

# 基于改进的多元离群检测方法的风机齿轮箱早期故障诊断

顾煜炯<sup>1</sup> 贾子文<sup>1</sup> 王 瑞<sup>1</sup> 任玉亭<sup>2</sup>

1.华北电力大学新能源电力系统国家重点实验室,北京,102206

2.国华能源投资有限公司,北京,100007

**摘要:**针对风电机组运行工况波动性以及机组早期故障特征不易提取的特点,提出一种基于改进的多元离群监测方法来实现风机齿轮箱故障的早期诊断。运用阶比重采样方法对原始振动信号进行预处理,并对处理结果进行量纲一因子分析;通过马氏距离建立风电齿轮箱的早期故障识别模型;利用多元线性回归改进多元离群检测算法进行实际数据的分析计算。结果表明,该方法较原始方法能够更早地察觉出风电齿轮箱早期故障。

**关键词:**阶比重采样;量纲一因子分析;多元线性回归;多元离群检测

**中图分类号:**TK83

**DOI:**10.3969/j.issn.1004-132X.2016.14.011

## Early Fault Diagnosis for Wind Turbine Gearbox Based on Improved Multivariate Outlier Detection

Gu Yujiong<sup>1</sup> Jia Ziwen<sup>1</sup> Wang Rui<sup>1</sup> Ren Yuting<sup>2</sup>

1.State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing, 102206

2.Guohua Energy Investment Limited Company, Beijing, 100007

**Abstract:** An improved method of multivariate outlier detection was used in early fault diagnosis for wind turbine gearboxes, which might extract the early fault features under fluctuation working conditions. First, the primitive vibration signals were preprocessed by order resampling, and the processed results were analyzed by dimensionless parameter analysis. Second, a model of early fault diagnosis was created based on Mahalanobis distance, which was used for turbine gearboxes. Finally, the actual data was analyzed by the method of multivariate outlier detection, which was improved by multiple linear regression. The results show that the new method may detect the gearbox faults earlier than original one.

**Key words:** order resampling; dimensionless parameter analysis; multiple linear regression (MLR); multivariate outlier detection

## 0 引言

风电场通常建设在地处偏远、交通闭塞、环境因素变化剧烈的区域,这使得风电机组的健康运行遭受极大挑战<sup>[1]</sup>。风机齿轮箱作为整个系统的传动机构,内部结构紧凑,部件之间耦合性较强,在运行过程中长期受到交变载荷与冲击载荷作用,容易造成齿轮点蚀、磨损等故障。同时,齿轮箱故障维修过程较为复杂,维修时间长,长时间的停机维修给风场业主带来很大的经济损失<sup>[2]</sup>。因此,进行风电机组齿轮箱早期故障诊断的研究,寻找故障早期信号特征,确定故障模式,在故障还没有发展到严重程度时及时排除安全隐患,对保证机组正常运行和提高风场经济效益具有重要

意义。

目前,针对齿轮箱故障诊断的方法有很多,如经验模态分解(EMD)与隐马尔科夫模型<sup>[3]</sup>、小波分析<sup>[4]</sup>、神经网络技术<sup>[5]</sup>等方法。这些方法虽然能够实现齿轮箱故障的诊断,但都主要面对设备的中晚期故障问题,对齿轮箱早期故障特征提取和分析的过程较少。风电机组齿轮箱早期故障诊断具有如下特性:①风电机组运行工况具有波动性、间歇性的特点,造成现场采集的振动信号具有明显的非线性和非平稳特征,需要有对应的方法进行处理,否则在很大程度上会影响振动信号特征提取的效果;②机组齿轮箱早期故障信号特征表现不明显,需要找到对早期故障数据敏感的参数并将其作为征兆,实现机组齿轮箱的早期诊断。针对这一现状,本文提出一种改进的多元离群检测方法对风电齿轮箱进行故障的早期诊断。通过对齿轮箱故障特征参数进行隶属度划分,建立多

收稿日期:2015-09-10

基金项目:神华集团科技创新项目(GTKJ-12-02);华能集团科学技术项目(HNKJ-H27)

元统计数理模型,运用多元离群检测方法实现风电机组齿轮箱故障的在线诊断。

## 1 阶比重采样角域信号特征的获取

有量纲幅域参数受到工作负载、转速等运行条件的影响,使得故障趋势判断存在误差。峰值、翘度、波形等量纲一幅域参数指标对幅值能量变化不敏感,即与设备运行条件关系不大,但对设备故障较为敏感,并且计算过程简单,容易应用到实际工程中<sup>[6]</sup>。郭厚明等<sup>[7]</sup>运用量纲一因子实现了矿用低速重载齿轮的故障诊断;岑少起等<sup>[8]</sup>通过对监测参数进行量纲一因子转换,实现了滑动轴承的动力学分析。目前,量纲一因子在旋转机械早期故障诊断领域的应用较少,所以,笔者利用量纲一因子的特性,通过提取齿轮箱振动数据信息,计算齿轮箱振动数据的因子值,并将其作为故障早期诊断的特征。

