

# 风电机组定子绕组温度传感器状态自确认研究<sup>\*</sup>

周 凌 黄 倩 曾进辉 黄浪尘 龙霞飞

(湖南工业大学电气与信息工程学院 株洲 41200)

**摘 要:**针对提高智慧风场风电机组运行可靠性和传感器状态自确认问题,以风电机组定子绕组温度传感器为研究对象,提出了一种融合多源信息与智能算法的传感器状态自确认方法。首先,基于灰色关联分析理论,利用传感器的相关性和信息融合技术,通过计算某风场异常定子绕组温度传感器与同机同类传感器之间的灰色关联度,实现传感器异常状态识别。其次,利用皮尔逊相关性和专家系统判断,筛选出和定子绕组温度传感器关联性较强的参数,建立长短期记忆神经网络(long short-term memory, LSTM)多参数输入单输出异常数据恢复模型,并通过麻雀算法(sparrow search algorithm, SSA)对LSTM模型的超参数进行优化。为验证数据恢复模型的精度,通过模拟异常数据恢复表明该模型的精度达到了99.69%。最后对该定子绕组温度传感器异常数据进行了恢复,基于贝叶斯动态不确定度评估方法,对恢复数据进行置信度分析,从而实现对传感器状态的动态自确认。

**关键词:**风电机组;温度传感器;自确认;信息融合技术;异常数据恢复

**中图分类号:** TP183;TK83;TN98

**文献标识码:** A

**国家标准学科分类代码:** 420.66

## Research on self-validation of temperature sensor status for stator winding of wind turbine

Zhou Ling Huang Qian Zeng Jinhui Huang Langchen Long Xiafei

(College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412002, China)

**Abstract:** In order to improve the operational reliability of wind turbines in smart wind farms and the self-validation of sensor status, a novel self-validation method for sensor status is proposed using the stator winding temperature sensor of wind turbines as an example. First, based on grey relational analysis theory, and utilizing sensor correlation and information fusion technology, the grey correlation degree between the abnormal stator winding temperature sensor of a specific wind field and the same type of sensor on the same machine is calculated to achieve sensor anomaly state recognition. Second, using Pearson correlation and expert system judgment, parameters with strong correlation to the stator winding temperature sensor are identified. A long short-term memory (LSTM) multi-parameter input, single-output abnormal data reconstruction model is then established and optimized using the sparrow search algorithm (SSA) to improve the model's accuracy. To verify the model's reconstruction accuracy, simulations of abnormal data recovery showed that the accuracy reached 99.69%. Finally, the abnormal data of the stator winding temperature sensor was recovered, and the dynamic validation uncertainty of the recovered data was calculated using a Bayesian algorithm, achieving self-validation of the sensor's state.

**Keywords:** wind turbines; temperature sensor; self-validation; information fusion technology; restructure

## 0 引 言

大型风电机组都安装有各种监控系统,通过监控系统中各种传感器可获取风电机组的实时环境参数和状态

参数数据<sup>[1]</sup>。而传感器种类和数量繁多,当运行时间过长时,传感器出现状态异常的几率大大增加,研究表明风电机组故障40%以上由传感器及相关系统的故障引起的<sup>[2]</sup>。传感器的故障会直接导致控制系统收到错误信息而发出错误的控制指令,甚至危及设备安全和人员安全,

从而增加风电场的运维成本,因此有必要对传感器的状态进行识别与评估<sup>[3]</sup>。但目前对风电场风电机组传感器的状态识别与评估研究还比较少,而且基本上都是集中在传感器输出信号异常上,很少涉及到传感器自身的状态异常。

与传统传感器相比,传感器状态自确认不仅能实现测量值的输出,而且当输出结果出现状态异常时,能检测出异常状态的类型,并能对异常值进行最佳估计,用最佳估计值代替异常测量值,并以确认的不确定度表示当前确认值的准确度,完成状态自确认<sup>[4]</sup>。传感器自确认的理念一旦被提出,便引起了众多学者的关注,并在理论研究、技术研究以及系统结构等领域都有了很好的研究成果。文献[5]提出了一种基于测量数据统计相关性的传感器自诊断设计方法。利用传感器测量数据建立其统计关系模型,借助自编码器提取传感器数据特征并将其编码为二进制形式。在同时考虑传感器测量数据统计独立和统计相关两种情况下,实现有参数和无参数的传感器故障自诊断,两种模式下测量系统中健康传感器的故障检测后验概率达到了 0.92。文献[6]针对金属氧化物半导体气体传感器的状态异常问题,提出了基于 T 分布随机邻域嵌入和随机森立的气体传感器自确认方法,并开发了一套气体监测实验系统,结果表明,与其他方法相比,该方法异常诊断准确率为 99.78%。文献[7]针对高 g 加速度传感器的损坏会使得航空航天热冲击试验失败的问题,提出了一种基于机器学习法来实现传感器的自确认。采用组合集成学习模型识别传感器的健康状态,深度神经网络实现异常值的重构,提高了热冲击实验的可靠性。从上述应用可以看出,通过广大研究者的不断努力,自确认传感器的应用越来越广泛,涉及的领域也越来越多元,大大促进了自确认传感器状态异常诊断技术、数据恢复与数据重构等技术的发展,同时也给其他场合和其他类型自确认传感器的研究发展提供了技术支持和思路拓展,极大地丰富了自确认传感器理论。风电场传感器不仅在种类上繁多,数量也十分庞大,同时传感器在整个风电系统中的作用十分重要。但是目前很少有风电场传感器自确认方面的研究,因此对建设智慧风场而言,风电场传感器进行状态自确认显得十分必要。

传感器状态自确认技术主要包括状态检测与隔离技术、异常状态辨识技术、数据重构恢复技术和确认不确定度计算,通过异常状态检测判断传感器是否工作在正常范围内,当检测到传感器不是正常测量时,识别异常状态类型并对异常状态数据进行隔离和数据恢复,保证系统能正常运行<sup>[8]</sup>。传感器异常状态检测技术是实现传感器状态自确认的基础,状态检测系统的核心是检测方法,目前传感器状态检测方法的主要方法有硬件冗余法、数学模型法、信号处理法、专家知识法和神经网络法等<sup>[9-12]</sup>。

但大部分研究只是针对传感器的信号异常识别,而不是准确意义上的传感器状态异常识别,两者之间虽然有共性之处,但是存在本质的区别<sup>[13]</sup>。针对传感器信号异常识别和状态异常识别的区别,除了利用单个传感器参数之间的关联特性之外,还需要充分利用多个相同类型传感器之间的相关性。

传感器测量值输出结果的正确是保证系统正常运行的关键,当传感器发生状态异常时,数据重构恢复技术能根据传感器输出的正常数据估计和替换异常数据输出,从而使系统能继续稳定运行。其核心思路是利用传感器之间的相关性和关联性,根据正常传感器或者是历史正常数据来恢复故障的传感器数据<sup>[14-17]</sup>。而风电场传感器参数之间不仅具有相关性,而且数据不会突然发生跳变,具有强关联时间和空间序列信息,因此可以很好的实现异常数据恢复。

确认的不确定度(the validated uncertainty, VU)表示测量不确定度或确认测量值的可能误差,误差主要有校准误差、数据采集误差和测量数据处理误差。传感器输出测量值的不确定度不仅受到传感器本身的影响,还包括环境因素和被测对象的影响,要综合考虑各种导致误差因素。不确定度是对当前传感器输出值准确度的一个定量表征,也可以看成是对测量值“质量”的好坏评定,不仅能反映当前传感器是否出现异常状态,也能反映出该异常状态对测量值准确度的影响程度。

