

# 风力机主轴承故障监测方法\*

郑玉巧, 魏剑峰, 朱凯, 董博

(兰州理工大学机电工程学院 兰州, 730050)

**摘要** 根据风力机主轴承发生故障时引起其温度异常波动的特点,提出一种基于温度模型的故障监测方法。首先,分别建立正常运行状态下主轴承温度的多元线性回归预测模型、灰色预测模型、支持向量机回归预测模型及其组合预测模型;其次,在最佳预测模型基础上引入滑动窗口方法,研究主轴承温度预测残差统计特性;最后,通过对比温度残差均值或标准差的置信区间与设定临界值判断主轴承是否发生故障。研究表明,主轴承温度组合预测模型预测效果最佳,其判定系数分别较多元线性回归模型、灰色预测模型及支持向量机回归模型提高 0.049 3, 0.002 7 和 0.000 2,通过滑动窗口方法所求解的温度预测残差统计特性可及时反映出主轴承运行状态,这对风力机主轴承故障状态的高效监测及制定科学健康的维护策略提供依据。

**关键词** 故障监测;主轴承;组合预测;滑动窗口法;判定系数

**中图分类号** TH17;TK83

## 引言

风力机主轴承作为风力机传动系统的核心部件,承受轮毂、叶片等部件带来的多种变载冲击作用,极易发生故障<sup>[1-2]</sup>。由于风力机主轴承在距离地面几十米甚至上百米高的机舱内,一旦发生故障维护难度大、成本高,若未及时发现并解决故障会给风力机运行带来严重影响甚至造成安全事故。因此,实现风力机主轴承的故障监测,及时发现故障征兆并开展维护,对于风力机健康运行、降低运维成本具有重要意义。

传统风力机部件故障监测方法有振动分析<sup>[3-5]</sup>、声发射监测<sup>[6-7]</sup>和电气信号监测<sup>[8-9]</sup>等。其中振动监测法和声发射监测法需要安装大量的传感器来采集信号,成本昂贵;此外安装的传感器一旦发生故障,不仅导致采集的数据可靠性降低,还增加额外的运维成本。相比风力机部件的振动信号和声信号,电气信号的采集不需要额外安装传感器<sup>[10]</sup>,但电气信号中所包含的故障信息往往比较微弱,信噪比低,监测结果误差较大。因此,上述方法难以在风力机部件实际监测中大范围推广。

目前,多数风电场通过安装数据采集与监控系统(supervisory control and data acquisition,简称 SCADA)采集储存风力机各部件运行参数及环境

参数,因此选择 SCADA 系统所记录的风力机主轴承运行参数开展其故障监测研究。主轴承故障表现形式为磨损、点蚀或塑性变形等,故障早期主轴承滚过表面损伤点产生的异常摩擦又加剧故障程度,引起主轴承温度特征的异常波动。鉴于此,本研究以 SCADA 系统所采集的风力机主轴承运行参数为研究对象,分别采用多元线性回归(multiple linear regression,简称 MLR)、灰色模型(grey model,简称 GM)及支持向量机回归(support vector machine regression,简称 SVR)建立主轴承温度单一预测模型及它们的组合预测模型,引入滑动窗口方法实时监测温度预测残差的均值和标准差变化情况,通过温度预测残差均值(或标准差)的置信区间与设定临界值的对比来判断主轴承运行状态,从而实现对风力机主轴承的故障监测。

## 1 主轴承温度组合预测模型的构建

### 1.1 问题描述与数据预处理

风力机 SCADA 系统通常以 10 min 为采样周期记录并传输风速、输出功率及主轴承温度等运行参数<sup>[11]</sup>,由于风力机 SCADA 系统输出的各项运行参数单位不同,若直接用于建模分析对预测结果造成较大误差,因此需要对原始数据进行归一化处理,消

\* 国家自然科学基金资助项目(51965034);兰州市人才创新创业资助项目(2018-RC-25)

收稿日期:2020-08-15;修回日期:2020-09-04

除数据间的量纲影响。公式如下

$$x^* = (x_i - x_{\min}) / (x_{\max} - x_{\min}) \quad (1)$$

其中: $x_{\max}$ 为监测数据中的最大值; $x_{\min}$ 为监测数据中的最小值; $x_i$ 为原始监测值; $x^*$ 为归一化处理后的数据, $x^*$ 的取值范围为 $[0, 1]$ 。

## 1.2 多元线性回归 (MLR) 模型

MLR模型阐述一个因变量依赖多个自变量的变化规律<sup>[12]</sup>。风力机主轴轴承温度的变化视为多个变量参数影响下的共同作用结果,因此可将主轴轴承温度作为因变量(即建模输出变量),影响主轴轴承温度变化的运行参数作为自变量(即建模输入变量),建立主轴轴承温度的多元线性回归预测模型步骤如下。