受风电机组工况影响,齿轮箱振动信号存在较为明显的非线性特征。所以本文先通过阶比重采样技术实现非平稳时域信号向平稳角域信号的转化,建立了角域信号量纲一因子变化趋势指标。

### 1.1 阶比重采样技术

风电机组运行工况的特殊性导致机组振动信号具有明显的非线性、非平稳性特征,运用传统的信号分析方法很难提取出反映设备运行状况的振动特征。阶比重采样通过对振动信号和转速信号进行同步采集,并结合插值计算的方法保证设备每一转采样的点数相等,实现非平稳时域信号向平稳角域信号的等价转换<sup>[9]</sup>。

假设机组角度编码器码盘孔数为  $n$ ,则孔间角度  $\Delta\theta = 360/n(^{\circ})$ ,角度编码器计数脉冲每秒脉冲数序列记为  $\{m_1, m_2, \dots, m_k\}$ ,  $k$  为时间序列序号。可以得出风电机组第  $i$  秒的平均转速 (r/min):

$$R_i = 60m_i \Delta\theta \quad (1)$$

则可依据每秒平均转速得出累计转角曲线。假设等时间间隔  $\{t_1, t_2, \dots, t_k\}$  采样下的秒级时域序列为  $\{x_1, x_2, \dots, x_k\}$ , 其中,  $x_k$  为时域振动幅值。则时域序列转化成角域序列的公式如下:

$$\theta_i = \frac{1}{60} \sum_{j=1}^k R_j \quad (2)$$

将等时间时域信号  $\{x_1, x_2, \dots, x_k\}$  转化为等时间间隔采样下的角域信号  $\{y_1, y_2, \dots, y_k\}$ , 因为机组在此时间不一定是恒定转速运行的,所以此角域信号非严格意义上的等角度间隔角域信号序列,为解决此问题,采用数值插值方法对其进行

进一步计算。

重构等时间间隔下的角度序列  $\{\theta'_1, \theta'_2, \dots, \theta'_p\}$ , 序列中的  $\theta'_i$  为等角度间隔采样中第  $i$  个角度序列,  $p$  可以看成等角度间隔采样频率,理论上  $p$  的值应大于传动机构最大阶比的 2 倍(采样定理),工程中,  $p$  的值通常为传动机构最大阶比的 6~10 倍。通过数值差值方法,寻找到  $\{\theta'_1, \theta'_2, \dots, \theta'_p\}$  在等时间间隔下对应时刻的角度序列  $\{\theta_1, \theta_2, \dots, \theta_k\}$  中的位置,计算出对应幅值,得出最终的等角度间隔角域信号。

### 1.2 齿轮箱早期故障特征提取

新的量纲一因子值趋势分析技术主要解决风电机组齿轮箱早期故障振动信号特征值提取困难的问题。通过对齿轮箱等角域信号量纲一因子值的计算,分析量纲一数值变化规律来反映风电齿轮箱的故障发展趋势<sup>[10]</sup>: 机组运行正常时,各个量纲一因子值没有明显的变化趋势;机组出现早期故障时,各个量纲一因子的值会出现不同的单调性变化,通过这些因子值的变化趋势以及数值变化程度,实现机组早期故障模式的甄别。

笔者为克服传统有量纲幅域参数与能量有关,或与能量无关但是只能定性分析特性的缺点,基于时域波动统计分析,引入对能量不敏感的量纲一幅域参数,将其作为故障特征参数。各因子的具体描述如下:

(1) 翘度  $K_f$ 。  $K_f$  与旋转部件的尺寸、设计参数、运行工况等无关,对冲击信号反应敏感,其数值随故障发展程度加深而增大,且变化较为明显。  $K_f$  的计算公式为

$$K_f = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{\sigma} \right)^4 \quad (3)$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

式中,  $x_i$  为数据数列中的第  $i$  个数据的值;  $\sigma$  为数据标准差;  $N$  为数据个数。

(2) 波形裕度  $C_L$  与偏态因子  $S_K$ 。这两个因子对振动信号形状和趋势的细微波动变化敏感,大量实验证明,这两个因子可以作为判断旋转机械早期故障的指标,其计算公式分别为

$$C_L = \frac{N^2 X_{\max}}{\left( \sum_{i=1}^N \sqrt{|x_i|} \right)^2} \quad (4)$$

$$S_K = \frac{\sum_{i=1}^N x_i^3}{\sqrt{\frac{1}{N} \left( \sum_{i=1}^N x_i^2 \right)^3}} \quad (5)$$