本文以定子绕组温度传感器为例,提出了风电机组传感器状态自确认的方法,解决传感器状态识别问题。首先利用传感器的相关性和信息融合技术,计算异常定子绕组温度传感器与同机同类传感器之间的灰色关联度,实现传感器状态异常确定。然后通过皮尔逊相关系数法和专家系统判断,得到定子绕组温度传感器关联性较强的参数,建立基于麻雀优化算法(sparrow search algorithm, SSA)优化长短期记忆神经网络(long short-term memory, LSTM)的多输入单输出风电机组传感器状态异常输出数据恢复模型。为验证 SSA-LSTM 对异常输出数据恢复精度,在正常数据采集与监控(supervisory control and data acquisition, SCADA)系统数据中加入故障数据,实现模拟故障,计算出数据恢复的精度,验证了 SSA-LSTM 模型的先进性。并为了验证模型的性能,对比分析了深度置信网络(deep belief network, DBN)模型、鲸鱼优化算法(whale optimization algorithm, WOA)优化 LSTM(WOA-LSTM)模型和未优化的 LSTM 模型,证明 SSA-LSTM 模型的准确度更高。最后用动态不确定度评估方法(贝叶斯算法)计算恢复数据的确认不确定度值,从而实现传感器的状态自确认<sup>[18]</sup>。本文方法不仅实现了对于异常数据的高精度恢复,还在动态不确定度约束下,确保系统级故障信号不被掩盖,提升了传感器数据的

使用安全性与系统容错性,为风电机组的健康运行评估与远程运维提供重要支持。

## 1 传感器状态自确认理论

自确认传感器(self-validating sensor, SEVA sensor)的模型如图 1 所示。由图 1 可知,与传统的传感器相比,传感器自确认的优势是增加了故障诊断环节和输出数据生成环节,输出的量也比传统的传感器更加丰富和优化,其可靠性和功能性也要比传统传感器要强大。

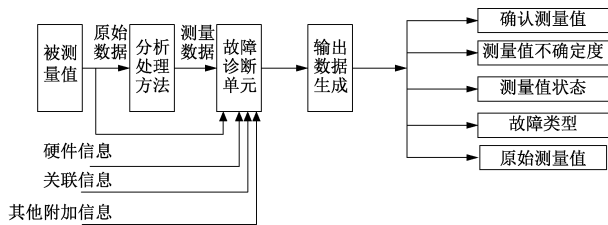


图 1 传感器状态自确认模型

Fig. 1 Sensor state self-validation model

自确认传感器模型应具备 3 个方面的基础:自诊断检测技术、数字通信技术应用和计量科学的不确定性分析,三者相互关系如图 2 所示。与传统传感器相比,自确认传感器不仅有获取测量信号的能力,同时测量输出数据通过 3 个特征来描述度量质量。

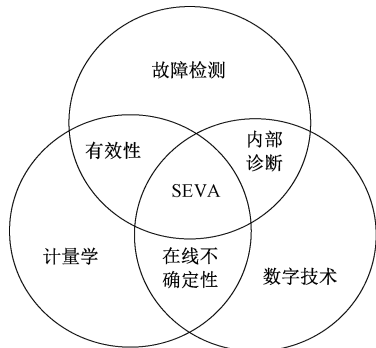


图 2 SEVA 模型的组成

Fig. 2 Composition diagram of SEVA model

## 2 风电机组定子绕组温度传感器异常状态识别

### 2.1 传感器信息融合技术

风电场风电机组传感器不仅种类繁多,数量也庞大。同类传感器之间存在相关性,不同类型的传感器之间也存在相关性。对于同类型传感器具有同质和同精度,因此可以充分利用多机同类传感器和单机同类传感器的信息,采用关联性相关数值指标进行传感器信息特征提取,

并进行分析得出最终结论。同时传感器本身异常是局部性的,只会出现本传感器的数据异常,虽然和其他传感器具有关联性,但并不会对其他传感器的输出数据产生很大影响。而风电机组本体异常是全局性的,会影响多个类型传感器的数据输出。因此当传感器出现异常时,而关联传感器和同类型传感器都是正常情况,通过异常传感器数据特征分析,来辨别异常是由传感器本身产生的。基于特征的信息融合能把具有局部性和全局性数据特征提取出来,特征提取可以采用关联性指标来实现。

而风电场风电机组传感器数据是实时变化的,通过灰色关联度分析可以很好分析各个传感器之间的数值关系,灰色关联度越高,表明关联性越强。当某个传感器本身发生故障时,与其他传感器之间的灰色关联度相比于正常时会出现比较大的变化。反过来,如果是风电机组本体发生故障,所有传感器的输出数据都会受到影响,相应的这些传感器之间的灰色关联度不会发生较大变化。

灰色系统关联分析从以下几个方面来实现:

确定参考数列和比较数列。本文的参考数列为出现异常的传感器数据序列,比较数列为同类型传感器数据序列。

无量纲化参考数列和比较数列。通过无量纲处理,可以提高比较结果的准确度。

1) 求参考数列与比较数列的灰色关联系数  $\xi_i$ 。二条曲线间的形状相差多少可以通过关联程度表现出来。因此可以用关联程度来表示二条曲线的差值大小。如果参考数列  $X_0$  来表示,其中有比较数列  $X_1, X_2, \dots, X_i, i$  表示比较数列传感器的数量。各比较数列与参考数列在各个时刻的关联系数  $\xi_{0i}(k)$  如式(1)所示。

$$\xi_{0i}(k) = \frac{\Delta(\min) + \rho\Delta(\max)}{\Delta_{0i}(k) + \rho\Delta(\max)} \quad (1)$$

式(1)表示第  $i$  个传感器的测量值在第  $k$  时刻与比较数列传感器的灰色关联度,  $\rho$  为分辨系数,一般取  $0.5^{[19]}$ 。  $\Delta(\min)$  是待检异常传感器与其他比较传感器绝对差值的最小值,  $\Delta(\max)$  是待检异常传感器与其他比较传感器绝对差值的最大值。  $\Delta_{0i}(k)$  表示异常传感器与各比较传感器数列上的点与点之间的绝对差值。

2) 求关联度  $R_i$ 。定义比较传感器与异常传感器二个数列在不同时间(即曲线中的各点)的关联程度值为关联系数,通过求解各个时刻的关联系数平均值,把这个平均值叫做关联度,其数量当成是比较传感器与异常传感器二个数列间关联程度,关联度  $R_i$  公式如式(2)所示。

$$R_i = \frac{1}{N} \sum_{k=1}^N \xi_i(k) \quad (2)$$

式中:  $N$  为总的关联系数数目<sup>[20]</sup>。



2.2 基于信息融合技术的定子绕组温度传感器异常状态识别

为准确判断传感器异常数据是否会导致风电机组的故障,采用基于信息融合技术的多传感器数据分析方法。当某个传感器的温度数据异常时,首先计算该传感器与其他关键传感器数据的关联性。如果异常数据与多个关键传感器的输出存在较大差异,说明可能是传感器本身出现了故障;反之,如果多个传感器同时出现异常,则可能意味着风电机组的某个部件发生了故障。结合信息融合技术,能够动态监测风电机组的状态,识别局部故障与全局故障的关联。

数据来源某风场提供的 23 台风电机组 SCADA 数据,其中选取 2016 年 7 号风电机组的数据来进行分析。根据信息融合技术原理,当本体出现异常时,多个传感器的测量数据都会出现异常,而传感器本身出现异常时,只会对本身测量值产生较大的变动,虽然会影响关联传感器的输出,但影响并不会很大,而且会根据关联程度的强弱体现测量值的变化。风电机组定子三相绕组总共有 6 个传感器,当本体出现异常时,6 个传感器都会受到影响,关联灰色度与正常时相差不会很大。当某一传感器出现异常,而其他 5 个正常时,该异常传感器与其他 5 个正常传感器之间的关联灰色度会出现较大的变化。根据

传感器信息融合技术,计算正常情况下与 7 号风电机组处于同一位置的 5 号、6 号、8 号、9 号风电机组中定子绕组 6 号传感器与其他 5 个传感器之间的灰色关联度,选取数据区间为发生异常数据段 15 818~35 818,该 4 个风电机组中定子绕组 6 号传感器与其他 5 个绕组温度传感器灰色关联度如表 1 所示。

