

# 基于负荷轴工况划分的发电厂关键设备非平稳状态监测

沙万里<sup>1</sup>, 陈军豪<sup>2</sup>, 赵春晖<sup>2</sup>

(1. 浙江浙能嘉华发电有限公司, 浙江 嘉兴 314201;

2. 浙江大学控制科学与工程学院, 杭州 310027)

**摘要:** 在发电厂热力系统实际运行过程中, 发电机负荷的频繁变化会使得热力系统关键设备部分运行参数随时间发生改变, 存在明显的非平稳特性, 由此导致设备的变工况问题非常普遍, 故障特征很容易被非平稳趋势所掩盖, 给及时检测设备异常带来较大困难。针对该问题, 对步进有序时段划分算法进行推广, 提出了一种基于负荷轴的工况划分算法, 准确自动划分火电关键设备负荷信号, 从而识别出不同的工况。基于划分出的不同工况, 分别建立主元分析监测模型, 在应用时根据新来样本的负荷调用相应的监测模型, 从而实现对非平稳过程的状态监测。所提出的划分方法以监测模型的性能优劣来指示工况划分, 相比一般基于聚类算法的工况划分方法更具有合理性。该算法的有效性已在某发电厂一次风机的故障案例得到验证。

**关键词:** 非平稳过程; 工况划分; 状态监测; 火力发电

文章编号: 1007-1881(2019)12-0031-08

DOI: 10.19585/j.zjdl.201912006

中图分类号: TK227

文献标志码: B

开放科学(资源服务)标识码(OSID):



## Nonstationary Operation Condition Monitoring for Key Machines of Power Plant Based on Load Axis Operating Condition Division

SHA Wanli<sup>1</sup>, CHEN Junhao<sup>2</sup>, ZHAO Chunhui<sup>2</sup>

(1. Zhejiang Zheneng Jiahua Power Generation Co., Ltd., Jiaxing Zhejiang 314201, China;

2. College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China)

**Abstract:** During the operation of the thermodynamic system of a power plant, the frequent load change causes variation of operating parameters of key equipment and obvious nonstationary operating characteristics. As a result, the variable working conditions of the equipment are very common, and the fault characteristics are easily covered by nonstationary trends, which brings great difficulties to the timely detection of abnormalities of the equipment. To solve the problem, this paper generalizes the stepwise sequential phase partition (SSPP) algorithm and proposes a load axis based operating condition division algorithm to automatically divide the thermal key equipment load signals to identify different working conditions. Under the divided working conditions, the principal component analysis monitoring models are established respectively. In the application, the corresponding monitoring model is called according to the load of the new sample, thereby realizing the state monitoring of the nonstationary process. The division method proposed in this paper uses the performance of the monitoring model to indicate the division of working conditions, which is more reasonable than the general clustering method. Finally, the effectiveness of the proposed algorithm is verified in a failure case of a primary fan in a power plant.

**Keywords:** nonstationary process; condition division; condition monitoring; thermal power generation

## 0 引言

在我国的电源结构中, 火电一直是主力电

源。截至 2018 年的统计数据, 火电发电量占全国发电量的比例超过 60%, 其中燃煤发电是火电的主体, 因此火电机组的安全稳定运行对保障我国电力供应至关重要。热力系统作为火电机组的重要组成部分, 是一个极其复杂的工业系统, 整个工业过程生产流程长、单元装置多、空间分布广、

基金项目: NSFC-浙江两化融合联合基金(U1709211), 浙江省重点研发项目(2019C01048)

安全要求高,其中可能会发生的故障复杂多样。对热力系统关键设备进行状态监测能够帮助发电厂充分了解设备运行状况、及时进行故障排查与检修、减少停机检修的时间,从而提升生产效率、防止生产事故的发生。因此,迫切需要对热力系统的关键设备实施有效的状态监测。

在火电热力系统实际运行过程中,负荷的频繁变化会使得热力系统关键设备部分运行参数的均值和方差随时间发生改变,呈现出明显的非平稳特性,由此导致机组的变工况问题非常普遍。频繁出现稳态工况与过渡工况交替的现象,给运行异常的及时监测带来了较大困难。一方面,故障特征很容易被非平稳趋势所掩盖。另一方面,传统的状态监测方法不能准确地描述非平稳变量间的关系,由此可能导致错误的监测结果。此外,由于变工况现象,尤其是过渡工况的存在,实际火电热力系统运行过程的状态监测相比单一稳态工况下的状态监测难度更大要求更高,存在很多问题和挑战需要作深入研究。