式中,  $X_{\max}$  为数据数列中的最大值。

(3) 重复性因子  $R_f$ 。该指标适合对波形的重

复性进行定量分析,随故障的发展其波形重复性变差,因子值发生变化。将数据按整周期截取成  $s$  段,每段  $t$  个数,则可得数据  $\{x_{11}, x_{12}, \dots, x_{1t}; x_{21}, x_{22}, \dots, x_{2t}; \dots; x_{s1}, x_{s2}, \dots, x_{st}\}$ , 每个时间段的数据可以看作一个角域序列,则对应的平均重复性波形定义为

$$x = \frac{1}{s} \sum_{i=1}^s x_{ij} \quad j = 2, 3, \dots, t-1 \quad (6)$$

计算重复波形平均差分值,运用链码技术对波形差分符号进行编辑,规则如图 1 所示,图中,0 表示波形下降,差分为负值;1 表示波形不变,差分为 0;2 表示波形上升,差分为正值。对平均重复波形编码进行比较,得到编码中不相同点的个数  $k$ ,则重复性因子为

$$R_t = k/t \quad (7)$$

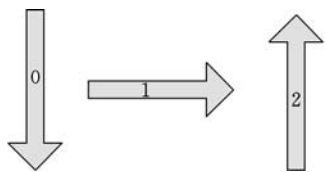


图 1 波形差分编码规则

(4) 相似性因子  $F_f$ 。此因子涉及到分形理论内容,它运用盒维数的概念反映故障信号早期变化情况。首先对角域信号进行符合其运算要求的标准化处理:

$$x'_i = \frac{Kx_\gamma(t_i)}{\max |x_\gamma(j)|} \quad j = 1, 2, \dots, n \quad (8)$$

式中,  $x_\gamma(t_i)$  为  $t_i$  时刻的幅值;  $\max |x_\gamma(j)|$  表示记录数据中幅值绝对值最大的值;  $K$  为比例放大因子,一般取整数,需根据经验与实际情况而定。

分形盒维数以单元方格的方式对信号波形进行覆盖,要求覆盖单元有较强相似性,且整个信号波形有严格的自相似性。设  $F$  是实数集合  $R_n$  中任意非空有限子集,记  $N(F, \varphi)$  为最大边长  $\varphi$  能覆盖  $F$  集合的最小数字,则  $F$  的盒维数计算表达式为

$$\dim_B F = \lim_{\varphi \rightarrow 0} \frac{\ln N(F, \varphi)}{\ln \varphi} \quad (9)$$

相似性因子为

$$F_f = \dim_B F \quad (10)$$

(5) 跳跃性因子  $J_f$ 。此因子主要反映振动波形的幅度调制,通过式(8)对原始角域信号进行标准化处理,对处理后的数据进行整周期截取,此过程类似重复性因子对数据的分段处理,计算方差值:

$$J_f = \frac{1}{N} \sum_{i=1}^N (x_{ip} - \bar{X})^2 \quad (11)$$

式中,  $x_{ip}$  为分段数据标准化后第  $i$  段数据中的最小值;  $\bar{X}$

为所有标准化最小值的平均值。

### 1.3 实际案例分析

以沧州某风电场 1.5 MW 风电机组为例,经调查,2013 年 2 月初的第 3 次全场巡检发现某一风电机组一级行星太阳轮发生较严重的磨损。从风场中继室数据库调取该机组整个 2 月内齿轮箱一级行星轮位置测点全部振动数据,对原始振动数据进行阶比重采样计算,将时域信号转化成平稳的等角度角域信号后,以天为单位对 2 月份 29 天的历史数据进行传统量纲一因子和新量纲一因子的分析计算。

由图 2 可以看出:在机组故障初期,新的量纲一因子数值随故障发展有较为明显的变化。对比图 2、图 3 可以看出:新量纲一因子较传统量纲一因子的变化趋势更为明显,数据波动小,说明新的量纲一因子对机组早期故障信号的微弱波动变化敏感。所以通过量纲一因子对角域信号的特征提取,能够准确分析出风机齿轮箱早期故障。