表 1 4 个风电机组中定子绕组 6 号传感器与其他 5 个绕组温度传感器灰色关联度  
Table 1 Grey correlation degree between stator winding No. 6 sensor and other 5 winding temperature sensors in 4 wind turbines

风电机组	传感器 1	传感器 2	传感器 3	传感器 4	传感器 5
5	0.896 6	0.935 0	0.851 0	0.961 3	0.844 8
6	0.811 9	0.946 5	0.927 7	0.965 9	0.921 4
8	0.836 6	0.964 8	0.916 0	0.985 2	0.974 1
9	0.877 3	0.791 8	0.833 9	0.944 7	0.862 9

由表 1 可得定子绕组 6 号传感器温度测量值与其他 5 个绕组温度测量值之间的关联度正常值范围。并计算 2016 年 7 号风电机组定子绕组 6 号传感器与其他 5 个定子绕组温度测量值之间的灰色关联度,计算数据范围为 15 818~35 818,结果如表 2 所示。

表 2 正常状态下定子绕组 6 与其他 5 个绕组温度传感器灰色关联度范围和 2016 年 7 号风电机组定子绕组温度传感器之间灰色关联度

Table 2 Gray correlation degree range between stator winding 6 and other 5 winding temperature sensors under normal conditions and grey correlation between stator winding temperature sensors of No. 7 wind turbine generator unit in 2016

状态	传感器 1	传感器 2	传感器 3	传感器 4	传感器 5
$R$ (正常状态)	0.811 9~0.896 6	0.791 8~0.964 8	0.833 9~0.927 7	0.944 7~0.985 2	0.844 8~0.974 1
$R$ (2016 年 7 号)	0.665 2	0.657 0	0.658 1	0.652 8	0.657 2

由表 2 可知,2016 年 7 号风电机组定子绕组 6 号温度传感器与其他 5 个传感器在该段时间之内的灰色关联度明显发生了变化,不在正常值范围之内,表明定子绕组 6 号温度传感器与其他 5 个传感器的关联性发生了破坏,而其他传感器之间关联度没有出现异常,由于各个传感器所处的环境是一样的,排除其异常是由环境因素引起的。根据信息融合技术的判断依据,基本可以判定传感器出现异常状态的原因是传感器本身所引起的,从而实现传感器异常状态识别。

3 传感器异常状态数据恢复模型的构建

3.1 LSTM 网络

LSTM 是在循环神经网络 (recurrent neural network, RNN) 的基础上改进过来的,将上一个时刻的输出加上本

时刻的输入,经过一个双曲正切激活层 (hyperbolic tangent function, tanh),就成了下一个时刻的输出。图 3 为 RNN 模型多输入单输出形式图,该形式是在基本模型进行扩展,在最后一个时间点的输出层加入一个 Softmax 激活函数。而 LSTM 是改进版的 RNN,在 RNN 基础上增加了输入门、遗忘门、输出门,由于它有 3 个门,结构就稍稍复杂了一些,标准 LSTM 细胞结构如图 4 所示。

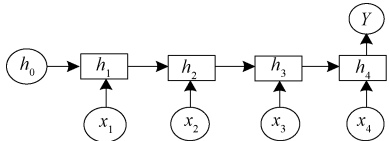


图 3 RNN 的多输入单输出模型  
Fig. 3 Multi input single output model of RNN

如图 4 所示,LSTM 的流程是为:每一个时刻都有一

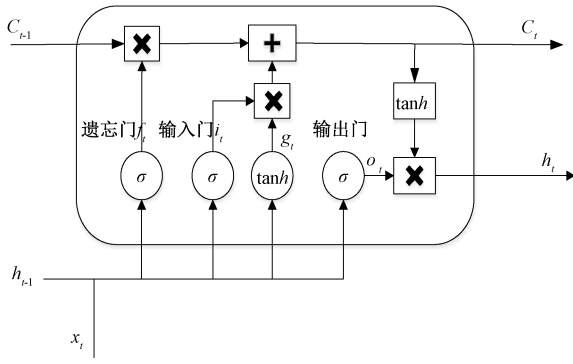


图 4 标准 LSTM 细胞结构

Fig. 4 Structure diagram of standard LSTM cells

个输出和细胞状态,输出就是上面的  $h$ ,状态用  $C$  来表示,那么,对于下一个时刻,把上一个时刻的输出连带着这一个时刻的输入作为整体,当作细胞的输入,同时也是控制 3 个门的输入。其各个门之间的表达式如式(3)~(8)所示。

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (4)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

$$g_t = \tanh(W_g x_t + U_g h_{t-1} + b_g) \quad (6)$$

$$C_t = f_t C_{t-1} + i_t g_t \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

式中: $f_t$  为遗忘门的激活值; $i_t$  为输入门的激活值; $o_t$  为输出门的激活值; $h_t$  和  $h_{t-1}$  为时间步  $t$  和时间步  $t-1$  时记忆单元的输出; $C_t$  和  $C_{t-1}$  分别为时间步  $t$  和时间步  $t-1$  记忆单元的状态; $W$  值表示输入数据的权重矩阵; $U$  值表示上一时刻 LSTM 单元输出的权重矩阵; $b$  为状态偏移量<sup>[21]</sup>。

### 3.2 麻雀搜索优化算法

LSTM 是具有记忆长短期信息的能力的神经网络,传统的训练方式采用梯度下降。然而,梯度下降方法容易陷入局部最优,群智能优化算法具有较强的全局搜索能力。本文通过对比分析采用麻雀搜索算法直接优化 LSTM 模型。麻雀搜索优化算法简单,易于实现和操作,同时具有较强的全局搜寻能力,能够快速找到全局最优解,有效地处理高维优化问题,还可以自适应调整搜索空间,优化 LSTM 的隐藏单元数目、最大训练周期、初始学习率和  $L2$  参数,从而极大的提高 LSTM 求解精度。

麻雀通常是群居的鸟类,它的能力主要是捕食和繁殖下一代,这两种行为类似于捕食者和繁殖者。麻雀算法建立模型的规则参考文献[22]。

麻雀的位置矩阵表示为:

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & \cdots & m_{1,j} \\ m_{2,1} & m_{2,2} & \cdots & \cdots & m_{2,j} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{i,1} & m_{i,2} & \cdots & \cdots & m_{i,j} \end{bmatrix} \quad (9)$$

其中, $i$  是麻雀的数量, $j$  是待优化变量的维数。式(9)中每一行的值代表个体的适应值,随着捕食者数量的增加,其位置不断更新,在每次迭代中,捕食者位置的更新表达式如式(10)所示。

$$F_M = \begin{bmatrix} f([m_{1,1} & m_{1,2} & \cdots & \cdots & m_{1,j}]) \\ f([m_{2,1} & m_{2,2} & \cdots & \cdots & m_{2,j}]) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f([m_{i,1} & m_{i,2} & \cdots & \cdots & m_{i,j}]) \end{bmatrix} \quad (10)$$

$$M_{a,b}^{t+1} = \begin{cases} M_{a,b}^t \cdot \exp(\frac{-a}{\rho \cdot n}), P_2 < T \\ M_{a,b}^t + Q \cdot L, P_2 \geq T \end{cases} \quad (11)$$