截至目前,已有的状态监测方法可以归为2类:基于模型的<sup>[1-2]</sup>和数据驱动的<sup>[3-5]</sup>。其中,数据驱动的状态监测方法不依赖过程知识,能够通过分析过程数据、提取其中所包含的过程信息来描述过程的运行状态,从而实现有效的过程监测。因此,数据驱动的方法被广泛运用于解决火电热力系统状态监测问题<sup>[4,6-10]</sup>。针对非平稳过程的监测方法,较为常用的有2种:基于协整分析的方法<sup>[10-13]</sup>和基于工况划分的方法<sup>[6-7]</sup>。

Chen 等人<sup>[11]</sup>首先将协整分析方法运用于一个工业蒸馏单元的状态监测。Li 等人<sup>[12]</sup>使用协整分析导出了多个残差变量,并计算了 T2 统计量的值来检测异常行为。Zhang 等人<sup>[10]</sup>提出了一种基于协整分析和主元分析的两级非平稳过程状态监测方法。在下级模型中,首先区分平稳变量和非平稳变量,使用主元分析和协整分析分别对平稳变量和非平稳变量进行监测。为了对平稳变量和非平稳变量间的关系进行监测,文章利用协整分析得到的平稳残差序列和平稳变量一起构成新的数据矩阵,在上级模型中使用主元分析来进行监测。但是,协整分析仅仅能够应用于具有协整关系的非平稳变量的监测,对不具备协整关系的非平稳变量的监测无能为力。

黄孝彬等<sup>[6]</sup>在分析火电机组负荷变化特点后提出,火电机组的典型运行状况可以归纳为大部分时间是处于稳定工作状态,而不同稳定状态是连续变化和过渡的,进而提出了一种累计递推的 PCA 建模方法,即“不断的积累数据并用递推的方式估计变量的统计特性:方差和相关系数矩阵。当递推的结果表现为变量统计参数的估计值趋于稳定时,此时可以认为用这些数据建模就能抓住过程的变动特性。”<sup>[6]</sup>文章将该方法应用于锅炉过程传感器的故障检测,取得了较好的效果。华北电力大学的牛征等人<sup>[7]</sup>针对火电厂变工况的特点,提出了另外一种工况划分思路:首先用 K-means 聚类方法将数据进行聚类,认为得到的各个类即为不同稳态下的数据集;然后对各个数据集分别建立主元模型;最后将待监测样本进行模糊划分,提出了一种模糊动态主元模型的方法计算其适用的主元模型,并在锅炉数据上验证了该方法的有效性。此外,许仙珍等<sup>[14]</sup>利用混合高斯模型来对多工况过程进行工况划分,然后对每个工况建立主元分析模型,实现了对多工况过程的状态监测。以上方法都利用了聚类的方法来进行工况划分。由于聚类时会利用多个变量而不仅仅是负荷这一变量,因此,划分的工况内很可能存在着负荷差距很大的样本,而这明显是不符合实际情况的;其次,聚类算法的聚类标准并不是针对状态监测这一问题提出的,因此,使用一般聚类算法来进行工况划分很有可能对后续的状态监测带来不利影响。

为解决上述方法的不足,本文提出了一种基于负荷轴工况划分算法的非平稳状态监测方法,来解决火电关键设备的非平稳状态监测问题。首先,考虑到同一工况下负荷必然是相近的,为避免将负荷差别较大的样本划入同一工况而负荷差别较小的样本划入不同工况,本文所提的方法将时间轴变为负荷轴,即,将初始时以时间顺序排列的样本集按负荷从小到大重新构造。其次,对于过程监测来说,以监测模型的性能优劣来指示工况划分更为合理,因此,提出基于负荷轴的工况划分算法来进行工况划分。划分完毕后,所得到的各个工况内的数据可近似认为是平稳的,且具有相似的特性。因此,可以利用主元分析方法在每个工况内分别建立状态监测模型。最后,根

据新来样本的负荷调用相应的监测模型,进行在线的状态监测。

## 1 基于负荷轴工况划分的非平稳状态监测方法

本文提出的基于负荷轴工况划分的非平稳状态监测方法,包括4个部分:构造负荷片以进行工况划分;基于负荷片的自动有序工况划分;基于主元分析的多工况建模方法;在线监测方法。

### 1.1 负荷片构造

考虑火电热力系统某设备数据集  $X(N \times J \times P)$ , 其中  $N$  表示样本数,  $J$  表示变量数,  $P$  表示负荷。在每一个负荷  $p$  处,可以得到一个二维的数据矩阵  $X_p(n_p \times J)$ , 称之为负荷片,作为基本数据分析单元,它是由所有负荷等于  $p$  的样本点组成的矩阵。在每个负荷片内,可以认为过程特性未发生显著变化,通过分析每个负荷片的数据特性,可以得到过程在每一个负荷处的内在运行状态以及过程状态随负荷的变化情况。