设因变量(即风力机主轴轴承温度)为 $y$ , $m$ 个自变量(即影响主轴轴承温度变化的变量)分别为 $x_1, x_2, \dots, x_m$ 的多元线性回归方程可表示为

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_m x_m \quad (2)$$

其中: $\theta_0$ 为回归常数; $\theta_1, \theta_2, \dots, \theta_m$ 为回归系数。

假设在实际问题中已测得 $n$ 组监测数据( $y_i; x_{i1}, x_{i2}, \dots, x_{im}$ )( $i=1, 2, \dots, n$ ),则式(2)可表示为

$$\begin{cases} y_1 = \theta_0 + \theta_1 x_{11} + \theta_2 x_{12} + \dots + \theta_m x_{1m} \\ y_2 = \theta_0 + \theta_1 x_{21} + \theta_2 x_{22} + \dots + \theta_m x_{2m} \\ \vdots \\ y_n = \theta_0 + \theta_1 x_{n1} + \theta_2 x_{n2} + \dots + \theta_m x_{nm} \end{cases} \quad (3)$$

采用最小二乘法<sup>[12]</sup>可求解得到式(3)的回归参数,将所求回归参数代入式(2)后,需要对回归方程进行统计量检验(F检验、T检验及多重共线性检验),若不满足检验要求则需要重新寻找影响主轴轴承温度变化的变量以便重新建模,满足要求则可直接输出模型。建模流程如图1所示。

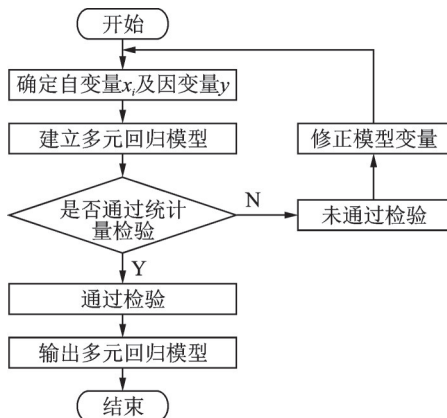


图1 多元线性回归建模流程图

Fig.1 Flowchart of Multiple linear regression modeling

## 1.3 灰色预测模型 GM (1, N)

灰色模型(GM)<sup>[13]</sup>具有建模数据少、预测效果好等优点。传统的GM(1, 1)模型是针对单一变量的预测,然而影响风力机主轴轴承温度的因素是多变量共同作用的结果,因此GM(1, N)模型更适合建立风力机主轴轴承温度预测模型。

假设一组风力机主轴轴承原始SCADA数据序列为 $X_j^{(0)} = \{X_1^{(0)}, X_2^{(0)}, \dots, X_m^{(0)}; j=1, 2, \dots, m\}$ ,其中 $X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n)\}$ 为系统特征数据序列(即风力机主轴轴承温度),而 $X_2^{(0)} = \{x_2^{(0)}(1), x_2^{(0)}(2), \dots, x_2^{(0)}(n)\}$ ,  $X_3^{(0)} = \{x_3^{(0)}(1), x_3^{(0)}(2), \dots, x_3^{(0)}(n)\}$ ,  $\dots$ ,  $X_m^{(0)} = \{x_m^{(0)}(1), x_m^{(0)}(2), \dots, x_m^{(0)}(n)\}$ 为相关因素数据序列(即影响主轴轴承温度变化的变量)。

$X_j^{(0)}$ 的一阶累加序列为 $X_j^{(1)} = \{x_j^{(1)}(1), x_j^{(1)}(2), \dots, x_j^{(1)}(n); j=1, 2, \dots, m, k=1, 2, \dots, n\}$ ,其中

$$x_j^{(1)}(k) = \sum_{i=1}^k x_j^{(0)}(i) \quad (4)$$

则系统特征数据的一次累加序列 $X_1^{(1)}$ 的紧邻均值数列为 $Z_1^{(1)} = \{z_1^{(1)}(2), z_1^{(1)}(3), \dots, z_1^{(1)}(n); k=2, 3, \dots, n\}$ ,其中

$$z_1^{(1)}(k) = \frac{1}{2} x_1^{(1)}(k) + \frac{1}{2} x_1^{(1)}(k-1) \quad (5)$$

由此可建立GM(1, N)模型为

$$x_1^{(0)}(k) + a z_1^{(1)}(k) = \sum_{j=2}^m b_j x_j^{(1)}(k) \quad (6)$$

其中: $a$ 为系统发展系数,反映预测值的发展趋势; $b_j$ 为驱动系数,反映原始数据的内在变化。

设置参数列 $\mu = (a, b_2, b_3, \dots, b_m)^T$ ,根据最小二乘法可求解式(6)得到参数列 $\mu$ 的值。

对GM(1, N)模型(式6)进行白化可得

$$\frac{dx_1^{(1)}}{dt} + a x_1^{(1)} = b_2 x_2^{(1)} + b_3 x_3^{(1)} + \dots + b_m x_m^{(1)} \quad (7)$$