其中, $t$  表示当前迭代位置, $b=1,2,\dots,j$ 。 $M_{a,b}^t$  表示在迭代位置处于  $t$  时第  $a$  个麻雀所在第  $b$  维。 $\rho$  是一个范围在  $0 \sim 1$  的随机数, $P_2$  表示警报值,它的范围是  $[0, 1]$ , $T$  是表示安全阈值, $n$  为当前迭代次数, $Q$  是随机数, $L$  表示  $1 \times j$  的矩阵。当时, $P_2 < T$  表示麻雀群周边没有威胁者,捕食者可以进入大范围的搜索模式,当  $P_2 \geq T$  时,表示周边有威胁者,所有麻雀都需要迅速飞到其他安全区域去。搜寻者位置更新的表达式如式(12)所示。

$$M_{a,b}^{t+1} = \begin{cases} Q \cdot \exp(\frac{M_w^t - M_{a,b}^t}{a^2}), a > \frac{i}{2} \\ M_q^{t+1} + |M_{a,b}^t - M_q^{t+1}| \cdot A^+ \cdot L, a \leq \frac{i}{2} \end{cases} \quad (12)$$

其中, $M_w^t$  是表示该群麻雀中最不理想的位置, $M_q^t$  是捕食者占据的最佳位置, $A$  为一个  $1 \times j$  的矩阵,动态分配为  $-1$  或  $1$ , $A^+ = A^T(AA^T)^{-1}$ 。当  $a > i/2$  时,表明适应性较差的第  $a$  个搜寻者最有可能被饿死。在进行训练时,把整个种群中  $10\% \sim 20\%$  的麻雀定义为警惕性高的麻雀,并且自动生成它们的初始位置,数据如式(13)所示。

$$M_{a,b}^{t+1} = \begin{cases} M_{best}^t + \beta \cdot |M_{a,b}^t - M_{best}^t|, & f_a > f_g \\ M_{a,b}^t + k \cdot (\frac{|M_{a,b}^t - M_w^t|}{(f_a - f_w) + \varepsilon}), & f_a = f_g \end{cases} \quad (13)$$

其中, $M_{best}^t$  是群体的中心位置,代表着最佳值。 $\beta$  是一个参数,表示控制步长的大小。 $k$  的值是任意得到的,其范围在  $[-1, 1]$ 。 $f_a$  是当前麻雀的适应度值。 $f_g$  和  $f_w$  是当前麻雀的最佳适应度值和最差适应度值。在式(13)中,当  $f_a = f_g$  时,如果不加  $\varepsilon$  值,会使得整个分母为  $0$ ,因此,为了防止这种情况的出现,在分母中加一个非常小的值, $f_a > f_g$  表示目前这只麻雀离群体的位置有点远,基本上处于群体的边缘外部。 $f_a = f_g$  表示位置中心的

麻雀意识到有危险,要向其他地方靠近。

### 3.3 SSA-LSTM 模型构建

#### 1) LSTM 网络数据恢复原理

风电场监控系统的 SCADA 数据特点是时序性强,上一个数据与下一个数据之间在时间上具有关联性,数据不会发生突然的跳变。而 LSTM 多输入单输出网络可以把多个参数前面一定时间之间的数据、现在的状态和当前的输入信息等融合在一起,实现对下一个数据进行预测,由于 LSTM 网络在处理数据上具有强关联时间和空间序列信息,因此 LSTM 网络可以使实现状态异常数据的恢复。

当检测出某个传感器某时段数据出现异常时,利用 LSTM 网络多输入单输出实现传感器状态异常输出数据恢复。它的基本思想是在给定的采样周期内,采用不同的方式获取传感器相应的相关参数。比如关联参数选取 3 个,取参数前正常  $m$  个数据,来预测异常传感器的第  $m+1$  个数。第 1 个参数标记为  $x'_1, x'_2, \dots, x'_m$ , 第 2 个参数标记为  $x''_1, x''_2, \dots, x''_m$ , 第三个参数标记为  $x'''_1, x'''_2, \dots, x'''_m$ , 构成第一个  $3 \times m$  的细胞元,一起作为 LSTM 网络的第一个输入值。然后数据往后推移一位,第 1 个参数标记为  $x'_2, x'_3, \dots, x'_{m+1}$ , 第 2 个参数标记为  $x''_2, x''_3, \dots, x''_{m+1}$ , 第三个参数标记为  $x'''_2, x'''_3, \dots, x'''_{m+1}$ , 构成第二个  $3 \times m$  的细胞元,一起作为 LSTM 网络的第 2 个输入值,以此类推,构成一个多数列的时间序列集。建立 LSTM 多输入单输出预测模型,利用构造的时间序列集进行模型训练。同理,建立异常状态数据的多参数时间序列集,利用训练好的 LSTM 模型进行预测,经过 LSTM 网络多次预测之后,可以得到预测数据集  $y_1, y_2, \dots, y_m$ , 其异常状态数据用对应时段内的预测数据集来替代,实现异常状态数据的恢复。

#### 2) 模型输入参数确定

SCADA 数据中,参数众多,不是每一个参数都与定子绕组温度有关系,因此为了提高数据恢复精度,该模型的输入参数选定通过专家系统判定和 Pearson 相关系数法计算得到。经过专家系统判断,从 SCADA 数据中 150 个参数选出 42 个不同类型的状态参数,然后利用 Pearson 相关系数法计算,从 42 个参数中再进一步确定输入参数, Pearson 相关系数法计算公式如下:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \cdot \sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \quad (14)$$

式中:  $r$  表示 Pearson 相关系数;  $n$  表示条件参数的长度;

$x_i, y_i$  分别表示不同参数  $i_{th}$  的采样点;  $\bar{x}, \bar{y}$  表示条件参数的平均值。  $x$  为风电机组定子绕组温度 6 测量值,  $y$  为选出的 42 个不同类型的状态参数, 由于风电机组运行的时候, 各个部件之间存在着关联特性, 状态参数会相互影响。经过反复试验, 选取 Pearson 相关性评判标准取值为 0.75, 得到模型的输入参数为轮毂转速、机舱气象站风速、变频器网侧有功功率、发电机侧功率、发电机运行频率、发电机电流、发电机定子温度 2、主轴承温度 1 和机舱温度, 共计 9 个参数<sup>[23]</sup>。

#### 3) SSA-LSTM 模型建立

SSA 优化 LSTM 网络的流程具体步骤如下:

(1) 参数初始化。对 SSA 参数初始化设置, 根据多次实验选取迭代次数为 30 次, 种群规模为 50 只, 捕食者的比例为 20%, 安全阈值 T1 取 0.8, 意识到危险的麻雀数量为种群的 10%~20%, 这里取 10 只。P2 为警报值, 它的范围是  $[0, 1]$  的随机值。

(2) 参数赋值。空间维度为 4, 最大迭代次数为 200, 搜索下边界设置为  $[2, 2, 10^{-4}, 10^{-5}]$ , 上边界设置为  $[200, 100, 1, 1]$ 。

(3) 优化参数设置。随机初始化麻雀种群的位置。

(4) 计算适应度值。根据目标函数计算麻雀的适应度值, 更新捕食者位置, 更新加入者位置, 更新警戒者位置, 计算适应度并更新麻雀位置, 并保存最佳位置个体。

(5) 优化结束。根据定义的最大迭代次数, 判断是否满足停止条件, 满足则退出, 输出最佳参数向量, 否则, 返回执行步骤(2)重复上述步骤直到得到最优参数。

运行 SSA 算法, 得到最佳适应度值, 把最佳适应度值 Best\_pos 中的值赋予 LSTM 的隐藏单元数目 (num hidden units)、最大训练周期 (max epochs)、初始学习率 (initial learn rate) 和  $L_2$  参数 ( $L_2$  regularization), 其表达式如下。

$$\text{numHiddenUnits} = \text{Best\_pos}(1) \quad (15)$$

$$\text{maxEpochs} = \text{Best\_pos}(2) \quad (16)$$

$$\text{InitialLearnRate} = \text{Best\_pos}(3) \quad (17)$$

$$\text{L2Regularization} = \text{Best\_pos}(4) \quad (18)$$

基于 SSA-LSTM 网络的异常状态输出数据恢复模型建立如下:

(1) 数据选择, 选取前面分析的风速仪和定子绕组 6 号温度传感器的相关参数正常历史数据。

(2) 数据处理, 建立数据序列, 确定训练输入样本和输出样本向量, 以及测试输入样本和输出样本向量。样本数据中的  $m$  个值, 通过多次试验, 通过误差最终来判定。对选取的数据进行归一化处理, 数据范围为  $[0, 1]$ 。

(3) 模型拓扑结构, 设置麻雀的种群规模、迭代次数、空间维度、边界范围, 设置好 LSTM 参数的输入输出维度等。



(4) 参数优化,通过 SSA 算法寻找最优值,并把最优值赋给 LSTM 网络的隐藏单元数目、最大训练周期、初始学习率和  $L_2$  参数,得到 LSTM 最优初始向量。

(5) 训练模型,把前期建立的训练数据序列输入模型进行训练,再用建立的测试数据序列进行模型测试,来验证建模的准确性。

(6) 状态异常输出数据恢复,根据前面的分析,选取状态异常的定子绕组 6 号传感器温度数据,应用训练好的模型进行数据恢复。

## 4 SSA-LSTM 模型训练和异常输出数据恢复流程

### 4.1 模型训练与测试

当 SSA-LSTM 模型建立好之后,对模型进行训练和测试,具体过程如下:

1) 选取研究传感器的相关参数历史正常数据 20 000,其中随机的 80% 作为训练数据,每一个相关参数时间序列记为  $(a_1, a_2, \dots, a_{16\,000})$ ,所研传感器的时间序列记为  $(b_1, b_2, \dots, b_{16\,000})$ ,剩余 20% 作为测试数据,所研传感器的时间序列记为  $(y_1, y_2, \dots, y_{4\,000})$ 。

2) 训练数据时间序列的构造。假设研究传感器相关参数的个数为  $n$ ,采用相关参数前面的  $m$  个数据预测所研传感器第  $m+1$  个数据,训练数据输入模型时间序列样本由  $16\,000-m$  个元胞组成,记为  $(X_1, X_2, \dots, X_{16\,000-m})$ ,每一个元胞都是一个  $n \times m$  的矩阵。其中  $X_1$  的表达式如 (19) 所示,  $X_2$  的表达式如 (20) 所示,依次类推,  $X_{16\,000-m}$  的表达式如 (21) 所示。输出模型时间序列样本向量记为  $(Y_1, Y_2, \dots, Y_{16\,000-m})$ ,其中  $Y_1$  为  $b_{m+1}$ ,  $Y_2$  为  $b_{m+2}$ ,同理得到  $Y_{16\,000-m}$  为  $b_{16\,000}$ 。

$$X_1 = \begin{bmatrix} a_1^1 & a_1^2 & \dots & a_1^m \\ a_2^1 & a_2^2 & \dots & a_2^m \\ \vdots & \vdots & \ddots & \vdots \\ a_n^1 & a_n^2 & \dots & a_n^m \end{bmatrix} \quad (19)$$

$$X_2 = \begin{bmatrix} a_2^1 & a_3^1 & \dots & a_{m+1}^1 \\ a_2^2 & a_3^2 & \dots & a_{m+1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_2^n & a_3^n & \dots & a_{m+1}^n \end{bmatrix} \quad (20)$$

$$X_{16\,000-m} = \begin{bmatrix} a_{16\,000-m}^1 & a_{16\,000-m+1}^1 & \dots & a_{15\,999}^1 \\ a_{16\,000-m}^2 & a_{16\,000-m+1}^2 & \dots & a_{15\,999}^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{16\,000-m}^n & a_{16\,000-m+1}^n & \dots & a_{15\,999}^n \end{bmatrix} \quad (21)$$

3) 测试数据时间序列的构造。类似训练数据时间序列,测试数据输入模型时间序列样本由  $4\,000-m$  个元胞

组成,记为  $(Z_1, Z_2, \dots, Z_{4\,000-m})$ 。其中  $Z_1, Z_2, \dots, Z_{4\,000-m}$  的矩阵表达式类似于式 (19)。输出模型时间序列样本向量记为  $(C_1, C_2, \dots, C_{4\,000-m})$ ,其中  $C_1$  为  $y_{m+1}$ ,  $C_2$  为  $y_{m+2}$ ,同理得到  $C_{4\,000-m}$  为  $y_{4\,000}$ 。

4) 时间序列归一化处理。对训练和测试数据按照式 (22) 进行归一化处理。

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (22)$$

5) SSA-LSTM 模型设置。根据前面的介绍,设置好 SSA-LSTM 模型参数,误差用均方根误差 (root mean square error, RMSE) 表示。

6) 模型训练。选择不同的  $m$  值,由于选取的 SCADA 数据是 10 min 数据,取  $m$  范围为  $[4, 18]$ ,表示用相关参数前面 0.4~3 h 的数据来预测所研传感器的  $m+1$  个数。对 SSA-LSTM 模型进行训练,训练完成之后,保存模型。

7) 模型测试。根据训练好的模型,利用测试数据进行测试模型性能,计算每次不同  $m$  得到的测试数据预测值与实际值的 RMSE 值,选取 RMSE 最小时对应的  $m$  和相应的模型。

### 4.2 状态异常输出数据恢复流程

根据对 SSA-LSTM 模型的测试,选取合适的  $m$  值,取定子绕组 6 号温度传感器状态异常输出数据,利用 SSA-LSTM 模型对异常输出数据进行恢复,具体过程如下:

1) 选取定子绕组 6 号温度传感器的状态异常输出数据和 9 个相关参数正常数据,把相关参数正常数据标记为  $(x_1, x_2, \dots, x_k)$ ,  $k$  为数据长度。

2) 状态异常数据时间序列构造,输入模型时间序列样本由  $k-m$  个元胞组成,记为  $(X_1, X_2, \dots, X_{k-m})$ 。其中  $X_1, X_2, \dots, X_{k-m}$  的矩阵表达式类似于式 (19)~(21)。

3) 输入数据时间序列按照式 (22) 进行归一化处理。

4) 利用训练好的 SSA-LSTM 模型,对输入状态异常数据  $(X_1, X_2, \dots, X_{k-m})$  进行预测,模型预测结果记为  $(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{k-m})$ 。

5) 把预测结果作为状态异常恢复数据并保存,作为系统的输入数据。

## 5 基于 SSA-LSTM 的定子绕组温度传感器异常状态输出数据恢复

### 5.1 定子绕组温度传感器模拟异常状态输出数据恢复

通过设置不同的  $m$  值,得到每一次测试数据的 RMSE 值,根据 RMSE 的值,取最小值时对应的  $m$  构成时间序列。输入参数选取处理过的 SCADA 数据 40 000 组,其中训练数据选取 4 000 组,测试数据选取 1 000 组,按照式 (19)~(21) 设置训练和测试输入输出参数数据

矩阵。对 SSA-LSTM 模型进行训练测试,当取不同  $m$  时对应的 RMSE 值如表 3 所示。从表 3 可以看出,当  $m$  取 11 时,对应的 RMSE 值最小,因此用前 11 个数据预测第 12 个数据得到的数据恢复精度更高。采用相似度系数  $R$  来衡量数据恢复精度, $R$  的表达式如式(23)所示。

$$R = \frac{\sum_i \sum_n (A_i - \bar{A})(B_n - \bar{B})}{\sqrt{(\sum_i \sum_n (A_i - \bar{A}))^2 (\sum_i \sum_n (B_n - \bar{B}))^2}} \quad (23)$$

其中, $A$  表示真实值, $B$  表示预测值, $i$  和  $n$  分别表示数值在  $A$ 、 $B$  数列中所处的位置,  $\bar{A} = \text{mean2}(A)$ ,  $\bar{B} = \text{mean2}(B)$ 。  $R$  值越接近 1,表明二列数据的相似度越高,恢复精度也就越高。