在实际应用时,无法对负荷进行无限细分,因此使用一个足够小的负荷区间  $[p, p+\Delta p]$  内的样本集合来代替负荷  $p$  处的样本集合作为基本数据分析单元是合理的。而由于数据量的限制,在每一个负荷区间  $[p, p+\Delta p]$  内得到的二维数据矩阵很可能存在样本数量不足的问题,不足以揭示同一负荷区间下的统计波动特性,因此,还需要对构造的基本数据分析单元进行一些调整。将样本数较少的基本数据分析单元与其相邻负荷区间的基本数据分析单元进行合并。如果合并过多的基本数据分析单元,会使得负荷方向上的波动信息过大而导致同一负荷区间下不同样本间的波动特性被掩盖。一般来说,为了提供可靠的数据标准化信息,通常基本数据分析单元内的样本数要等于过程变量个数的2~3倍<sup>[15]</sup>。在每个调整后的基本数据单元中,可以近似认为过程特性未发生显著变化,能够提供较为可靠的数据标准化信息。

### 1.2 基于负荷片的自动有序工况划分

考虑到相同的工况内,负荷相近、过程特性变化较小、变量间的相关关系近似相等,因此,利用变量间相关关系的相似程度来对调整后的负荷片进行合并,将过程特性相似的相邻负荷片合并,得到不同工况,分析每个工况的过程特性,

分别建立状态监测模型,可以提高对监测过程的理解及监测结果的准确性。

针对间歇过程的多时段特性,Zhao等<sup>[16]</sup>提出了一种SSPP(自动步进有序时段划分)方法,通过不断比较时间片模型和由时间片依次构成的时间块模型的SPE指标来确定时段划分点。由于SPE指标可以灵敏地反映变量相关性并且直接和监测性能相关,该方法有效改善了监测性能。李文卿<sup>[15]</sup>针对小样本情况,将SSPP算法进行了推广,提出了一种GSSPP(基于泛化时间片的步进有序时段划分)算法,通过衡量泛化时间片中的数据特性对监测模型性能的影响进行时段识别。

对于过程监测来说,应当以监测模型的性能来指示工况划分。因此,借鉴上述GSSPP算法划分间歇过程时段的思想,提出一种基于负荷轴的有序工况划分算法,以火电机组热力系统关键设备的负荷变量来指示工况划分,从而识别出不同的工况。基于负荷轴的有序工况划分算法的具体步骤如下:

(1)步骤1:构造负荷片。以一个较小的  $\Delta p$  为间隔,划分负荷区间为  $[P_{\min}+(i-1)\Delta p, P_{\min}+i\Delta p]$ ,  $i=1, 2, \dots, \text{floor} \frac{P_{\max}-P_{\min}}{\Delta p}$ 。负荷片记为  $X_k, k=1, 2, \dots, \text{floor} \frac{P_{\max}-P_{\min}}{\Delta p}$ , 每个负荷片由负荷处于相应负荷区间的所有样本构成。

(2)步骤2:对负荷片进行合并调整,使得每个负荷片内样本数为过程变量个数的2~3倍。将调整后的负荷片记为  $X_k, k=1, 2, \dots, K$ ,并在每个负荷片内对其进行标准化,记为  $\bar{X}_k(n_k \times J) (k=1, 2, \dots, K)$ 。

(3)步骤3:对每个负荷片进行PCA建模。将步骤2中经过标准化处理的每一个负荷片矩阵  $\bar{X}_k(n_k \times J) (k=1, 2, \dots, K)$  进行PCA分解,从而建立PCA模型:

$$\bar{X}_k = T_k P_k^T + E_k, k=1, 2, \dots, K \quad (1)$$

式中:  $T_k(n_k \times R_k)$  与  $P_k(J \times R_k)$  代表保留  $R_k$  个主元变量的得分矩阵和负载矩阵;  $E_k$  为残差矩阵。这里,保留的主元个数按以下方法确定:计算每一个负荷片PCA模型累计方差达到99%的主元个数,选取其中出现次数最多的主元个数作为所有负荷片PCA模型的主元个数,记为  $R$ ,即  $R_k=R$

( $k=1, 2, \dots, K$ )。

(4)步骤4: 负荷片 PCA 模型 SPE 统计量控制限确定。对每一个负荷片 PCA 模型, 计算其 SPE 统计量控制限。计算公式如下:

$$Ctrl_k = \delta_{\alpha}^2 = g \chi_{h, \alpha}^2, \quad k=1, 2, \dots, K, \quad (2)$$

式中:  $g = \frac{\theta_2}{\theta_1}$ ;  $h = \frac{\theta_1^2}{\theta_2^2}$ ;  $\chi_{h, \alpha}^2$  是自由度为  $h$ 、置信度为

$\alpha$  的  $\chi_{h, \alpha}^2$  分布临界值;  $\theta_i = \sum_{j=R_i+1}^J \lambda_j^i$  ( $i=1, 2, 3$ );  $\lambda_j$  为负荷片  $X_k$  的协方差矩阵  $S$  的特征值。