求解式(7)得到时间响应函数,表示为

$$\hat{x}_1^{(1)}(k+1) = \left( x_1^{(0)}(1) - \frac{1}{a} b_j x_j^{(1)}(k+1) \right) e^{-ak} + \frac{1}{a} \sum_{j=2}^m b_j x_j^{(1)}(k+1) \quad (8)$$

最终可得累减还原式(即灰色预测值)为

$$\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k) \quad (9)$$

## 1.4 支持向量机回归 (SVR) 模型

支持向量机回归的核心思想是建立一个最优分类面,使得所有训练样本离最优分类面的误差最小<sup>[14]</sup>。风力机主轴轴承温度的SVR模型的训练与预

测借助 LIBSVM 工具箱实现。首先,选定一组风力机主轴轴承 SCADA 数据作为原始数据,对其进行归一化处理;然后,根据工具箱提供的交互检验(cross validation, 简称 VC)功能寻求最优惩罚参数  $c$  和核函数参数  $g$ (由函数 SVMcgForRegress 实现);最后,根据得到的最优参数  $c, g$  对 SVR 模型进行预测。

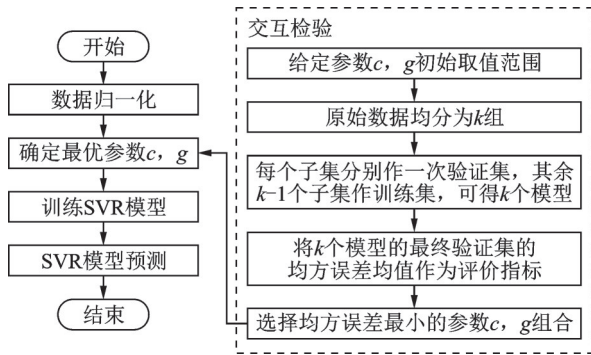


图 2 SVR 算法流程图

Fig.2 Flow chart of SVR algorithm

### 1.5 组合预测模型

假设某预测问题共有  $f$  种单一预测模型, 令  $y_t$  表示  $t$  时刻的实际监测值,  $\hat{y}_{it}$  表示第  $i$  种预测模型在  $t$  时刻的预测值(其中:  $i=1, 2, \dots, f; t=1, 2, \dots, n$ ), 第  $i$  种预测模型在  $t$  时刻的相对误差为  $e_{it}$

$$e_{it} = \left| (y_t - \hat{y}_{it}) / y_t \right| \quad (10)$$

由于每种单一预测模型具有其不足之处, 而组合预测模型将不同类型的单一模型按照一定的权重比例重新组合可发挥各种单一模型优势, 有效削弱各种单一模型的不足以达到提高预测精度的目的。因此, 建立基于熵值法<sup>[15]</sup>的组合模型, 步骤如下:

1) 求解第  $i$  种单一预测模型在  $t$  时刻的相对误差比重  $p_{it}$

$$p_{it} = e_{it} / \sum_{t=1}^n e_{it} \quad (11)$$

2) 求解第  $i$  种单一预测模型与实际值的相对误差熵值  $g_i$

$$g_i = -\frac{1}{\ln n} \sum_{t=1}^n p_{it} \ln p_{it} \quad (12)$$

3) 求解相对误差的变异系数  $d_i$

$$d_i = 1 - g_i \quad (13)$$

4) 求解各种单一预测模型的权重  $w_i$

$$w_i = g_i / \sum_{i=1}^m g_i \quad (14)$$

5) 求解组合预测模型值

$$\hat{y}_t = \sum_{i=1}^m \sum_{t=1}^n (w_i \hat{y}_{it}) \quad (15)$$

### 1.6 各模型综合评价指标

为实现对各个预测模型的定量评价, 选取均方根误差(root mean square error, 简称 RMSE)、平均绝对百分误差(mean absolute percentage error, 简称 MAPE)、平均绝对误差(mean absolute error, 简称 MAE) 及决定系数( $R^2$ ) 作为评价指标, 求解公式如下

$$RMSE = \sqrt{\sum_{t=1}^n (y_t - \hat{y}_t)^2 / n} \quad (16)$$

$$MAPE = \sum_{t=1}^n |(y_t - \hat{y}_t) / y_t| / n \quad (17)$$

$$MAE = \sum_{t=1}^n |y_t - \hat{y}_t| / n \quad (18)$$

$$R^2 = 1 - \sum_{t=1}^n (y_t - \hat{y}_t)^2 / \sum_{t=1}^n (y_t - \sum_{i=1}^m y_{it} / n)^2 \quad (19)$$