为了验证 SSA-LSTM 模型数据恢复的精度,构造一段定子绕组温度传感器状态异常数据。取 2013 年 19 号风电机组的定子绕组 6 温度传感器 SCADA 数据作为输出参数,测试数据范围为 28 001~32 000,总计 4 000 组数据,其中在第 1 001~1 500 个数加偏置量为 15 ℃,构造模拟故障数据段。按照前面的设置,取  $m$  值为 11,构造训练数据时间序列样本,训练数据选取 8 000 组。表示 9 个输入参数 1~11 个数预测定子绕组 6 温度传感器的第 12 个数据,输入参数 2~12 个数预测定子绕组 6 温度传感器的第 13 个数据,依此类推,输入参数 7 988~7 998 个数预测定子绕组 6 温度传感器的第 7 999 个数据。构造一个输入为 7 988×1 的元胞输入数组,每一个元胞由一个 9×11 的矩阵组成。

表 3 不同  $m$  对应的 RMSE 值

Table 3 RMSE values corresponding to different $m$					
$m$ 值	RMSE	$m$ 值	RMSE	$m$ 值	RMSE
4	1.085 7	5	1.558 7	6	1.165 1
7	1.115 4	8	1.887 7	9	1.682 5
10	1.810 5	11	1.061 4	12	1.062 5
13	1.914 7	14	1.131 5	15	1.930 2
16	1.524 5	17	1.121 5	18	1.587 5

对 SSA-LSTM 模型进行训练,用训练好的模型对模拟异常数据进行恢复,其数据恢复结果如图 5 所示,实际值与恢复值之间的差值曲线如图 6 所示,可以看出对于非故障部分,恢复值曲线与实际值曲线基本重合,而对于 500 个故障数据段,模型在真实数据上的精度与恢复后数据上的精度也非常接近,温度重构误差控制在±3 ℃以内,平均相对误差不超过 1.86%,完全满足系统的要求。通过式(23)计算出 SSA-LSTM 模型异常数据段恢复值与实际值的  $R$  值,同时对比了未优化 LSTM 模型、DBN 模型和 WOA-LSTM 模型的  $R$  值,其结果如表 4 所示。可以看出 SSA-LSTM 模型的数据恢复精度是最高的,其恢复精

度比未优化模型高近 4.2%,比 FISSA-DBN 模型高 2.2%,而且恢复精度接近 99.7%,表明该方法是完全可行的,可以用来对实际异常数据进行恢复。

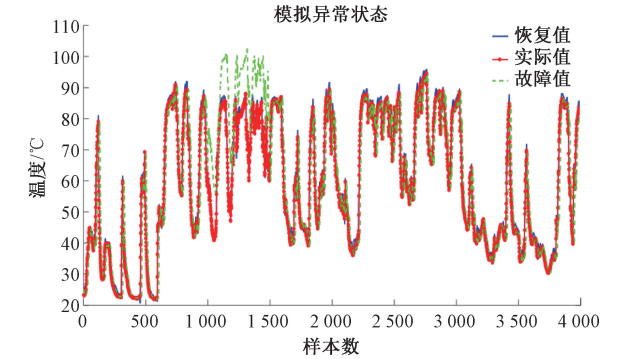


图 5 模拟状态异常数据恢复图

Fig. 5 Reconstruction of abnormal data in simulated state

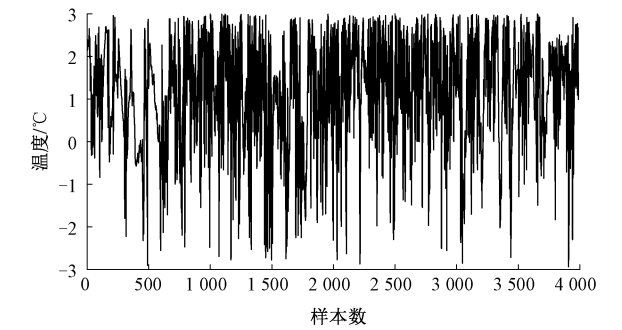


图 6 模拟异常状态实际值与恢复值的差值曲线

Fig. 6 Curve of the difference between the actual value and the recovered value of the simulated abnormal state

表 4 4 种模型计算  $R$  值对比表

Table 4 Comparison of $R$ values calculated by four models				
模型	LSTM	DBN	WOA-LSTM	SSA-LSTM
$R$ 值	0.955 2	0.957 8	0.974 2	0.997 1

5.2 定子绕组温度传感器实际异常输出数据恢复

利用 SSA-LSTM 模型对前面 2016 年 7 号风电机组的定子绕组 6 温度传感器进行异常数据恢复,数据段数据 15 818~35 818。 $m$  取值为 11,即用 9 个输入参数的 15 818~15 828 预测定子绕组 6 温度传感器的第 15 829 个数,输入参数的 15 819~15 829 预测定子绕组 6 温度传感器的第 15 830 个数,依此类推,一直到输入参数的第 35 807~35 817 预测 35 818 个数,构造一个 19 900×1 的元胞输入数组,每一个元胞由 9×11 的矩阵组成。经过训练好的 SSA-LSTM 模型数据恢复后,其结果如图 7 所示,明显可以看出,在未发生异常时段,恢复值曲线与异常值曲线非常接近,在异常数据段 19 808~30 487,恢复值与



异常值相差较大,其恢复值与异常值的差值如图 8 所示,表明该数据的故障非常明显,比正常数据大的多。

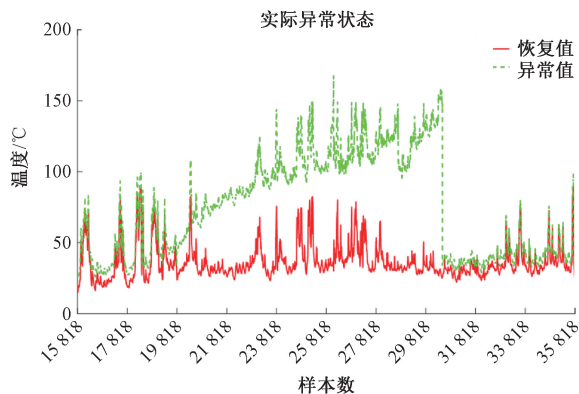


图 7 实际异常数据恢复曲线

Fig. 7 Actual abnormal data recovery curve graph

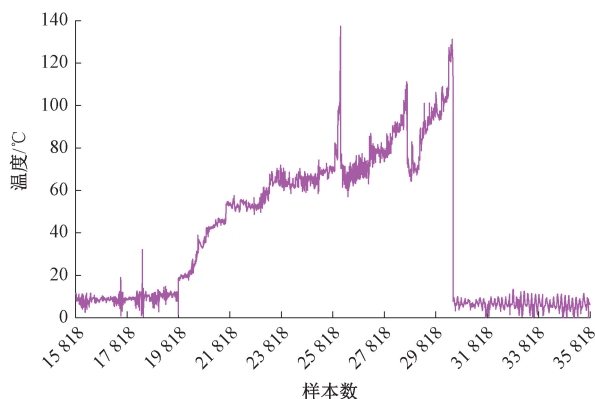


图 8 异常数据恢复值与实际异常数据的差值曲线

Fig. 8 The difference curve between the recovered value of abnormal data and the actual abnormal data

## 6 确认不确定度计算

测量不确定度评定是传感器自确认技术准确性的评定指标,反映了自确认技术对传感器故障检测的正确率。在风电机组传感器状态自确认中,传感器数据的不确定度不仅受传感器自身性能影响,还受到环境噪声、系统动态变化等因素的干扰。传统的静态不确定度评估方法(如蒙特卡洛模拟)通常假设噪声分布固定,难以适应动态环境,因此本文采用动态不确定度评估方法(贝叶斯算法)。