(5)步骤5: 负荷段 PCA 建模。从第一个负荷片开始, 依次将下一个负荷片与之前的负荷片组合在一起, 得到负荷段数据矩阵  $X_{i,j} = [\bar{X}_i, \bar{X}_{i+1}, \dots, \bar{X}_{ij}]$ , 其中  $i$  从 1 开始,  $j$  也从 1 开始。对负荷段数据矩阵  $X_{i,j}$  进行 PCA 分解, 获得负载矩阵  $P_{i,j} (J \times R)$ 。使用负荷段 PCA 模型负载矩阵  $P_{i,j} (J \times R)$  对负荷段内每一个负荷片  $\bar{X}_i, \bar{X}_{i+1}, \dots, \bar{X}_{ij}$  进行分解, 并在每一个负荷片内, 计算 SPE 统计量及其控制限  $Ctrl'_i, Ctrl'_{i+1}, \dots, Ctrl'_{ij}$ 。这代表了负荷段 PCA 模型对其区间内的每一个负荷片的重建能力。

(6)步骤6: 比较负荷段 PCA 模型中每一个负荷片的 SPE 控制限  $Ctrl'_i, Ctrl'_{i+1}, \dots, Ctrl'_{ij}$  与每一个负荷片单独建立的 PCA 模型的 SPE 控制限  $Ctrl_i, Ctrl_{i+1}, \dots, Ctrl_{ij}$ 。找到连续 3 个负荷片呈现出  $Ctrl'_k > \alpha Ctrl_k$  时的  $j$ , 这就意味着新加入的第  $j$  个负荷片对负荷段的 PCA 模型及其监测性能有重大影响, 导致负荷段 PCA 模型的准确性相比负荷片 PCA 模型有显著降低, 也就是说, 只能将第  $i$  个至第  $i+j-1$  个负荷片划分在一起, 这样便划分出了一个工况。 $\alpha$  是一个被称为宽松因子的常数, 反映了与负荷片 PCA 模型相比, 允许负荷段 PCA 模型监测准确度损失的程度。

(7)步骤7: 更新过程分析数据, 迭代完成所有工况的划分。移除已划分好的工况所包含的负荷片, 将剩余负荷片作为步骤 5 的新的输入, 再次进行步骤 5 和步骤 6, 划分出下一个工况。迭代进行, 直至没有负荷片剩余。

通过以上步骤, 可以将不同工况所对应的负荷区间划分出来。接下来, 就可以在每一个工况内分别建立一个监测模型, 根据新样本的负荷值所处区间来调用相应的监测模型, 实现过程监测。

需要注意的是, 宽松因子  $\alpha$  的取值将会直接

影响工况的划分结果。 $\alpha$  越大, 意味着负荷段 PCA 模型所允许的准确度损失越大, 那么每个负荷段将包含更多的负荷片, 从而会获得更长的负荷段, 即划分出的工况数会更少; 而  $\alpha$  越小, 意味着负荷段 PCA 模型所允许的准确度损失越小, 监测模型对每个负荷片的表征就越精确, 从而得到的是更短的负荷段, 即划分出的工况数会更多。因此,  $\alpha$  反映的是模型准确度和模型复杂度的一种折中。一般来说,  $\alpha$  的值是通过比较验证集上的监测效果来确定的,  $\alpha$  的取值要使得每个工况包含合适的过程运行模式, 保证对过程特征变化的敏感性。

### 1.3 基于主元分析的多工况建模方法

在工况划分完毕之后, 将每个工况内的负荷片整合成为一个数据矩阵, 记为  $Y_i (N_i \times J)$ ,  $i=1, 2, \dots, M$ , 其中,  $M$  为划分出的工况个数。在每个工况内建立 PCA 模型, 计算出  $T^2$  和 SPE 统计量控制限。 $Y_i$  的每一行代表一个样本点, 每一列代表一个变量序列。具体步骤如下:

(1)步骤1: 使用 SVD (奇异值分解) 的方法建立主元分析模型:

$$\begin{cases} T_i = \bar{Y}_i P_i \\ \bar{Y}_i = T_i P_i^T + E_i \end{cases}, \quad (3)$$

式中:  $P_i$  是负载矩阵;  $T_i$  是相应的得分矩阵;  $E_i$  是提取主元后留下的残差矩阵。 $R_i$  表示保留的主成分个数, 由累计方差贡献率准则来确定。