其中:  $y_t$  为实际值,  $\hat{y}_t$  为预测值; RMSE, MAPE 及 MAE 用来检验实际值与预测值的偏差及波动性大小, 数值越小则表明预测模型精度越高;  $R^2$  为自变量所解释的变化量占总变化量的比例,  $R^2$  取值为  $[0, 1]$ ,  $R^2$  越接近于 1 则表明预测精度越高。

## 2 残差统计特性分析及故障监测

### 2.1 滑动窗口残差统计特性分析

假设风力机主轴轴承温度在  $t$  时刻的实际监测值为  $y_t$ , 预测模型在  $t$  时刻的预测值为  $\hat{y}_t$ , 则预测残差序列为  $[y_1 - \hat{y}_1, y_2 - \hat{y}_2, \dots, y_N - \hat{y}_N, \dots]$ , 预测残差序列的统计特性可由残差均值及残差标准差反映。为准确描述预测残差的统计特性(均值和标准差)变化情况, 采用滑动窗口方法<sup>[16]</sup>求解预测残差序列的窗口均值及窗口标准差, 滑动窗口如图 3 所示。

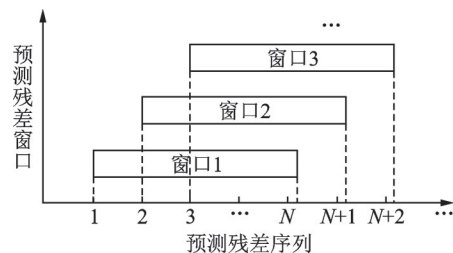


图 3 滑动窗口示意图

Fig.3 Schematic diagram of sliding window

设定宽度为  $N$  的滑动窗口, 则窗口内的  $N$  个残差的窗口均值  $\mu$  和窗口标准差  $\sigma$  求解公式为

$$\mu = \sum_{t=1}^N (y_t - \hat{y}_t) / N \quad (20)$$

$$\sigma = \sqrt{\sum_{i=1}^N (y_i - \hat{y} - \mu)^2 / (N - 1)} \quad (21)$$

## 2.2 主轴的故障监测

主轴承正常工作时,温度预测值与实际温度值接近,其残差序列分布稳定;然而主轴承发生故障会引起其温度的异常波动,进而带来预测残差统计特性的异常变化。具体表现为以下3种情况:①主轴承温度预测残差的窗口均值 $\mu$ 未发生较大改变,但其窗口标准差 $\sigma$ 变化显著,残差范围增大;②主轴承温度预测残差的窗口标准差 $\sigma$ 未发生较大改变,但其窗口均值 $\mu$ 变化显著,出现系统偏差;③主轴承温度预测残差的窗口均值 $\mu$ 及标准差 $\sigma$ 均变化显著。

基于上述情况,可根据预测残差序列的统计特性变化来监测主轴承运行状态,设定残差均值和残差标准差故障临界值分别为 $\mu_v$ 和 $\sigma_v$ 。当温度预测残差的均值或标准差超过某一临界值时,系统产生故障报警告知维护人员进行检修。假设滑动窗口的温度残差序列均值绝对值最大值为 $\mu_{\max}$ ,标准差绝对值最大值为 $\sigma_{\max}$ ,则主轴承温度故障的临界标准为

$$\begin{cases} \mu_v = \pm k_1 \mu_{\max} \\ \sigma_v = k_2 \sigma_{\max} \end{cases} \quad (22)$$

式(22)中 $k_1, k_2$ 可由现场运维人员根据经验确定。根据预测模型对主轴承温度进行预测时,会存在各种不确定性因素对输出结果造成干扰<sup>[17]</sup>。为简化起见,将预测残差序列假设为一组均值和方差未知的正态分布数据,在求解残差窗口均值和标准差时需要给出置信度为 $1 - \alpha$ 的均值和标准差的置信区间,表示为

$$A = \left[ \mu - \frac{\sigma}{\sqrt{N}} t_{\alpha/2}(N-1), \mu + \frac{\sigma}{\sqrt{N}} t_{\alpha/2}(N-1) \right] \quad (23)$$

$$B = \left[ \frac{\sqrt{N-1} \sigma}{\sqrt{\chi^2_{\alpha/2}(N-1)}}, \frac{\sqrt{N-1} \sigma}{\sqrt{\chi^2_{1-\alpha/2}(N-1)}} \right] \quad (24)$$

其中: $t_{\alpha/2}$ 为 $t$ 分布的 $\alpha/2$ 分位点; $\chi^2_{\alpha/2}$ 为 $\chi^2$ 分布的 $\alpha/2$ 分位点,其值在给定置信度 $1 - \alpha$ 后可通过查阅 $t$ 分布临界值表和 $\chi^2$ 分布临界值表得知。

