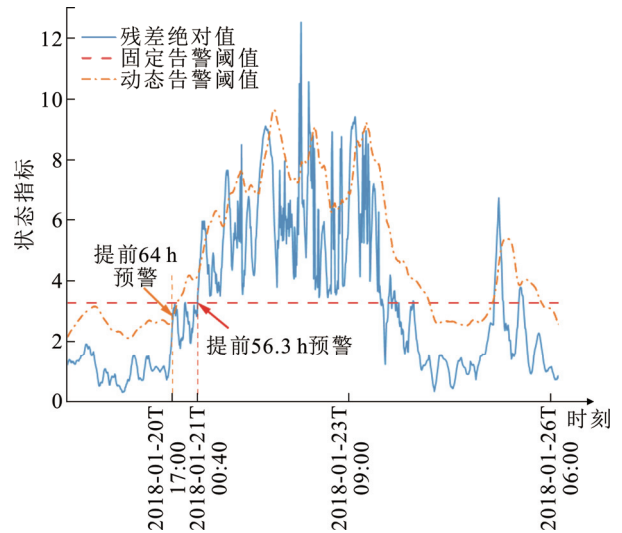


图 11 异常工况预警案例 1

Fig.11 Abnormal working condition warning case 1

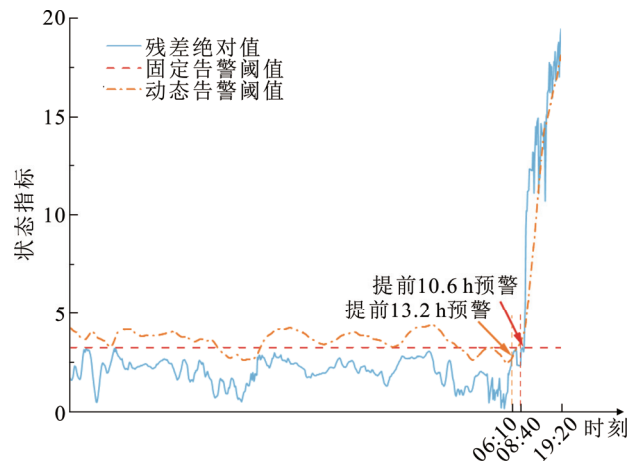


图 12 异常工况预警案例 2

Fig.12 Abnormal working condition warning case 2

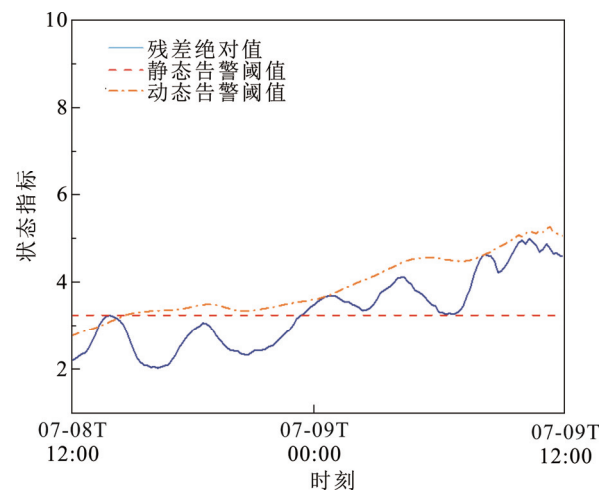


图 13 高温环境下动静态阈值对比案例 3

Fig.13 Comparison of dynamic and static thresholds in high temperature environments case 3

阈值能够更有效地适应数据分布的变化, 特别是在风机运行过程中, 它能及时响应由外部环境变化引起的残差波动, 从而提高了故障预警的准确性和提前性。

5 结论

1) 比较了 GRA-LSTM 模型和其改进模型, 其中改进模型以温度热模型求解温度作为重要输入特征之一。结果表明, 改进模型的评价指标 RMSE 和 MAE 分别降低 30.17% 和 24.26%, 验证了温度热模型分析的有效性。

2) 对比了 Stacking 集成算法和单一算法的预测性能, 结果显示 Stacking 集成算法的 RMSE 和 MAE 值比其他模型平均降低了 43.90% 和 39.38%, 预测拟合度达到 99.71%, 验证了所提融合模型具有更优的预测性能和更精确的预测结果。

3) 以国内某海上风电场的真实运行数据对所提故障预警方法进行验证, 结果表明该方法能够对发电机的异常工况实现提前 13.2~64 h 预警。动态阈值能够灵敏地反映外部环境变化, 在高温环境下, 对残差波动的影响具有较强的适应性。该方法展现出较强的泛化性, 可有效避免重大或不可逆故障的发生, 减少故障停机损失, 为海上风电安全可靠运行提供重要支撑。