(2)步骤2: 计算  $T^2$  和 SPE 统计量控制限。 $T_{\alpha, i}^2$  表示第  $i$  个工况下置信度为  $\alpha$  的  $T^2$  控制限;  $\delta_{\alpha, i}^2$  表示第  $i$  个工况下置信度为  $\alpha$  的 SPE 控制限。当  $T^2$  和 SPE 位于控制限内时, 认为过程是正常的, 如果超出控制限, 则表明有故障发生。2 个统计量的控制限计算公式如下:

$$\begin{cases} T_{\alpha, i}^2 = \frac{R_i(N_i^2-1)}{N_i(N_i-R_i)} F_{R_i, N_i-R_i, \alpha} \\ \delta_{\alpha, i}^2 = \theta_1 \left( \frac{c_{\alpha} \sqrt{2\theta_2 h_0^2}}{\theta} + 1 + \frac{\theta_2 h_0 (h_0-1)}{\theta_1^2} \right)^{\frac{1}{h_0}}, \end{cases} \quad (4)$$

式中:  $F_{R_i, N_i-R_i, \alpha}$  是带有  $R_i$  和  $N_i-R_i$  个自由度、置信水平为  $\alpha$  的  $F$  分布临界值;  $\theta_i = \sum_{j=R_i+1}^J \lambda_j^i$  ( $i=1, 2, 3$ );

$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$ ,  $\lambda_j$  为  $\bar{Y}_i$  的协方差矩阵  $S$  的特征值,

$c_{\alpha}$  为标准正态分布在置信水平  $\alpha$  下的阈值;  $R_i$  为

所建立的 PCA 模型保留的主元个数;  $J$  为  $\bar{Y}_i$  所含有的变量个数;  $N$  为当前工况下的样本数。

(3)步骤 3: 在实际应用中,使用 1 个指标比 2 个指标更为方便,因此,计算合成指标控制限。合成指标将 SPE 和  $T^2$  用以下方式进行了融合<sup>[17]</sup>:

$$\varphi(x) = \frac{\text{SPE}(x)}{\delta_\alpha} + \frac{T^2(x)}{\delta_\alpha} = x^T \Phi x \leq \zeta^2 \quad (5)$$

式中:  $\Phi = \frac{P_i \Lambda_i^{-1} P_i^T}{T_\alpha^2} + \frac{I - P_i P_i^T}{\delta_\alpha^2}$ , 需注意  $\Phi$  是对称且正定的。

利用近似分布可以计算这个混合指标的控制限  $\zeta_i^2$ :

$$\zeta_i^2 = g x_i^2 \quad (6)$$

式中:  $g = \frac{\text{tr}(S\Phi)^2}{\text{tr}(S\Phi)}$ ;  $h = \frac{[\text{tr}(S\Phi)]^2}{\text{tr}(S\Phi)^2}$ ;  $S$  为  $\bar{Y}_i$  的协方差矩阵。如果  $\Phi > \zeta^2$ , 则认为过程发生了故障。

### 1.4 在线监测方法

在上述步骤中,根据过程特性建立了各个工况下的状态监测模型。在线应用时,首先根据新来样本  $y_i (J \times 1)$  包含的负荷值确定其所处的工况,然后选取相应的监测模型来进行监测。具体步骤表述如下:

(1)步骤 1: 首先,根据负荷值确定其所属的负荷片,假设其处于第  $j$  个负荷片;其次,根据工况划分结果,确定  $y_i$  所处的工况,假设  $y_i$  处于第  $i$  个工况。

(2)步骤 2: 根据第  $j$  个负荷片的离线数据  $X_j$  的均值和方差对  $y_i$  进行标准化,然后,由第  $i$  个 PCA 模型来进行监测。在线监测统计量定义为:

$$\begin{cases} T_i^2 = t_i^T \Lambda_i^{-1} t_i \\ \text{SPE}_i = e_i^T e_i \\ \Phi_i = \frac{\text{SPE}_i}{\delta_{\alpha,i}^2} + \frac{T_i^2}{T_{\alpha,i}^2} \end{cases} \quad (7)$$

式中:  $t_i = P_i^T y_i$  是样本  $y_i$  的主成分;  $P_i$  和  $\Lambda_i$  是 2.3 节中建立模型时,使用正常运行状态下采集的第  $i$  个工况下的历史数据所计算得到的负载矩阵和主成分协方差矩阵;  $e_i = y_i - t_i$  是主元解释后的残差;  $\delta_{\alpha,i}^2$  和  $T_{\alpha,i}^2$  分别为 2.3 节建立模型时计算得出的第  $i$  个工况下的 SPE 和  $T^2$  统计量的控制限。