当残差均值或标准差的置信区间超过各自临界值时,发出风力机主轴承故障报警信息。

## 3 案例分析

### 3.1 建模变量的选取

根据文献[18]可知,关键部件温度相关联的特征量主要包括环境温度、风速、转速及功率等,此外

由于部件温度的热惯性特点使得前一时刻的部件温度对当前时刻部件温度有直接影响。因此,建立主轴承温度预测模型时初步选择风力机输出功率、风速、环境温度、主轴转速及上时刻主轴承温度作为输入变量,主轴承当前时刻温度( $y$ )作为输出变量。

选取某风力机2019年4月主轴承正常状态时的运行参数进行建模,根据所建模型对5月份主轴承温度进行预测。按照1.1部分所述步骤对原始SCADA数据进行归一化处理,归一化后的部分数据如表2所示。

表1 SCADA数据的归一化

Tab.1 Normalization of SCADA data

序号	时刻	功率	主轴 转速	主轴承 温度	环境 温度	...
1	00:10	0.159 0	0.344 3	0.769 2	0.040 7	...
2	00:20	0.163 2	0.337 6	0.743 6	0.038 0	...
3	00:30	0.224 0	0.473 7	0.743 6	0.032 0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

### 3.2 MLR, GM (1, N), SVR及组合预测模型的求解

#### 3.2.1 多元线性回归模型(MLR)

以风力机输出功率( $x_1$ )、风速( $x_2$ )、空气温度( $x_3$ )、主轴转速( $x_4$ )及上时刻主轴承温度( $x_5$ )为自变量,风力机主轴承当前时刻温度( $y$ )为因变量建立多元线性回归方程,根据1.1部分所述步骤求解得到的MLR模型为

$$y = 0.0321 - 0.0219x_1 + 0.0073x_2 + 0.0079x_3 + 0.0265x_4 + 0.9421x_5 \quad (25)$$

对式(25)进行统计量检验时发现输出功率( $x_1$ )所对应的方差扩大因子(variance inflation factor, 简称VIF)为15.6418,风速( $x_2$ )所对应的VIF为15.7134,表明所建模型存在严重的多重共线性(VIF > 10),可能会给回归结果带来较大误差,因此需要重新修正模型输入变量,以消除多重共线性。

消除多重共线性的常用方法则是剔除VIF最大值所对应的变量,即风速( $x_2$ );剔除后的MLR修正模型自变量为风力机输出功率( $x_1$ )、空气温度( $x_3$ )、主轴转速( $x_4$ )及上时刻主轴承温度( $x_5$ ),根据1.1部分所述步骤求解得到的MLR修正模型为

$$y = 0.0329 - 0.0180x_1 + 0.0079x_3 + 0.0269x_4 + 0.9418x_5 \quad (26)$$

据检验,式(26)已消除多重共线性,F检验及T检验各项指标均满足显著性条件。利用修正后的MLR模型对5月份主轴承温度进行预测,部分预测值及残差值见表2。

表 2 MLR 模型部分温度预测结果值

Tab.2 Temperature prediction results of MLR model

序号	实际值	MLR 预测值	残差
1	59	58.913 0	0.087 0
2	60	58.914 3	1.085 7
3	60	59.848 2	0.151 8
⋮	⋮	⋮	⋮

3.2.2 灰色预测模型 GM(1,N)的求解

基于 MLR 修正模型所得自变量及因变量,选择当前时刻主轴承温度、功率、空气温度、主轴转速及上时刻主轴承温度的时间序列为原始数据序列  $X_j^{(0)} = \{X_1^{(0)}, X_2^{(0)}, X_3^{(0)}, X_4^{(0)}, X_5^{(0)}\}$ 。其中:当前时刻主轴承温度数据序列  $X_1^{(0)}$  为系统特征数据序列;而功率数据序列  $X_2^{(0)}$ 、空气温度数据序列  $X_3^{(0)}$ 、主轴转速数据序列  $X_4^{(0)}$  以及上时刻主轴承温度序列  $X_5^{(0)}$  均为相关因素序列。选取 2019 年 4 月份 SCADA 数据依据 1.2 所述步骤进行建模,求解得到 GM(1,N) 模型参数列  $\mu$  的值为

$$\mu = (2.0145, -0.0006, 0.0625, 0.1465, 1.9723)^T$$

将参数列  $\mu$  代入式(8),得到时间相应函数,而后根据式(9)累减可得到 GM(1,N) 预测值,部分预测值及残差见表 3。

表 3 GM(1, N)模型部分温度预测结果值

Tab.3 Temperature prediction results of GM(1, N) model

序号	实际值	MLR 预测值	残差
1	59	59.141 3	-0.141 3
2	60	58.130 0	1.870 0
3	60	60.903 6	-0.903 6
⋮	⋮	⋮	⋮