附录见本刊网络版(<http://hve.epri.sgcc.com.cn>)。

参考文献 References

- [1] FAN Q X, WANG X, YUAN J, et al. A review of the development of key technologies for offshore wind power in China[J]. *Journal of Marine Science and Engineering*, 2022, 10(7): 929.
- [2] 刘 芳, 刘 威, 汪浩东, 等. 高比例新能源电力系统振荡机理及其分析方法研究综述[J]. *高电压技术*, 2022, 48(1): 95-113.
LIU Fang, LIU Wei, WANG Haodong, et al. Review on oscillation mechanism and analysis methods of high proportion renewable energy power system[J]. *High Voltage Engineering*, 2022, 48(1): 95-113.
- [3] CARROLL J, MCDONALD A, MCMILLAN D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines[J]. *Wind Energy*, 2016, 19(6): 1107-1119.
- [4] 赵洪山, 林诗雨, 孙承妍, 等. 考虑多时间尺度信息的风力发电机滚动轴承故障预测[J]. *中国电机工程学报*, 2024, 44(22): 8908-8919.
ZHAO Hongshan, LIN Shiyu, SUN Chenyan, et al. Fault prediction of wind turbine rolling bearing considering multi-time scale information[J]. *Proceedings of the CSEE*, 2024, 44(22): 8908-8919.
- [5] 孙丽玲, 王艳娟. 电网电压不对称骤升时双馈风力发电机定子磁链暂态全过程及控制策略研究[J]. *高电压技术*, 2019, 45(7): 2160-2166.
SUN Liling, WANG Yanjuan. Study on transient whole-process and control strategy of stator flux in doubly-fed induction generator with asymmetric voltage swell[J]. *High Voltage Engineering*, 2019, 45(7): 2160-2166.
- [6] ZHANG F H, CHEN M S, ZHU Y Z, et al. A review of fault diagnosis, status prediction, and evaluation technology for wind turbines[J]. *Energies*, 2023, 16(3): 1125.
- [7] 宋 海, 穆塔里夫·阿赫迈德, 尼加提·玉素甫, 等. 基于热网络与数值仿真模拟的风电轴承瞬态温度分析[J]. *轴承*, 2023(6): 51-58.
SONG Hai, MUTELLIP A, NIJAT Y, et al. Transient temperature analysis of wind turbine bearings based on thermal network and numerical simulation[J]. *Bearing*, 2023(6): 51-58.
- [8] 阚超豪, 张 恒, 方裕超, 等. 基于定子电流法的绕线转子无刷双馈发电机偏心故障研究[J]. *中国电机工程学报*, 2024, 44(8): 3260-3268.
KAN Chaohao, ZHANG Heng, FANG Yuchao, et al. Study on eccentricity fault of brushless doubly-fed generator with wound rotor based on motor current signature analysis[J]. *Proceedings of the CSEE*, 2024, 44(8): 3260-3268.
- [9] 吕世文, 张宏立, 马 萍, 等. 基于改进深度置信网络的风力发电机在线故障诊断[J]. *机械设计与制造*, 2025(5): 171-176.
LÜ Shiwen, ZHANG Hongli, MA Ping, et al. On-line fault diagnosis of wind turbines based on improved deep belief network[J]. *Machinery Design & Manufacture*, 2025(5): 171-176.
- [10] TANG H H, ZHANG K, WANG B, et al. Early bearing fault diagnosis for imbalanced data in offshore wind turbine using improved deep learning based on scaled minimum unscented Kalman filter[J]. *Ocean Engineering*, 2024, 300: 117392.
- [11] ZHANG F H, ZHU Y Z, ZHANG C J, et al. Abnormality detection method for wind turbine bearings based on CNN-LSTM[J]. *Energies*, 2023, 16(7): 3291.
- [12] 王进花, 韩金玉, 曹 洁, 等. 基于 AM 和 CNN 的多级特征融合的风力发电机轴承故障诊断方法[J]. *太阳能学报*, 2024, 45(5): 51-61.
WANG Jinhua, HAN Jinyu, CAO Jie, et al. Fault diagnosis method of wind turbine bearing based on am and CNN multi-level feature fusion[J]. *Acta Energetica Solaris Sinica*, 2024, 45(5): 51-61.
- [13] WANG H, XIE H, LIU S W, et al. A multi-view spatio-temporal feature fusion approach for wind turbine condition monitoring based on SCADA data[J]. *IEEE Access*, 2024, 12: 43948-43957.
- [14] RAMA V S B, HUR S H, YANG J M. Short-term fault prediction of wind turbines based on integrated RNN-LSTM[J]. *IEEE Access*, 2024, 12: 22465-22478.
- [15] LIU J R, YANG G T, LI X L, et al. A deep generative model based on CNN-CVAE for wind turbine condition monitoring[J]. *Measurement Science and Technology*, 2023, 34(3): 035902.
- [16] RAMA V S B, HUR S H, YANG J M. Predictive maintenance and anomaly detection of wind turbines based on bladed simulator models[J]. *IFAC-PapersOnLine*, 2023, 56(2): 4633-4638.
- [17] 蒙 康, 滕 伟, 彭迪康, 等. 运行机理与数据双驱动的风电齿轮箱系统故障预警[J]. *中国机械工程*, 2023, 34(12): 1476-1485.
MENG Kang, TENG Wei, PENG Dikang, et al. Operating mechanism and data driven approach for fault alarm of wind turbine gearbox systems[J]. *China Mechanical Engineering*, 2023, 34(12): 1476-1485.
- [18] 董立卓, 张思琪, 张 钊, 等. 机理-数据混合驱动的叶片加工变形预测方法[J]. *航空学报*, 2024, 45(13): 629037.
DONG Lizhuo, ZHANG Siqi, ZHANG Zhao, et al. Prediction method of blade machining deformation driven by mechanism-data hybrid[J]. *Acta Aeronautica et Astronautica Sinica*, 2024, 45(13): 629037.
- [19] 胡敏强, 黄学良, 黄允凯, 等. *电机学*[M]. 3 版. 北京: 中国电力出版社, 2014.
HU Minqiang, HUANG Xueliang, HUANG Yunkai, et al. *Electric*