当统计量在控制限以下时,样本被认为是正常的。否则说明样本存在异常,检测到了故障。

## 2 实验分析

### 2.1 数据说明

一次风是指煤粉燃烧时与煤粉一起送入炉膛的空气,对锅炉的燃烧起主导作用,一次风机是锅炉的重要辅机之一。本文数据集来源于浙江省某集团下属发电厂,选取了一次风机正常运行时采集的 12 364 个样本,相邻样本间的时间间隔为 3 min,其中 70% 作为训练集,用来建立监测模型,30% 作为验证集用来确定  $\alpha$  的取值。每个样本含有 36 个过程变量,简单起见,在表 1 中只列出前 14 个变量。训练集记为  $X_{\text{train}} (655 \times 36)$ ,验证集记为  $X_{\text{verify}} (3 709 \times 36)$ 。此外,选取了一次风机正常运行时另外采集的 900 个样本作为测试集 1,相邻样本间的时间间隔为 3 min,记为  $X_{\text{test1}} (900 \times 36)$ ;选取一次风机在故障发生前采集的 500 个样本作为测试集 2,相邻样本间的时间间隔为 3 min,记为  $X_{\text{test2}} (500 \times 36)$ 。此处,一次风机发生的故障为从第 100 个样本点处开始电机轴承温度升高。

表 1 一次风机部分变量

编号	变量名	单位
1	环境温度	℃
2	出口风压	kPa
3	电机定子线圈温度	℃
4	电机轴承水平振动	mm/s
5	电机轴承温度	℃
6	电流	A
7	动叶位置反馈	%
8	风机轴承垂直振动	mm/s
9	风机轴承水平振动	mm/s
10	润滑油压力	MPa
11	风机轴承温度	℃
12	出口风温	℃
13	动叶开度指令	%
14	总风量	t/h

### 2.2 建模过程

使用训练集  $X_{\text{train}} (8 655 \times 36)$  来建立监测模型的第一步便是划分工况。根据过程机理知识,对于一次风机来说,可以使用出口风压来作为区分其不同工况的划分变量,但根据本文所提的算法,计算监测统计量时会排除划分变量,为保留尽可能多的信息,本文考虑构造一个与出口风压十分

相关的变量代替出口风压作为划分变量,以使得计算监测统计量时出口风压也被考虑在内,分析训练数据可知风压开方后的值与风压的相关系数达到了0.99,因此,这里将出口风压开方后的值,作为划分变量对工况进行划分,记为 $\text{sqrt}(\text{风压})$ 。

训练集  $X_{\text{train}}(8\ 655 \times 36)$  包含了足够多的运行工况,其中  $\text{sqrt}(\text{风压})$  和一次风机电流随时间变化如图1所示。由图1可知,  $\text{sqrt}(\text{风压})$  随时间波动相当频繁,存在明显的非平稳特性,经检验亦为非平稳变量。相应地,过程变量例如一次风机电流也随之产生变化。

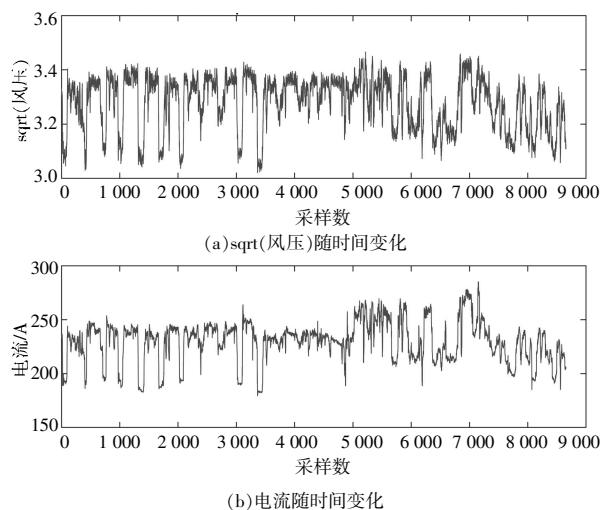


图1 正常运行时  $\text{sqrt}(\text{风压})$  及一次风机电流随时间变化

为进行工况划分,首先,将  $\text{sqrt}(\text{风压})$  最小值至最大值区间等分为了40个小区间,形成40个基础数据矩阵。考虑到每个基础数据矩阵内的样本数量需要为变量数的2~3倍才能较好地使用PCA方法提取到数据中的信息,因此,在这里确定每个基础数据矩阵内的最小样本数为80。之后,需要将不满足最小样本数量要求的基础数据矩阵进行合并,最终得到满足要求的基础数据矩阵共34个。对34个基础数据矩阵使用基于负荷轴的有序工况划分算法进行工况划分。以验证集的监测结果来确定 $\alpha$ 的取值,表2是 $\alpha$ 取不同值时,验证集上的FAR。

选取使得验证集FAR最小的作为 $\alpha$ 的值,综合确定 $\alpha=1.3$ 较为合适。 $\alpha=1.3$ 时,工况划分结果如图2所示。

表2 不同 $\alpha$ 取值时统计量的FAR及划分得到的工况数1

$\alpha$	$T_2$ 统计量 FAR/%	SPE 统计量 FAR/%	混合统计量 FAR/%	划分得到的 工况数/个
1.1	1.78	1.05	1.48	7
1.15	1.81	1.00	1.48	6
1.3	1.59	0.94	1.35	5
1.35	1.54	0.97	1.54	4
1.5	1.54	1.35	1.46	3
2	1.59	1.24	1.27	2
3	1.32	1.38	1.32	1