3.2.3 支持向量机回归模型(SVR)的求解

基于 MLR 修正模型所得自变量及因变量,选择 2019 年 4 月份的当前时刻主轴承温度、功率、空气温度、主轴转速及上时刻主轴承温度等序列数据用来训练 SVR 模型,用训练好的模型来预测 5 月份主轴承温度值。

给定参数  $c$  和  $g$  的初始范围均为  $[-8, 8]$ , 根据交互检验法求解得到 SVR 模型最优参数  $c$  和  $g$  为: Best  $c=1.4142$ , Best  $g=0.3535$ ; 根据所求最优参数  $c$  和  $g$  对 SVR 模型进行训练,然后对 5 月份数据进行预测,部分预测值及残差值见表 4。

3.2.4 MLR、GM(1,N)及 SVR 组合模型的求解

假设 MLR 修正模型所得主轴承温度预测值为  $\hat{y}_{1t}$ , GM(1,N) 模型所得预测值为  $\hat{y}_{2t}$ , SVR 模型所得

表 4 SVR 模型部分温度预测结果值

Tab.4 Temperature prediction results of SVR model

序号	实际值	MLR 预测值	残差
1	59	59.244 2	-0.244 2
2	60	59.247 5	0.752 5
3	60	60.238 2	-0.238 2
⋮	⋮	⋮	⋮

预测值为  $\hat{y}_{3t}$ , 则应用熵值法确定的各模型最优权重为  $w_1=0.1258, w_2=0.3105, w_3=0.5637$ , 因此可得组合模型表达式为

$$\hat{y}_t = 0.1258\hat{y}_{1t} + 0.3105\hat{y}_{2t} + 0.5637\hat{y}_{3t} \quad (27)$$

根据所求组合模型对 5 月份数据进行预测,部分预测值及残差见表 5。

表 5 组合模型部分温度预测结果值

Tab.5 Temperature prediction results of Combined model

序号	实际值	MLR 预测值	残差
1	59	59.244 2	-0.244 2
2	60	59.247 5	0.752 5
3	60	60.238 2	-0.238 2
⋮	⋮	⋮	⋮

3.3 各模型预测结果对比分析

根据各模型对 5 月份主轴承温度数据进行预测,实际值与各模型预测值变化趋势如图 4 所示。

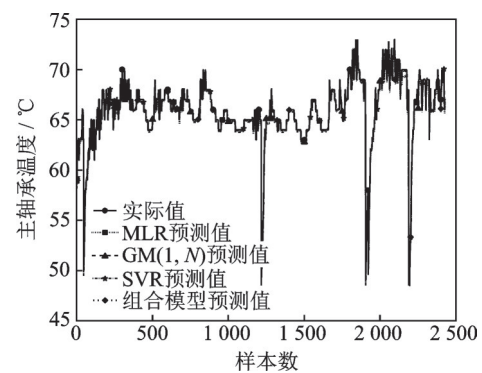


图 4 风力机主轴承温度预测图

Fig.4 Forecast figure of wind turbine main bearing temperature

图 4 显示在第 48, 1 220, 1 928 及 2 206 点附近处主轴承温度较低,查询主状态日志可知是由于风力机停机重新启动所造成。另外从图中可看到,不同模型所求主轴承温度预测值与实际值变化趋势相近,拟合程度较高;但根据图 4 不易直观判断哪种模

型预测值与实际值吻合度更好。为进一步对比各模型优劣性,进而求解各模型综合评价指标值,见表6。

表6 各模型指标对比值

Tab.6 Comparison results of each model index

评价指标	MLR模型	GM(1,N)模型	SVR模型	组合模型
RMSE	0.489 4	0.464 8	0.434 6	0.433 9
MPPE	0.005 2	0.004 8	0.004 2	0.004 0
MAE	0.349 1	0.318 4	0.272 2	0.265 4
R <sup>2</sup>	0.932 5	0.979 1	0.981 6	0.981 8

由表5的各指标数据对比可知,组合模型的RMSE,MPE及MAE最小,R<sup>2</sup>最大,表明组合模型偏差及波动性更小,在各模型中具有最高的预测精度;其中组合模型的R<sup>2</sup>值分别较其他模型提高0.049 3,0.002 7和0.000 2。因此,选择在组合预测模型基础上引入滑动窗口方法实现对风力机主轴轴承的故障监测。此外在单一预测模型中,SVR模型具有良好的预测精度,各项指标均和组合模型指标相近。

3.4 基于滑动窗口的主轴承温度故障状态监测

查询风力机主状态日志可知,5月份风力机主轴承未发生故障。为验证方法有效性,人为模拟主轴承因发生故障而导致主轴承温度升高的情况。选择5月份主轴承温度较高的12~15号3天累计568个温度点进行研究,取滑动窗口宽度N=20,置信水平 $\alpha=0.05$ ,根据式(23),(24)求解得到主轴承正常状态下预测残差均值及标准差的各自95%置信区间,它们的变化趋势如图5所示。