### 6.1 贝叶斯算法的基本原理

贝叶斯算法是一种基于概率的推理方法,通过先验分布和观测数据的结合,动态更新参数的后验分布。相比传统方法,贝叶斯方法具备动态更新能力,更适合风电运行中的非稳定环境。其核心公式为:

$$P(\theta | D) = \frac{P(D | \theta) \cdot P(\theta)}{P(D)} \quad (24)$$

式中:  $P(\theta)$  是参数的先验分布(基于历史数据或专家知识);  $P(D | \theta)$  是似然函数(描述观测数据与参数的关系);  $P(\theta | D)$  是后验分布(更新后的参数分布);  $P(D)$  是归一化常数。在动态不确定度评估中,贝叶斯算法通过不断更新后验分布,实现对不确定度的动态量化。

置信区间的计算公式为:

$$\mu \pm k^* \frac{\sigma}{n} \quad (25)$$

式中:  $\mu$  为恢复值;  $k$  为置信因子;  $\sigma$  为恢复数据的标准差;  $n$  为样本大小。根据标准正态分布,本文选取 95% 置信水平,此时置信因子  $k = 1.96$ 。

### 6.2 贝叶斯算法的步骤及计算

#### 1) 定义不确定度模型

假设温度传感器测量值  $y_i$  服从如下模型:

$$y_i = f(x_i, \theta) + \varepsilon_i \quad (26)$$

式中:  $x_i$  是输入变量温度;  $\theta$  是模型参数(如传感器增益、偏移量等);  $\varepsilon_i$  是噪声项,通常假设服从高斯分布  $\varepsilon_i \sim N(0, \sigma_i^2)$ 。不确定度  $\sigma_i^2$  是动态变化的,需要通过贝叶斯算法实时更新。

#### 2) 初始化先验分布

假设模型参数  $\theta$  和噪声方差  $\sigma_i^2$  的先验分布:

$$P(\theta) \sim N(\mu_0, \Sigma_0) \quad (27)$$

$$P(\sigma_i^2) \sim \text{Inverse-Gamma}(\alpha_0, \beta_0) \quad (28)$$

式中:  $\mu_0, \Sigma_0, \alpha_0, \beta_0$  基于历史数据或专家知识设定。

#### 3) 构建似然函数

假设观测数据  $D_i = \{y_1, y_2, \dots, y_i\}$  独立同分布,似然函数为:

$$P(D_i | \theta, \sigma_i^2) = \prod_{i=1}^i N(y_i | f(x_i, \theta), \sigma_i^2) \quad (29)$$

#### 4) 更新后验分布

利用贝叶斯公式,更新参数的后验分布:

$$P(\theta, \sigma_i^2 | D_i) \propto P(D_i | \theta, \sigma_i^2) \cdot P(\theta) \cdot P(\sigma_i^2) \quad (30)$$

由于解析解难以计算,通常采用马尔可夫链蒙特卡洛(MCMC)方法近似求解。

#### 5) 动态更新不确定度

在每个时间步  $t$ , 基于后验分布  $P(\theta, \sigma_i^2 | D_i)$  计算不确定度:

$$\sigma_i^2 = E[\sigma_i^2 | D_i] \quad (31)$$

式中:  $E[\cdot]$  表示期望值。

#### 6) 预测与校正

预测下一时刻的测量值及其不确定度:

$$y_{i+1} \sim N(f(x_{i+1}, \theta), \sigma_i^2) \quad (32)$$

当新观测数据  $y_{i+1}$  到达时,重复步骤 4) ~ 5), 动态更

新不确定度。

本文采样次数为 19 900, 噪声方差的先验形状参数取 2, 噪声方差的先验尺度参数取  $1^{[24]}$ 。为量化恢复数据的不确定度, 本文采用贝叶斯估计方法, 并通过计算恢复值的 95% 置信区间来进一步确认恢复数据的可靠性。

根据前面的分析, 对恢复数据进行不确定度计算得出平均不确定度为 1.252, 其不确定度曲线如图 9 所示, 对实际异常数据进行确认不确定度计算, 平均不确定度为 32.95, 恢复数据的平均不确定度仅为异常数据的 1/26, 恢复数据可信度较高, 其不确定度曲线如图 10 所示, 通过贝叶斯算法的动态不确定度评估, 能够实时更新传感器数据的可信度。该方法在风电机组传感器状态自确认中具有重要应用价值, 可显著提升异常检测与数据恢复的可靠性。

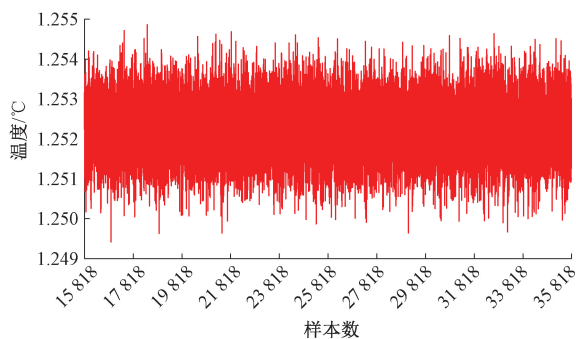


图 9 恢复数据的不确定度曲线

Fig. 9 Uncertainty curve for data recovery

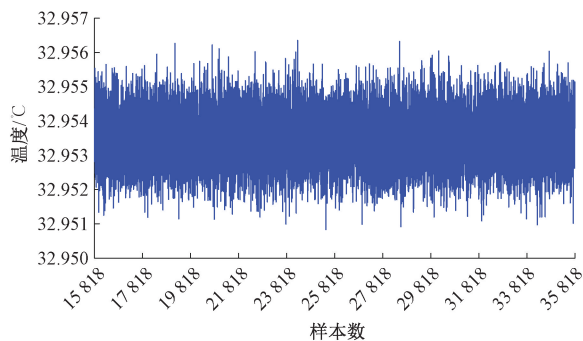


图 10 异常数据的不确定度曲线

Fig. 10 Uncertainty curve of abnormal data

## 7 结 论

本文针对风电机组定子绕组温度传感器的状态自确认问题, 提出了一种基于信息融合技术与深度学习的研究方法, 利用传感器之间的相关性和关联性, 计算传感器之间的灰色关联度, 实现了对温度传感器异常状态的精准识别; 通过构建基于 LSTM 神经网络的时序数据恢复模型, 利用 SSA 对模型超参数进行优化。实验结果表明,

在传感器发生偏置、间歇性失效等典型故障场景下, 其恢复精度接近 99.7%, 比未优化模型高近 4.2%, 比 FISSA-DBN 模型高 2.2%, 温度重构误差控制在  $\pm 3$  °C 以内, 平均相对误差不超过 1.86%, 完全满足系统的要求; 采用贝叶斯算法对恢复数据进行不确定度计算, 平均不确定度为 1.252, 反映了自确认技术对传感器故障检测的正确率高, 提升异常检测与数据恢复的可靠性。研究结果表明, 将信息融合技术与深度学习相结合, 不仅能有效解决传感器故障检测与数据恢复的技术难题, 其引入的不确定度量化框架更为智能诊断系统的可靠性评估提供了新思路, 为智慧风场建设提供技术支持。