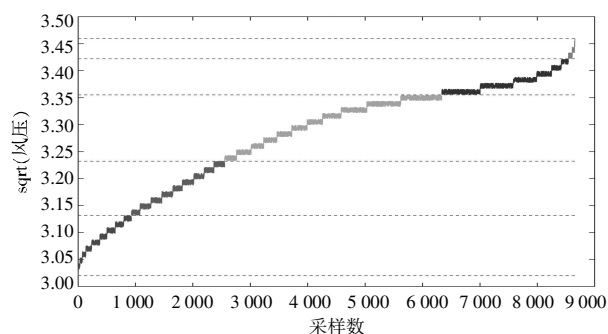


图2 工况划分结果

至此,得到了划分出的5个工况。根据工况划分结果,在每个工况内,分别建立1个PCA监测模型,最终得到5个子模型。

### 2.3 监测结果

建模完成后,在正常测试集1,  $X_{\text{test1}}(900 \times 36)$  上进行监测,查看监测效果。对处在不同工况区间内的样本调用不同的监测模型,最终得到的监测结果如图3所示。

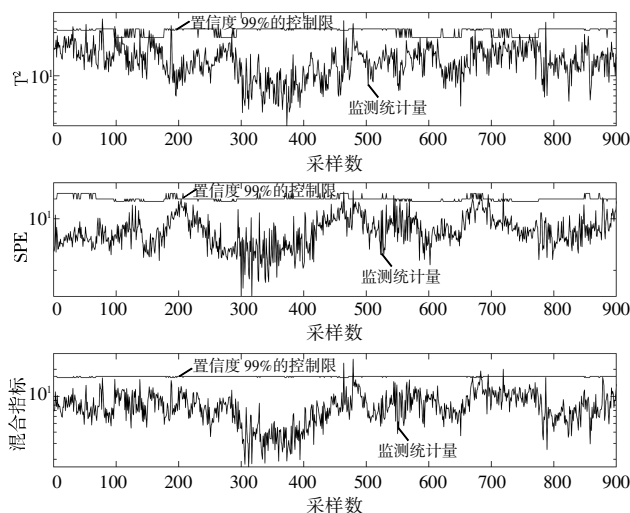


图3 测试集1监测结果

由图3可知,无论是 $T^2$ ,SPE还是两者的混合指标,只在个别样本点处略微超出控制限,且超限时间很短,不存在较长时间连续超限的现象。经统计,在 $T^2$ ,SPE以及两者的混合指标下,FAR分别为0.022 2,0.010 0和0.006 7。

在异常测试集2, $X_{test2}(500 \times 35)$ 上进行监测,查看监测效果。对不同工况下的样本应用不同的监测模型,最终得到的监测结果如图4所示。

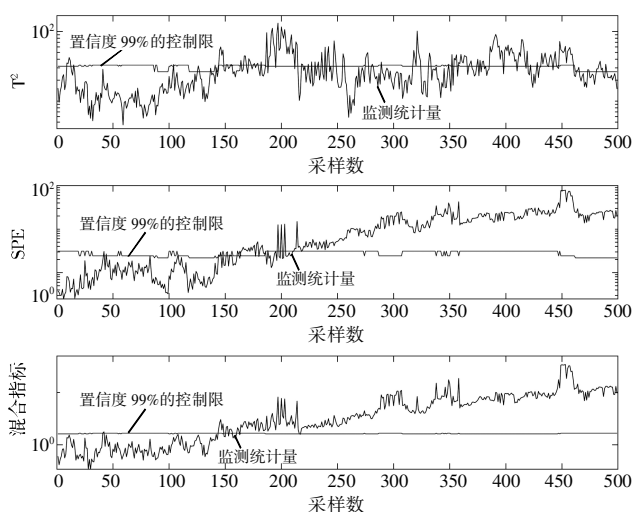


图4 测试集2监测结果

由图4可知,从第144个样本点开始,混合指标超出控制限,意味着检测到了故障,相比故障真正的发生时间延后44个样本点(132 min)。

为验证本文所提方法的有效性,与以下2种方法进行对比:第一种方法是使用高斯混合模型聚类后在各个类内使用PCA进行监测;第二种方法是不进行工况划分,将全部标准化后的负荷片合并后使用PCA进行监测。以正常测试集的FAR和异常测试集的检测时延作为评价指标,结果如表3所示。由结果可知,本文所提方法相对于基于一般聚类算法的监测方法,在保证较低FAR的同时,检测时延显著减少,能够更加灵敏地检测出故障。同时,与第二种方法的对比可以说明进行工况划分的必要性。