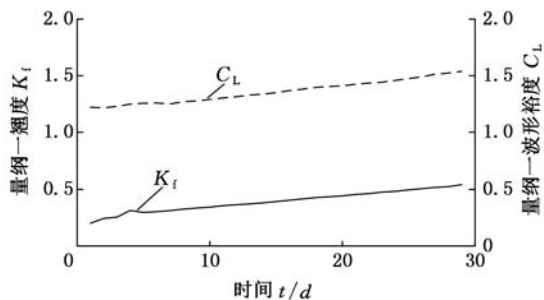


图 2 新量纲一翘度  $K_i$  和波形裕度  $C_L$  变化趋势图

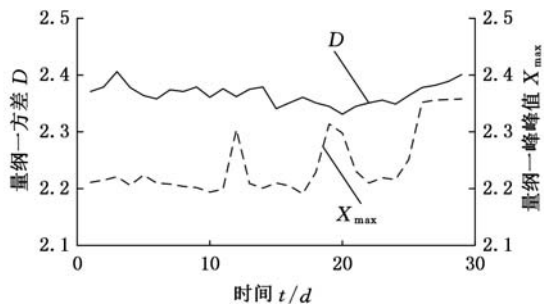


图 3 传统量纲一方差值  $D$  和峰峰值  $X_{max}$  变化趋势图

## 2 多元离群检测方法的改进

多元故障特征离群检测可以有效地将多元时序特征融合成综合评价指标,通过综合指标的差异反映设备故障的严重程度,实现对设备早期故障的等级划分。本文选取基于距离的离群检测方法来解决融合多元故障特征的风电机组早期故障模式预警问题。基于距离的离群点检测算法,一方面避免了数据分布模型不确定时,检测精确度较低的问题;另一方面,当处理属性较多的空间数据集时,基于距离的离群点检测算法比基于密度

的离群点检测方法效率更高,这对于将数据异常检测用于旋转机械设备的在线监测与数据分析至关重要。

离群检测中,距离求解方法很多,这里采用马氏(Mahalanobis)距离实现对机组齿轮箱运行状态的监测。传统的马氏距离方法在进行样本数据与被检测数据距离比较的时候,没有对影响故障模式的各个特征参数进行隶属度的划分。因为故障数据对各个故障特征波动的影响不同,同时,每个故障特征对数据变化的敏感程度也有差异,所以,找到各个故障模式下故障特征之间的隶属关系,可以提高对故障模式变化显著的因子的贡献度。

为解决计算马氏距离时各影响因素隶属度的问题,这里采用多元线性回归的方法。

## 2.1 多元线性回归方程

多元线性回归<sup>[11]</sup>方程是描述因变量  $Y$  的平均值或期望值如何依赖于自变量  $X_1, X_2, \dots, X_p$  的方程。多元线性回归方程形式为

$$E(Y) = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p \quad (12)$$

其中,  $b_i (i=0, 1, \dots, p)$  为待定参数,表示假定其他变量不变,仅  $X_i$  每变动一个单位时,  $Y$  的平均变量。

总体参数  $b_0, b_1, \dots, b_p$  是未知的,必须利用样本观测值去估计它们。估计的多元线性回归方程为

$$y'_i = b'_0 + b'_1 x_{1i} + b'_2 x_{2i} + \dots + b'_p x_{pi} \quad (13)$$

式中,  $b'_i$  为不同特征参数的隶属度;  $x_{pi}$  为特征参数  $p$  在第  $i$  个时刻的参数值;  $y'_i$  为在第  $i$  个时刻受各个参数影响的因变量。

## 2.2 参数最小二乘估计

多元线性回归分析采用最小二乘法来估计待定参数。满足使剩余平方和  $S_{SE} = \sum_{i=1}^n (y_i - y'_i)$  取得最小值。根据最小二乘法要求,可得求解各参数的方程:

$$\frac{\partial S_{SE}}{\partial b'_0} = \frac{\partial S_{SE}}{\partial b'_j} = 0 \quad j = 1, 2, \dots, p \quad (14)$$

式中,  $y_i$  为在第  $i$  个时刻实际观察的因变量数值。

计算出各个自变量的待定参数后,要对回归方程拟合优度、回归方程显著性和回归系数显著性进行检验。

## 2.3 改进马氏距离的计算

马氏距离<sup>[12]</sup>的计算公式为

$$d_{ij}^2(\mathbf{X}_i - \mathbf{X}_j) = (\mathbf{X}_i - \mathbf{X}_j)^T \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \mathbf{X}_j) \quad (15)$$

其中,  $d_{ij}^2(\mathbf{X}_i - \mathbf{X}_j)$  为第  $i$  个和第  $j$  个数据向量的马氏距离;  $\mathbf{X}_i, \mathbf{X}_j$  为第  $i$  个和第  $j$  个数据对象的属

性向量矩阵;  $\boldsymbol{\Sigma}^{-1}$  为属性向量的协方差矩阵。

在实际应用中,  $\mathbf{X}_i, \mathbf{X}_j$  均可量化处理出对应的数值,例如信号能量、平均温度等,作为多元线性回归方程的因变量。由式(14)计算得出  $\mathbf{X}_i, \mathbf{X}_j$  向量中各个数据的参数,假设有  $p$  个特征参数,则记参数隶属度序列为  $\{b_0, b_1, b_2, \dots, b_p\}$ 。这样就对参与计算的向量  $\mathbf{X}_i, \mathbf{X}_j$  中每一个数据完成隶属度划分,则改进的马氏距离为

$$d_{ij}^2(\mathbf{X}_i - \mathbf{X}_j) = (b_i \mathbf{X}_i - b_j \mathbf{X}_j)^T \boldsymbol{\Sigma}^{-1} (b_i \mathbf{X}_i - b_j \mathbf{X}_j) \quad (16)$$