- machinery[M]. 3rd ed. Beijing, China: China Electric Power Press, 2014: 7.
- [20] 曹梦楠, 邱颖宁, 冯延晖, 等. 基于等效热网络法的风力发电机故障诊断[J]. 工程热物理学报, 2019, 40(2): 306-313.
CAO Mengnan, QIU Yingning, FENG Yanhui, et al. Fault diagnosis of a wind generator based on equivalent thermal network method[J]. Journal of Engineering Thermophysics, 2019, 40(2): 306-313.
- [21] 张建忠, 姜永将. 基于等效热网络法的定频双转子永磁风力发电机的热分析[J]. 电工技术学报, 2015, 30(2): 87-97.
ZHANG Jianzhong, JIANG Yongjiang. Thermal analysis of constant frequency double rotor permanent magnet generator based on equivalent thermal network method[J]. Transactions of China Electrotechnical Society, 2015, 30(2): 87-97.
- [22] 何磊. 基于热阻网络法的电机温度场分析[D]. 上海: 同济大学, 2014.
HE Lei. Thermal analysis of electromotor based on thermal network method[D]. Shanghai, China: Tongji University, 2014.
- [23] TOGUN H, ABDULRAZZAQ T, KAZI S N, et al. A review of studies on forced, natural and mixed heat transfer to fluid and nanofluid flow in an annular passage[J]. Renewable and Sustainable Energy Reviews, 2014, 39: 835-856.
- [24] 陈万玮, 刘豪. 混合转子双定子风力发电机的温度场计算[J]. 机械设计与制造, 2024, 406(12): 81-86, 90.
CHEN Wanwei, LIU Hao. Temperature field calculation of hybrid rotor dual-stator wind generator[J]. Machinery Design & Manufacture, 2024, 406(12): 81-86, 90.
- [25] 黄国治, 傅丰礼. Y2系列三相异步电动机技术手册[M]. 北京: 机械工业出版社, 2004.
HUANG Guozhi, FU Fengli. Y2 series three-phase asynchronous motor technical manual[M]. Beijing, China: Machinery Industry Press, 2004.
- [26] 郑颖颖, 李鑫, 陈延旭, 等. 基于 Stacking 多模型融合的极端天气短期风电功率预测方法[J]. 高电压技术, 2024, 50(9): 3871-3882.
ZHENG Yingying, LI Xin, CHEN Yanxu, et al. Short-term wind power forecasting method in extreme weather based on Stacking multi-model fusion[J]. High Voltage Engineering, 2024, 50(9): 3871-3882.
- [27] YUAN T K, SUN Z F, MA S H. Gearbox fault prediction of wind turbines based on a stacking model and change-point detection[J]. Energies, 2019, 12(22): 4224.
- [28] WANG X Y, MAKIS V. Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov-Smirnov test[J]. Journal of Sound and Vibration, 2009, 327(3/5): 413-423.
- [29] 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): 2373-2382.
WEI Shurong, ZHANG Xin, FU Yang, et al. Early fault warning and diagnosis of offshore wind DFIG based on GRA-LSTM-Stacking model[J]. Proceedings of the CSEE, 2021, 41(7): 2373-2382.