表3 3种方法结果比较

	测试集1 FAR	测试集2 FAR
本文所提方法	0.67%	44
Gmm+PCA	0.33%	93
直接合并+PCA	1.56%	44

### 3 结语

本文针对火电热力系统关键设备的非平稳状态监测问题提出了一种基于负荷轴工况划分方法的监测策略。相比于基于协整分析的非平稳过程状态监测方法,该方法解决了不存在协整关系的非平稳变量的监测问题。相比于传统的以聚类为基础的工况划分方法,提出的方法考虑了同一工况下负荷相近、运行特性相似的特点,将时间轴转为负荷轴,构造了负荷片,并以监测模型的性能优劣来指示工况划分,更具有合理性。在每个划分出的工况内,可认为过程是平稳的,建立了主元分析模型来进行监测,从而实现了对非平稳过程的状态监测。利用浙江某发电厂一次风机实际的运行数据,验证了这种方法的可行性和有效性。该方法还可以扩展和推广到火电热力系统其他设备,乃至其他非平稳过程中。

### 参考文献:

- [1] KOPPEN-SELIGER B, KIUPEL N, KELLINGHAUS H S, et al. A fault diagnosis concept for a high-pressure-preheater line[C]//Proceedings of 1995 34th IEEE Conference on Decision and Control. 1995(3):2383-2388.
- [2] BARSZCZ T, CZOP P. A feedwater heater model intended for model-based diagnostics of power plant installations[J]. Applied Thermal Engineering, Elsevier Ltd, 2011, 31(8-9):1357-1367.
- [3] 周东华, 胡艳艳. 动态系统的故障诊断技术[J]. 自动化学报, 2009, 35(6):748-758.
- [4] CHEN K Y, CHEN L S, CHEN M C, et al. Using SVM based method for equipment fault detection in a thermal power plant[J]. Computers in Industry, Elsevier B.V., 2011, 62(1):42-50.
- [5] WISE B M, RICKER N L, VELTKAMP D F, et al. Theoretical basis for the use of principal component models for monitoring multivariate processes[J]. Process Control and Quality, 1990, 1(1):41-51.
- [6] 黄孝彬, 刘吉臻, 牛玉广. 主元分析方法在火电厂锅炉过程故障检测中的应用[J]. 动力工程学报, 2004, 24(4):542-547.
- [7] 牛征, 刘吉臻, 牛玉广. 基于主元分析的火电厂生产过程故障检测研究[J]. 华北电力大学学报(自然科学版), 2005, 32(4):33-37.
- [8] 李晓杰, 田亮. 基于多尺度主元分析的磨煤机堵塞检测[J]. 电力科学与工程, 2018, 34(1):41-45.

- [9] 张曦, 阎威武, 刘振亚, 等. 基于核主元分析和邻近支持向量机的汽轮机凝汽器过程监控和故障诊断[J]. 中国电机工程学报, 2007, 27(14):56-61.
- [10] ZHANG S, ZHAO C, ZHAO S. Two-level modeling strategy based incipient fault detection for nonstationary industrial processes[C]//2017 IEEE 56th Annual Conference on Decision and Control, CDC 2017.
- [11] CHEN Q, KRUGER U, LEUNG A Y T. Cointegration testing method for monitoring nonstationary processes[J]. Industrial and Engineering Chemistry Research, 2009, 48(7):3533-3543.
- [12] LI G, QIN S J, YUAN T. Nonstationarity and cointegration tests for fault detection of dynamic processes[M]. IFAC Proceedings Volumes (IFAC-PapersOnline), IFAC, 2014.
- [13] ZHAO C, HUANG B. A full-condition monitoring method for nonstationary dynamic chemical processes with cointegration and slow feature analysis[J]. AIChE Journal, 2018, 64(5):1662-1681.
- [14] 许仙珍, 谢磊, 王树青. 基于 GMM 的多工况过程监测方法[J]. 计算机与应用化学, 2010, 27(1):17-22.
- [15] 李文卿. 数据驱动的复杂工业过程统计过程监测[D]. 杭州: 浙江大学, 2018.
- [16] ZHAO C, SUN Y. Step-wise sequential phase partition (SSPP) algorithm based statistical modeling and online process monitoring[J]. Chemometrics and Intelligent Laboratory Systems, Elsevier B.V., 2013(125):109-120.
- [17] YUE H H. Reconstruction-based fault identification using a combined index[J]. Industrial & Engineering Chemistry Research, 2001, 40(20):4403-4414.

---

收稿日期: 2019-09-23

作者简介: 沙万里(1980), 男, 工程师, 主要从事火电厂智能生产、智慧管理方向的研究及应用。

赵春晖(1979), 女, 教授, 主要研究方向为大数据分析及其在不同领域的应用等。(通讯作者)

(本文编辑: 徐 晗)