## 参考文献

- [1] YANG W, TAVNER P J, CRABTREE C J, et al. Wind turbine condition monitoring: Technical and commercial challenges[J]. Wind Energy, 2014, 17(5): 673-693.
- [2] 李辉, 赵猛, 赵斌, 等. 双馈风电机组关键传感器的故障诊断方法[J]. 中国电机工程学报, 2011, 31(6): 73-78.  
LI H, ZHAO M, ZHAO B, et al. Fault diagnosis method for key sensors of doubly fed wind turbines [J] Chinese Journal of Electrical Engineering, 2011, 31(6): 73-78.
- [3] 王梓齐, 张书瑶, 刘长良. 基于增量式相对熵的风电机组实时状态监测[J]. 电子测量与仪器学报, 2020, 34(12): 125-132.  
WANG Z Q, ZHANG SH Y, LIU CH L. Real-time condition monitoring of wind turbine based on incremental relative entropy[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(12): 125-132.
- [4] HENRY M P, CLARKE D W. The self-validating sensor: Rationale, definitions and examples[J]. Control Engineering Practice, 1993, 1(4): 585-610.
- [5] 蒋栋年, 褚天锐, 高玉鑫. 基于冗余关系分析的传感器自诊断设计方法研究[J]. 仪器仪表学报, 2024, 45(3): 72-83.  
JIANG D N, CHU T R, GAO Y X. Research on sensor self-diagnosis design method based on redundancy relationship analysis [J]. Chinese Journal of Scientific Instrument, 2024, 45(3): 72-83.
- [6] XU P, SONG K, CHEN Y S, et al. Fault diagnosis method of self-validating metal oxide semiconductor gas sensor based on t-distribution stochastic neighbor embedding and random forest [J]. Review of Scientific Instruments, 2019, 90(5): 055002.
- [7] WEN J J, YAO H P, JI Z, et al. Self-validating high-g accelerometers through data-driven methods [J]. Sensors and Actuators A: Physical, 2021, 328: 112803.
- [8] SHAHNAI R, ZHAO Q. Adaptive fuzzy descriptor

- sliding mode observer-based sensor fault estimation for uncertain nonlinear systems [J]. *Asian Journal of Control*, 2016, 18 (4): 1478-1488.
- [9] 田宇, 高强, 李栋, 等. 测距传感器性能评估与故障诊断方法研究[J]. *电子测量与仪器学报*, 2020, 34(10): 172-178.
- TIAN Y, GAO Q, LI D, et al. Research on Performance Evaluation and Fault Diagnosis Methods of Distance Measuring Sensors [J]. *Journal of Electronic Measurement and Instrumentation*, 2020, 34 (10): 172-178.
- [10] 马永健, 冯云, 孙培伟, 等. 基于 PCA-SVM 的小型自然循环铅冷快堆传感器故障诊断方法研究[J]. *核科学与工程*, 2024, 44 (2): 464-471.
- MA Y J, FENG Y, SUN P W, et al. Research on Sensor Fault Diagnosis Method of a Small Natural Circulation Lead-cooled Fast Reactor Based on PCA-SVM [J]. *Nuclear Science and Engineering*, 2024, 44 (2): 464-471.
- [11] 李洁, 黄文新, 蔡一鸣, 等. 基于 DFPMM 的位置传感器故障诊断与容错控制[J]. *航空学报*, 2024, 45 (10): 265-276.
- LI J, HUANG W X, CAI Y M, et al. Fault diagnosis and fault tolerant control of position sensor based on DFPMM[J]. *Acta Aeronautica et Astronautica Sinica*, 2024, 45(10): 265-276.
- [12] 王家能, 周治平. 正则化 ISU-LSTM 的多传感器故障预测方法[J]. *传感技术学报*, 2021, 34 (2): 196-202.
- WANG J N, ZHOU ZH P. Multi sensor fault prediction method using regularized ISU-LSTM [J]. *Journal of Sensing Technology*, 2021, 34 (2): 196-202.
- [13] 周凌, 赵前程, 石照耀, 等. 风电场风电机组机载风速仪状态自确认[J]. *太阳能学报*, 2022, 43(11): 172-178.
- ZHOU L, ZHAO Q CH, SHI ZH Y, et al. Self-confirmation of onboard anemometer status of wind turbines in wind farms[J]. *Acta Energiæ Solaris Sinica*, 2022, 43(11): 172-178.
- [14] YANG X L, ZHAO Q F, WANG Y Z, et al. Fault signal reconstruction for multi-sensors in gas turbine control systems based on prior knowledge from time series representation[J]. *Energy*, 2023, 262: 124996.
- [15] LINDA H, MOHAMED G, ABEDLKIRIM N. Fault diagnosis and reconfiguration for mobile robot localization based on multi-sensors data fusion [J]. *Unmanned Systems*, 2022, 10(1): 69-91.
- [16] YUE Y, PENG M J, WANG H, et al. Improved PCA model for multiple fault detection, isolation and reconstruction of sensors in nuclear power plant [J]. *Annals of Nuclear Energy*, 2020, 148: 107662.
- [17] 林涛, 张达, 王建君. 改进 LSTM-RF 算法的传感器故障诊断与数据重构研究[J]. *计算机工程与科学*, 2021, 43(5): 845-852.
- LIN T, ZHANG D, WANG J J. Sensor fault diagnosis and data reconstruction based on improved LSTM-RF algorithm[J]. *Computer Engineering & Science*, 2021, 43(5): 845-852.
- [18] ISERMANN R. Bayesian methods for fault detection and diagnosis in sensor networks [J]. *Control Engineering Practice*, 2006, 14(6): 575-576.
- [19] 谢民, 邵庆祝, 王同文, 等. 基于 CPSO-PNN 及灰色关联度分析的电网分区故障诊断[J]. *安徽大学学报 (自然科学版)*, 2022, 46 (3): 94-101.
- XIE M, SHAO Q ZH, WANG T W, et al. Divisional fault diagnosis of power grid based on CPSO-PNN and gray relational analysis [J]. *Journal of Anhui University (Natural Science Edition)*, 2022, 46 (3): 94-101.
- [20] 刘晓悦, 魏宇册. 基于改进灰色关联分析的 BA-BP 短期负荷预测[J]. *科学技术与工程*, 2020, 20(1): 223-227.
- LIU X Y, WEI Y C. Short-term load forecasting of BA-BP neural network based on improved grey relational analysis [J]. *Science Technology and Engineering*, 2020, 20 (1): 223-227.
- [21] 李冬辉, 尹海燕, 郑博文, 等. 改进的 LSTM 方法在冷水机组传感器故障检测中的应用[J]. *电工技术学报*, 2019, 34(11): 2324-2332.
- LI D H, YIN H Y, ZHENG B W, et al. Application of Improved LSTM Method in Sensor Fault Detection of Chiller Unit [J]. *Journal of Electrical Technology*, 2019, 34 (11): 2324-2332.
- [22] XUE J, SHEN B. A novel swarm intelligence optimization approach: Sparrow search algorithm [J]. *Systems Science & Control Engineering*, 2020, 8 (1): 22-34.
- [23] 周凌, 赵前程, 朱岸锋, 等. 基于 FISSA-DBN 模型的风电机组运行状态监测[J]. *振动. 测试与诊断*, 2023, 43(1): 80-87, 199.
- ZHOU L, ZHAO Q CH, ZHU A F, et al. Wind Turbine Operation Status Monitoring Based on FISSA-DBN Model[J]. *Journal of Vibration, Measurement & Diagnosis*, 2023, 43(1): 80-87, 199.
- [24] VATHAVAYA S G, CHAKRABORTY S. Bayesian



uncertainty quantification for wind turbine health monitoring[J]. Wind Energy, 2018, 21(3): 876-891.

作者简介



周凌, 2004 年于湖南工程学院获得学士学位, 2007 年于湖南大学获得硕士学位, 2023 年于湖南科技大学获得博士学位, 现为湖南工业大学讲师, 主要研究方向为风电机组智能检测与故障诊断。  
E-mail: 10103@hut.edu.cn

**Zhou Ling** received his B. Sc. degree from Hunan Institute of Engineering in 2004, M. Sc. degree from Hunan University in 2007, and Ph. D. degree from Hunan University of Science and

Technology in 2023. Now he is a lecturer at Hunan University of Technology. His main research interests include intelligent detection and fault diagnosis of wind turbines.



黄倩(通信作者), 2024 年毕业于怀化学院, 现为湖南工业大学硕士研究生, 主要研究方向为风电机组智能检测与故障诊断。  
E-mail: 2815751907@qq.com

**Huang Qian** (Corresponding author) received the B. Sc. degree from Huaihua University in 2024. Now she is a M. Sc. candidate at Hunan University of Technology. Her main research interests include intelligent detection and fault diagnosis of wind turbines.