## 3 齿轮箱早期故障模型建立

对基于马氏距离的多元故障特征离群检测方法进行一定的优化,从而更好地解决风电机组的早期故障模式预警问题。本节首先对故障模式各个故障特征进行隶属度分析,在分析过程中,先将振动信号的能量  $E = \sum_{i=1}^m x_i^2$  作为各个故障特征变化的因变量,  $x_i$  为振动信号的幅值;再确定故障模式以及正常工况下各个故障特征值;最后通过马氏距离计算得出对应故障模式下的距离阈值。

### 3.1 故障特征隶属度计算

首先要确定风机齿轮箱故障的分析周期。假定检验周期数为  $m$ ,故障模式对应的故障特征元数目为  $n$ ,则风电机组某故障模式的待检测多元故障特征矩阵经过标准化处理后可表示为

$$\begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1m} \\ x'_{21} & x'_{22} & \dots & x'_{2m} \\ \vdots & \vdots & & \vdots \\ x'_{n1} & x'_{n2} & \dots & x'_{nm} \end{bmatrix}$$

然后应用改进的算法对影响故障模式的各个故障特征元素进行隶属度划分,获得各个故障模式对应特征的量化隶属度表达式,得到隶属度序列,即  $\rho = \{b'_0, b'_1, b'_2, \dots, b'_n\}$ ,将  $\rho_i$  中的数值与特征值对应相乘得:

$$\begin{bmatrix} b'_1 x'_{11} & b'_2 x'_{12} & \dots & b'_n x'_{1m} \\ b'_1 x'_{21} & b'_2 x'_{22} & \dots & b'_n x'_{2m} \\ \vdots & \vdots & & \vdots \\ b'_1 x'_{n1} & b'_2 x'_{n2} & \dots & b'_n x'_{nm} \end{bmatrix}$$

### 3.2 齿轮箱故障特征值的确定

已知风电机组多元故障特征指标中的传动链振动数据为 6 种量纲一幅域参数,分别对风机齿轮箱正常工况和不同故障情况下的数据进行收集,选取某机组不同时间段内的历史数据。为保证计算结果随数值的波动性变化不大,对不同工况、不同阶段的故障特征值取均值,确定机组齿轮

箱故障特征值:

$$\mathbf{x}_{0p} = [b_1x_{11} \quad b_2x_{22} \quad \cdots \quad b_nx_{nn}]^T$$

$$\mathbf{x}_{ip} = [b'_1x'_{11} \quad b'_2x'_{22} \quad \cdots \quad b'_nx'_{nn}]^T$$

式中,  $\mathbf{x}_{0p}$  为融合隶属度标准参考样本的特征矩阵;  $\mathbf{x}_{ip}$  为考虑了特征隶属度的待测样本矩阵。

### 3.3 基于马氏距离的多元离群检测的故障模式确定

基于马氏距离的多元故障特征离群检测公式为

$$d_b^2(\mathbf{x}_{ip} - \mathbf{x}_{0p}) = (\mathbf{x}_{ip} - \mathbf{x}_{0p})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_{ip} - \mathbf{x}_{0p})$$

计算正常工况与齿轮箱故障样本特征集的马氏距离, 定义  $d_b^2$  为某故障模式下的边界距离, 若待检测数据对象的多元故障特征计算距离超过边界距离, 则表示待检测数据属于离群数据, 判断对应故障模式的发生。

## 4 案例分析

依旧以某风电场 1.5 MW 风电机组一级行星太阳轮磨损为例, 选取 2013 年 1 月份的 31 天为研究时间段, 以每天同一时间记录下来的 10 s 振动数据为分析样本, 来挖掘该时间段内机组齿轮故障发生过程中的多元故障特征的演变情况。其中, 参与分析的数据包括: 信号能量  $E$  (单位为 J)、翘度  $K_f$ 、波形裕度  $C_L$ 、偏态因子  $S_K$ 、重复性因子  $R_f$ 、相似性因子  $F_f$ 、跳跃性因子  $J_f$ 、温度异常率  $\varphi$ 。

齿轮箱故障发生初期, 机组信号数据故障特征并不明显, 各量纲一因子的值变化较为缓慢, 而且每个因子值的变化趋势因故障模式不同的会略有差异, 不易统一划定出机组故障的判断特征。故障发生时, 油温综合了各故障模式的机理特性, 即故障引起的齿轮箱零件间振动与摩擦程度加大, 导致油温发生改变。故将齿轮箱油温异常率作为进行诊断的阈值, 同时将各个量纲一因子值变化趋势作为故障模式特征, 实现风电机组齿轮箱的早期故障诊断, 这里将温度异常率的阈值设为 0.1。