WEI Shurong
Ph.D., Professor
Corresponding author

魏书荣(通信作者)

1980—, 女, 博士, 教授

主要从事海上风电规划设计, 故障诊断和并网运行等方面工作

E-mail: wsrmail@163.com



FU Yang
Ph.D., Professor

符杨

1968—, 男, 博士, 教授

主要从事电力系统分析、风力发电并网技术及海上风电优化设计等方面工作

E-mail: mfdong@162.com

收稿日期 2024-11-28 修回日期 2025-03-06 编辑 曹昭君

附录 A

表 A1 部分模型参数

Table A1 Selected model parameters

参数名称	数值	单位	参数名称	数值	单位
机壳外径	1 270	mm	定子绕组电阻	2.6	mΩ
绝缘层厚度	0.4	mm	转子绕组电阻	2.4	mΩ
定子外直径	1 250	mm	铁芯密度	7650	kg/m ³
定子内直径	762	mm	空气导热系数	0.025 9	W/(m·K)
定子槽深	129	mm	绕组导热系数	380	W/(m·K)
定子槽数	72		铁芯导热系数	25	W/(m·K)
定子齿宽	18	mm	转轴直径	236	mm
转子外直径	758	mm	机壳导热系数	168	W/(m·K)
转子内直径	236	mm	轴导热系数	63	W/(m·K)
转子槽深	90	mm	绝缘层导热系数	0.3	W/(m·K)
转子槽数	60		机壳轴向长度	2 002	mm

附录 B

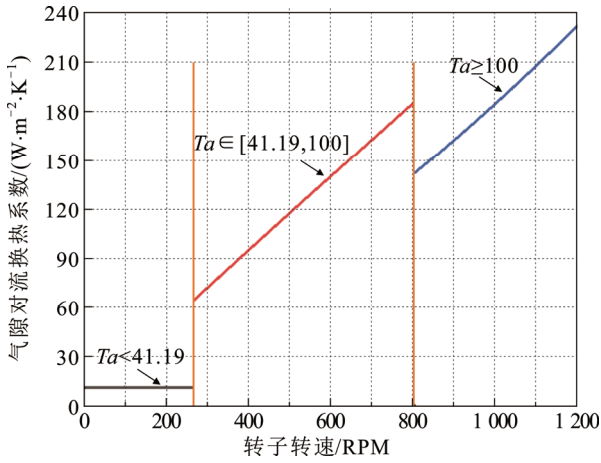


图 B1 气隙对流换热系数随转子转速变化

Fig.B1 Variation of air gap convection heat transfer coefficient with rotor speed

附录 C

R^2 、RMSE 和 MAE 的计算式分别为：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \in [0, 1] \quad (C1)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \in [0, +\infty) \quad (C2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \in [0, +\infty) \quad (C3)$$

式中， y_i 为第 i 个时刻定子绕组温度的实际值； y'_i 为第 i 个时刻定子绕组温度的预测值； \bar{y}_i 为定子绕

组温度平均值。 R^2 表示预测模型的拟合优度，其值越接近 1 表示拟合程度越高；RMSE 和 MAE 表示预测结果准确性，其值越小代表预测结果越准确。

附录 D

移动平均法设定固定阈值计算流程：

窗口内的平均绝对误差计算表达式为：

$$T_{\text{MAE}(t_i)} = \frac{1}{w} \sum_{j=t_i-w}^{t_i} |T_j - T'_j| \quad (D1)$$

式中， w 为滑动窗口的宽度； T_j 和 T'_j 分别为 j 时刻模型的实际值和预测值。 $T_{\text{MAE}(t_i)}$ 的大小反映绕组温度的实时状态，数值越大偏离正常工况越明显，反之则表明运行更稳定。本文采用统计学常见的区间估计方法，通过计算置信上限来设定预警阈值，公式如下：

$$T_{\text{th}} = \overline{T_{\text{MAE}(t_i)}} + \frac{S}{\sqrt{c}} t_a (c-1) + 3S \quad (D2)$$

式中， $\overline{T_{\text{MAE}(t_i)}}$ 为机组正常运行时的状态指标均值； S 为状态指标的标准差； c 为滑动窗口数量， $c = (n-b)/\omega$ ， n 为输入样本总数， b 为窗口宽度， ω 为移动步长。

附录 E

表 E1 各基模型超参数选择

Table E1 Hyperparameter selection for each base model

模型名称	超参数集	模型名称	超参数集
XGBoost	弱学习器数 1 000 树深度 5 子节点中实例的最小 权重和 1 样本采样比例 0.8 特征采样比例 0.8	RF	弱学习器数 130 叶节点最小样本 数 13 树深度 7
GBDT	弱学习器数 70 叶节点最小样本数 15 树深度 8	SVM	惩罚参数 1 150 核函数系数 10^{-4} 核函数为高斯径 向基函数
LASSO	L1 正则化系数 1	KNN	近邻数 9
RNN	第一层神经元数 64 第二层神经元数 100 学习率 0.001	CNN	卷积核数 64 卷积核大小 3 全连接层神经元 数量 100 学习率 0.001
BP	学习率 0.001 神经元数 64		