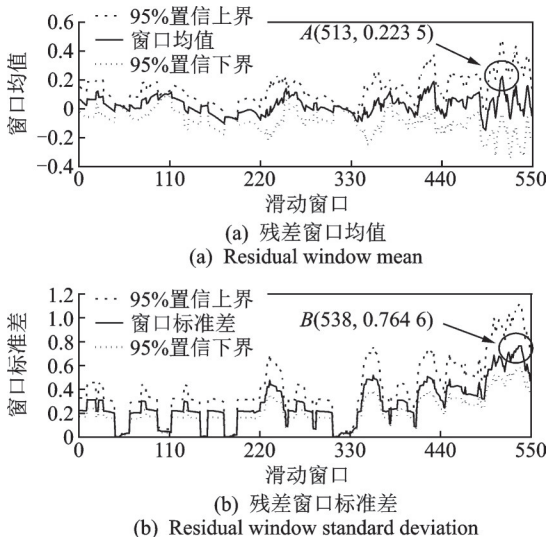


图5 残差窗口均值和标准差

Fig.5 Residual window mean and standard deviation

图5(a)为主轴承温度残差窗口均值变化趋势图,从图5(a)可知主轴承温度残差窗口均值最大值为 $\mu_{max}=0.223 5$ (即A点);图5(b)为主轴承温度残差窗口标准差变化趋势图,由图5(b)可知主轴承温度残差窗口标准差最大值为 $\sigma_{max}=0.764 6$ (即B点),由于是人为模拟故障,所以残差标准差并未发生变化。故选择残差窗口均值展开分析,取 $k_1=2$ ,则窗口均值临界值根据式(22)可得 $\mu_c=\pm 0.447 0$ 。

从第303个温度点开始人为加入步长为0.015℃的温度累积偏移量,基于组合模型的主轴承温度预测残差均值变化趋势如图6所示。

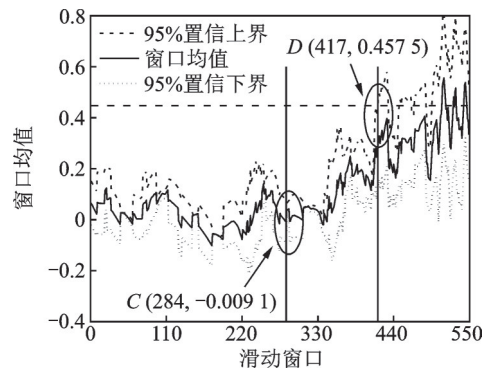


图6 模拟故障状态的残差窗口均值

Fig.6 Residual window mean value of simulated fault state

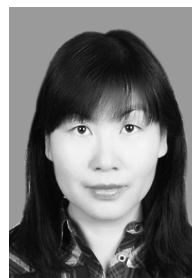
图6显示,主轴承正常运行时残差均值95%置信区间均未超过临界值,在C点即第284个窗口(303-20+1)加入温度偏移后,滑动窗口残差均值逐渐增大,在D点残差窗口均值95%置信上界超过临界值0.447 0,引发主轴承故障报警。综上可得,当主轴承发生故障导致其温度异常时,基于滑动窗口方法的故障监测方法可及时发现故障并报警,有效监控主轴承异常状态。

4 结束语

基于风力机SCADA数据分别建立主轴承温度的MLR,GM(1,N),SVR 3种单一预测模型及它们的组合预测模型;其中组合模型有着更优越的预测性能,其决定系数分别比其他模型提高0.049 3,0.002 7及0.000 2。此外SVR模型在单一预测模型中有着更高的预测精度。在组合预测模型基础上引入滑动窗口方法求解分析温度预测残差的统计特性(均值和标准差),当窗口均值或标准差的置信区间超过设定临界值时可判断出主轴承发生故障,有效监测主轴承运行状态。笔者所提方法可用于工程实际中风力机或其他机械设备的关键部件温度预测及故障监测。