通过对数据进行多元线性回归分析计算, 获得的风电机组一级太阳轮磨损故障与多元故障特征指标间的量化隶属关系式为

$$E = 19\,382.43 + 178.12K_f + 235.65C_L + 453.07S_K + 231.54R_f + 603.11F_f + 112.45J_f \quad (17)$$

对式(17)进行拟合优度评定、回归方程显著性检验(F 检验)和回归系数显著性检验(t 检验)。

计算拟合优度评定系数为

$$R^2 = S_{SR}/S_{ST} = 0.9640$$

式中,  $S_{SR}$  为总离差平方和;  $S_{ST}$  为回归平方和。

调整后的拟合优度判定系数为

$$R_a^2 = 1 - \frac{n-1}{n-p-1}(1-R^2) = 0.9550$$

$$n = 31 \quad p = 6$$

其中,  $R_a^2$  取值在 0 至 1 之间, 其值越接近于 1, 表明回归方程对实际观测值的拟合优度越高。

### 4.1 F 检验

$$H_0 : b'_i = 0, H_1 : b'_i \text{ 不全为零}$$

$$(i = 1, 2, \dots, 6)$$

由表 1 可知,  $F = 107.18$ , 显著水平  $\alpha = 0.05$ , 查表得  $F_{0.05}(6, 24) = 2.51$ , 明显小于  $F$ 。所以拒绝  $H_0$ , 接受  $H_1$ , 即故障特征隶属度值不全为零, 模型线性关系在 95% 置信水平下显著成立。

表 1 方差分析表

方差来源	平方和	自由度	均方和	F
回归 $S_{SR}$	551 215.17	6	$V_r = 91\,869.20$	107.18
误差 $S_{SE}$	20 571.25	24	$V_e = 857.12$	
总平方和	571 786.42	30		

### 4.2 t 检验

因为模型自由度为 6, 这里只对第一个自由度  $t_1$  进行假设检验:

$$H_0 : b_1 = 0; H_1 : b_1 \neq 0$$

给定一个显著性水平  $\alpha = 0.05$ , 查 t 分布表, 得到临界值  $T_a = 0.5000$ 。因为  $t_1 = 2.144 > T_a$ , 所以拒绝  $H_0$ , 接受备择假设, 即回归系统  $b_1 \neq 0$ 。对于其他回归系数  $b_i (i = 2, 3, \dots, 6)$ , 用上述同样方法可得出各回归系数是显著不为 0 的。

将同一台风机 2011 年 5 月 ~ 9 月的正常数据作为样本数据, 得出正常工况下 6 个量纲一因子阈值集合和隶属度集合:

$$\{K_f, C_L, S_K, R_f, F_f, J_f\} =$$

$$\{0.525, 5.211, 1.103, 0.017, 12.342, 5.178\}$$

$$\{b_1, b_2, b_3, b_4, b_5, b_6\} =$$

$$\{162.76, 213.77, 412.32, 221.21, 569.01, 100.99\}$$

常数项  $b_0 = 17\,942.76$ 。经过多元线性回归统计检验, 各项指标均符合要求, 线性方程线性显著成立。主轴轴承温度异常率故障的阈值为 0.1, 对异常率小于 0.1 的各个量纲一因子取均值, 得出标准故障特征集合。将多元故障特征量化隶属度代入多元故障特征离群检测距离边界公式, 求得距离边界  $d_b^2 = 1.543$ 。

对机组 1 月份 31 天数据进行马氏距离计算, 结果如表 2 所示。

表 2 1 月份故障数据与样本数据马氏距离

日期	马氏距离	日期	马氏距离	日期	马氏距离
1	1.454	12	1.614	23	1.641
2	1.433	13	1.621	24	1.614
3	1.523	14	1.606	25	1.629
4	1.511	15	1.611	26	1.644
5	1.544	16	1.634	27	1.651
6	1.579	17	1.628	28	1.659
7	1.602	18	1.599	29	1.667
8	1.586	19	1.641	30	1.887
9	1.594	20	1.614	31	1.889
10	1.598	21	1.635		
11	1.624	22	1.625		

由表 2 可得到,该机组 2013 年 1 月实际数据与标准样本数据的马氏距离从第 5 天开始,直到月末 26 天均属于离群点,因此可判断该机组从 1 月 5 日开始逐渐检测出一级太阳轮发生磨损故障。

最后,计算多元离群检测因子:

$$O_{DF} = W/A$$

式中, $W$  为待测对象的离群点数目; $A$  为所有待测数据总数。

从 5 日到 31 日,因子值由 1% 增长到 24%,说明机组故障程度正在不断加深。

应用传统的马氏距离的多元离群检测技术,其对应故障模式距离边界  $d_b^2 = 1.443$ 。对同一组实测数据进行计算,结果如表 3 所示。

表 3 传统方法计算马氏距离

日期	马氏距离	日期	马氏距离	日期	马氏距离
1	1.215	12	1.472	23	1.571
2	1.251	13	1.485	24	1.584
3	1.233	14	1.494	25	1.599
4	1.254	15	1.509	26	1.584
5	1.253	16	1.518	27	1.601
6	1.258	17	1.524	28	1.599
7	1.311	18	1.536	29	1.602
8	1.369	19	1.541	30	1.627
9	1.387	20	1.544	31	1.665
10	1.394	21	1.535		
11	1.469	22	1.525		

通过对比可以发现,传统算法在第 11 天发现机组齿轮箱出现一级太阳轮磨损故障,比改进方法晚了 5 天,说明改进方法能够更快速有效地识别风电机组齿轮箱早期的故障。

## 5 结论

(1) 应用软件算法实现了阶比重采样计算,将非平稳时域信号转化成平稳角域信号,不仅保证了后续计算结果的准确性,同时在实际工程应用中,节省了硬件费用。

(2) 利用量纲一因子对设备监测数据波动变化敏感的特点,将因子趋势变化作为风机齿轮箱早期故障的数据基础,为后续的故障模式识别打下坚实基础。

(3) 改进的多元离群检测算法对影响机组故障的各个特征参数进行了隶属度分析计算,使得基于距离的多元离群检测更有说服力,计算结果更为可靠,有很好的工程应用前景。

## 参考文献:

- [1] 刘德顺,戴巨川,胡燕平,等.现代大型风电机组现状与发展趋势[J].中国机械工程,2013,24(1):125-134.  
Liu Deshun, Dai Juchuan, Hu Yanping, et al. Status and Development Trends of Modern Large-scale Wind Turbines[J]. China Mechanical Engineering, 2013,24(1):125-134.
- [2] 张亮.风力发电机组齿轮箱早期故障诊断方法研究[D].大连:大连理工大学,2010.
- [3] 何邵灿,高宏力,许明恒.基于隐马尔科夫模型的机床部件故障预警技术[J].机械设计与制造,2012(8):159-161.  
He Shaocan, Gao Hongli, Xu Mingheng. Components of Machine Tools' Failure Warning Based on Hidden Markov Model[J]. Machinery Design and Manufacture, 2012(8):159-161.
- [4] 李蓉,于德介,陈向民,等.基于阶次分析与循环平稳解调的齿轮箱复合故障诊断方法[J].中国机械工程,2013,24(10):1320-1327.  
Li Rong, Yu Dejie, Chen Xiangmin, et al. A Compound Fault Diagnosis Method for Gearbox Based on Order Tracking and Cyclostationary Demodulation[J]. China Mechanical Engineering, 2013, 24(10):1320-1327.
- [5] 祁丽婉,梁庚,童国炜.基于果蝇算法优化 BP 神经网络的齿轮箱故障诊断[J].电网与清洁能源,2014,30(9):31-42.  
Qi Liwan, Liang Geng, Tong Guowei. A Gearbox Diagnosis Method Based on Fruit Fly Optimization Algorithm to Optimize the BP Neural Network[J]. Power System and Clean Energy, 2014,30(9):31-42.
- [6] 马怀祥,徐明新,金珏.应用油液分析法和振动法诊断往复机械故障[J].石家庄铁道学院学报,1998,11(2):39-43.  
Ma Huaixiang, Xu Mingxin, Jin Jue. Breakdown Diagnosis of Reciprocating Machinery by Means of Oil Analysis and Vibration Analysis[J]. Journal of Shijiazhuang Railway Institute, 1998,11(2):39-43.

(下转第 1916 页)

Learning and One-class SVM[J]. China Mechanical Engineering, 2013, 24(5): 628-633.