## 参 考 文 献

- [1] HU Y G, LI H, SHI P P, et al. A prediction method for the real-time remaining useful life of wind turbine bearings based on the Wiener process [J]. *Renewable Energy*, 2018, 127: 452-460.
- [2] STROMBERGSSON D, MARKLUND P, BERGLUND K, et al. Bearing monitoring in the wind turbine drivetrain: a comparative study of the FFT and wavelet transforms [J]. *Wind Energy*, 2020, 23(6): 1391-1393.
- [3] 程静, 王维庆, 樊小朝, 等. 基于二值双谱和模糊聚类的风电轴承故障诊断[J]. *振动、测试与诊断*, 2018, 38(4): 765-771.  
CHENG Jing, WANG Weiqing, FAN Xiaochao, et al. Bearing fault pattern recognition of wind turbine based on two-value bispectrum feature-fuzzy clustering method[J]. *Journal of Vibration, Measurement & Diagnosis*, 2018, 38(4): 765-771. (in Chinese)
- [4] ZHENG Y Q, CAO Y Y, ZHANG C C, et al. Structural optimization design of large wind turbine blade considering aeroelastic effect [J]. *Mathematical Problems in Engineering*, 2017, 2017(9): 1-7.
- [5] 汤宝平, 罗雷, 邓蕾, 等. 风电机组传动系统振动监测研究进展[J]. *振动、测试与诊断*, 2017, 37(3): 417-425.  
TANG Baoping, LUO Lei, DENG Lei, et al. Research progress of vibration monitoring for wind turbine transmission system[J]. *Journal of Vibration, Measurement & Diagnosis*, 2017, 37(3): 417-425. (in Chinese)
- [6] GUO Y, ZHAO Z, SUN R, et al. Data-driven multiscale sparse representation for bearing fault diagnosis in wind turbine [J]. *Wind Energy*, 2019, 22(4): 587-604.
- [7] TZIAVOS N I, HEMIDA H, DIRAR S, et al. Structural health monitoring of grouted connections for offshore wind turbines by means of acoustic emission: An experimental study [J]. *Renewable Energy*, 2020, 147: 130-140.
- [8] JIN X H, QIAO W, PENG Y Y, et al. Quantitative evaluation of wind turbine faults under variable operational conditions [J]. *IEEE Transactions on Industry Applications*, 2016, 52(3): 2061-2069.
- [9] SHAHRIAR M R, BORGHESANI P, LEDWICH G, et al. Performance analysis of electrical signature analysis-based diagnostics using an electromechanical model of wind turbine [J]. *Renewable Energy*, 2017, 116: 15-41.
- [10] 李辉, 胡姚刚, 李洋, 等. 大功率并网风电机组状态监测与故障诊断研究综述[J]. *电力自动化设备*, 2016, 36(1): 6-16.  
LI Hui, HU Yaogang, LI Yang, et al. Overview of condition monitoring and fault diagnosis for grid-connected high-power wind turbine unit [J]. *Electric Power Automation Equipment*, 2016, 36(1): 6-16. (in Chinese)
- [11] 张帆, 刘德顺, 戴巨川, 等. 一种基于SCADA参数关系的风电机组运行状态识别方法[J]. *机械工程学报*, 2019, 55(4): 1-9.  
ZHANG Fan, LIU Deshun, DAI Juchuang, et al. An operating condition recognition method of wind turbine based on SCADA parameter relations [J]. *Journal of Mechanical Engineering*, 2019, 55(4): 1-9. (in Chinese)
- [12] FANG T T, LAHDELMA R. Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system [J]. *Applied Energy*, 2016, 179: 544-552.
- [13] YIN K D, GENG Y, LI X M. Improved grey prediction model based on exponential grey action quantity [J]. *Journal of Systems Engineering & Electronics*, 2018, 29(3): 560-570.
- [14] CHENG Y, LIU Q, WANG J, et al. Distributed fault detection for wireless sensor networks based on support vector regression [J]. *Wireless Communications & Mobile Computing*, 2018, 2018:1-8.
- [15] MUHAMMAD W, MUHAMMAD A, TAEWOONG K. Development of a new composite drought index for multivariate drought assessment [J]. *Journal of Hydrology*, 2015, 527: 30-37.
- [16] 郭鹏, DAVID Infield, 杨锡运. 风电机组齿轮箱温度趋势状态监测及分析方法[J]. *中国电机工程学报*, 2011, 31(32): 129-136.  
GUO Peng, DAVID Infield, YANG Xiyun. Wind turbine gearbox condition monitoring using temperature trend analysis [J]. *Proceedings of the CSEE*, 2011, 31(32): 129-136. (in Chinese)
- [17] GAN C B, WANG Y H, YANG S X. Nonparametric modeling on random uncertainty and reliability analysis of a dual-span rotor [J]. *Journal of Zhejiang University-Science A (Applied Physics & Engineering)*, 2018, 19(3): 189-202.
- [18] 李辉, 胡姚刚, 李洋, 等. 基于温度特征量的风电机组关键部件劣化渐变概率分析[J]. *电力自动化设备*, 2015, 35(11): 1-7.  
LI Hui, HU Yaogang, LI Yang, et al. Gradual deterioration probability analysis based on temperature characteristic parameters for critical of wind turbine generator system [J]. *Electric Power Automation Equipment*, 2015, 35(11): 1-7. (in Chinese)



**第一作者简介:**郑玉巧,女,1977年7月生,博士、副研究员、硕士研究生导师。主要研究方向为机械结构动态设计与优化、机械系统动力学及生产系统工程等。曾发表《基于均匀设计的风力机塔架结构参数优化研究》(《太阳能学报》2020年第41卷第4期)等论文。  
E-mail: zhengyuxiaolul@163.com。