- [9] 欧璐, 于德介. 基于拉普拉斯分值和模糊 c 均值聚类的滚动轴承故障诊断[J]. 中国机械工程, 2014, 25(10): 1352-1357.
- Ou Lu, Yu Dejie. Rolling Bearing Fault Diagnosis-Based on Laplacian Score and Fuzzy C-means Clustering[J]. China Mechanical Engineering, 2014, 25(10): 1352-1357.
- [10] 苏祖强, 汤宝平, 刘自然, 等. 基于正交半监督局部 fisher 判别分析的故障诊断[J]. 机械工程学报, 2014, 50(18): 7-13.
- Su Zuqiang, Tang Baoping, Liu Ziran. Fault Diagnosis Method Based on Orthogonal Semi-supervised Local Fisher Discriminant Analysis[J]. Journal of Mechanical Engineering, 2014, 50(18): 7-13.
- [11] Su Zuqiang, Tang Baoping, Ma Jinghua, et al. Fault Diagnosis Method Based on Incremental Enhanced Supervised Locally Linear Embedding and Adaptive Nearest Neighbor Classifier[J]. Measurement, 2014, 48: 136-148.
- [12] Wang Yi, Xu Guanghua, Liang Lin, et al. Detection of Weak Transient Signals Based on Wavelet Packet Transform and Manifold Learning for Rolling Element Bearing Fault Diagnosis[J]. Mechanical Systems and Signal Processing, 2015, 54: 259-276.
- [13] Belkin M, Niyogi P. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation [J]. Neural Computation, 2003, 15(6): 1373-1396.
- [14] Belkin M, Niyogi P. Semi-supervised Learning on Riemannian Manifolds [J]. Machine Learning, 2004, 56(1/3): 209-239.
- [15] Kouropteva O, Okun O, Hadid A, et al. Beyond Locally Linear Embedding Algorithm, MVG-01-2002[R]. Oulu, Finland: University of Oulu, 2002.
- [16] He Qingbo. Time - frequency Manifold for Non-linear Feature Extraction in Machinery Fault Diagnosis[J]. Mechanical Systems and Signal Processing, 2013, 35(1/2): 200-218.
- [17] Loparo K A. Bearings Vibration Data Set [DB/OL]. Cleveland, Ohio: Case Western Reserve University. [2014-10-28]. <http://Csegroups.Case.Edu/Bearingdatacenter/Home>.
- [18] Vibroacoustic Gear Signatures with Time-Frequency Spectrograms[DB/OL]. [2015-05-09]. <http://www.ulb.ac.be/polytech/laborulb/gearvi/gearjtfa.htm>.

(编辑 张 洋)

作者简介: 江 丽, 女, 1980 年生. 武汉理工大学机电工程学院助理研究员、博士. 主要研究方向为机械状态监测与故障诊断、模式识别. 发表论文 3 篇. 郭顺生(通信作者), 男, 1963 年生. 武汉理工大学机电工程学院教授、博士研究生导师.

(上接第 1910 页)

- [7] 郭厚明, 行志刚. 量纲一参数在矿用低速重载齿轮故障诊断中的应用[J]. 煤炭科学技术, 2006, 34(8): 28-31.
- Guo Houming, Xing Zhigang. Dimensionless Parameters Applied to Fault Diagnosis of Mine Low Speed Heavy Loaded Gear [J]. Coal Science and Technology, 2006, 34(8): 28-31.
- [8] 岑少起, 于卫东, 李瑞珍. 动静压滑动轴承基本方程和特性系数量纲一化研究[J]. 机械传动, 2008, 32(3): 4-7.
- Cen Shaoqi, Yu Weidong, Li Ruizhen. The Study of Dimensionless Basic Equation and Characteristics Coefficients of Hydrodynamic Hydrostatic Bearings [J]. Journal of Mechanical Transmission, 2008, 32(3): 4-7.
- [9] Hochman D, Sadok M. Theory of Synchronous Averaging[C]//Proceedings of the 2004 IEEE Aerospace Conference. Big Sky, Montana, USA, 2004: 3636-3653.
- [10] Wonlen K, Manson G, Fieller N J. Damage Detection Using Outlier Analysis[J]. Journal of Sound and Vibration, 2000, 229(3): 647-667.
- [11] 扈静, 刘明周, 龚任波, 等. 基于主成分分析的汽车操纵力舒适性分析与评价[J]. 中国机械工程, 2011, 22(20): 2456-2459.
- Hu Jing, Liu Mingzhou, Gong Renbo, et al. Evaluation of Vehicle Operating Force Comfort Based on Principal Component Analysis[J]. China Mechanical Engineering, 2011, 22(20): 2456-2459.
- [12] 申志刚, 何宁, 李亮. 高速硬铣削加工刀具磨损监测研究[J]. 中国机械工程, 2009, 20(13): 1582-1586.
- Shen Zhigang, He Ning, Li Liang. Monitoring of Tool Wear in Hard Milling Process[J]. China Mechanical Engineering, 2009, 20(13): 1582-1586.

(编辑 张 洋)

作者简介: 顾煜焱, 男, 1968 年生. 华北电力大学能源动力与机械工程学院教授、博士研究生导师. 主要研究方向为电站大型旋转机械振动监测与故障诊断. 发表论文 80 余篇. 贾子文, 男, 1986 年生. 华北电力大学能源动力与机械工程学院博士研究生. 王 瑞, 男, 1989 年生. 华北电力大学能源动力与机械工程学院硕士研究生. 任玉亭, 男, 1966 年生. 国华能源投资有限公司副总工程师.