

基于全局局部保留投影与测地距离的气化炉故障检测方法

庄稼¹, 韦炜², 朱书奔², 李扬³, 鲍涛¹, 王伟¹, 常雪丁¹, 王村松⁴

(1. 国家能源集团宁夏煤业有限责任公司煤制油分公司, 宁夏 灵武 750411; 2. 浙江中控软件技术有限公司, 杭州 310053;
3. 中控技术股份有限公司, 杭州 310053; 4. 南京工业大学智能制造研究院, 南京 210009)

摘要: 高温、高压、强腐蚀工作环境下的气化炉易发生仪表测量故障, 进而影响煤制油、煤制甲醇等生产工艺, 甚至危及人员安全。针对上述问题, 提出了一种基于全局局部保留投影(global-local preserving projection, GLPP)算法与测地距离的气化炉故障检测方法。首先, 采用 GLPP 算法提取样本邻域确定的数据局部特征; 然后, 提出一种基于测地距离度量样本的非近邻关系的数据全局特征提取方法; 进一步, 利用提取的全局特征构建相应的统计量来进行故障检测。最后, 分别通过田纳西伊斯曼(Tennessee Eastman, TE)与真实气化炉 2 个案例验证所提出方法的有效性和可行性。

关键词: 故障检测; 气化炉; 测地距离; 全局局部保留投影

中图分类号: TP277 **文献标志码:** A **文章编号:** 1671-1807(2024)14-0266-08

随着全球能源需求的不断增长和传统石油资源逐渐枯竭, 将煤转化为液体燃料的煤制油技术具有重要的战略意义^[1]。煤制油作为一种可持续发展的能源转化方式, 不仅可以有效利用煤炭资源, 还可以减少对传统石油资源的依赖, 提供稳定的能源供应。在煤制油工艺中, 作为关键核心装置的气化炉将煤炭在高温、高压和强腐蚀环境下进行热解和转化^[2-4]。然而, 气化炉的恶劣工况和极端操作条件容易导致仪表设备的故障, 从而妨碍操作人员对炉况的准确账款。这不仅影响生产运行效益产生, 甚至可能引发严重的安全风险, 危及人员生命。因此, 气化炉设备的精准故障检测对安全生产和稳定运行至关重要^[5-6], 具有重要的理论和实践意义。

现阶段, 由于集散控制系统(distributed control system, DCS)与仪表的广泛应用, 大量运行数据被采集, 为数据驱动的过程监测技术提供了良好的基础。以主元分析法(principal component analysis, PCA)为代表的统计过程监测方法是应用最广泛的

数据驱动的过程监测技术, 通过将原始数据投影到低维空间提取数据的方差信息用于监测^[7-9]。例如, 张谦谦等^[10]使用加权差分方法处理提取的数据信息, 提出了一种动态加权差分主成分分析法(dynamic weighted differential principal component analysis, DWDPCA), 解决了工业过程中存在的动态特性和多模态特性问题。郭金玉等^[11]提出一种基于特征值变化的工业过程实时故障检测方法, 揭示了建模数据之间的内在相关性, 特征值能够过程故障的波动而变化。宋冰等^[12]提出了一种变分自编码器-正交典型相关分析(variational automatic encoder-orthogonal CCA, VAE-OSCCA)方法, 实现了对输入数据进行无监督自适应学习和高维非线性过程变量的特征提取。然而, 尽管这类方法有适用于非线性、动态、多模态等多种过程特征的改进, 但其只关注数据的全局特征, 忽视了数据以邻域信息为代表的局部特征。

目前, 以局部保留投影(local preserving pro-

收稿日期: 2024-03-30

基金项目: 国家能源集团项目(煤制油分公司服务(咨询)〔2023〕001号)

作者简介: 庄稼(1979—), 男, 宁夏银川人, 工程师, 硕士研究生导师, 研究方向为自控仪表及智能化; 韦炜(1992—), 男, 浙江东阳人, 工程师, 研究方向为自动化、先进预测控制; 朱书奔(1992—), 男, 温州龙港人, 工程师, 研究方向为自动化、先进预测控制; 李扬(1997—), 男, 山西长治人, 硕士, 工程师, 算法工程师, 研究方向为自动化、先进预测控制; 鲍涛(1985—), 男, 宁夏银川人, 工程师, 研究方向为自动化、测控技术与仪器; 王伟(1991—), 男, 回族, 宁夏银川人, 助理工程师, 研究方向为自动化、测控技术与仪器; 常雪丁(1997—), 男, 宁夏石嘴山人, 助理工程师, 研究方向为自动化; 王村松(1990—), 男, 黑龙江伊春人, 博士, 副教授, 硕士研究生导师, 研究方向为故障预测与健康管理。

jecting, LPP)、邻域保持嵌入(neighborhood preserving embedding, NPE)等方法为代表的基于流形的特征提取算法因其保留了数据的邻域信息而被应用于监测领域^[13-15]。例如, He 和 Xu^[16]提出一种基于统计模式分析与 LPP 的监测方法应用于非高斯过程。Yang 等^[17]提出一种考虑数据动态特征的时间邻域保持嵌入方法来监测动态过程故障,成功应用于 CE117 过程和 TE 过程。徐静等^[18]提出一种基于 DGLPP-SVDD(dynamic global-local preserving projection-support vector data description)算法的化工过程故障检测方法,提高了动态监测能力和检测准确率。宋冰等^[19]提出一种基于双子空间并行回归(twin-space parallel regression, TSPR)的质量相关故障检测方法,实现了对数据的流形特征和变化速度信息的同时提取。然而,上述方法虽然保留数据局部特征,却忽视了数据的全局特征。

为解决上述问题,部分学者提出了一些能够同时保留数据全局与局部特征的监测方法。例如, Zhang 等^[20]通过权重系数整合 PCA 与 LPP 的目标函数提出了全局-局部结构分析(global-local structure analysis, GLSA)监测方法,解决了 LPP 中可能出现的奇异性问题,并且很好地保持了 PCA 方法的正交特性。Yu^[21]通过最大化 PCA 与 LPP 目标函数的比值提出了局部-全局 PCA(local and global principle component analysis, LGPCA)监测方法,用于化工过程的动态监测。Luo^[22]将保留局部特征的 LPP 目标函数扩展到了全局非近邻点特征的保持,提出一种全局局部保留投影(global-local preserving projection, GLPP)监测方法,解决了局部和全局双目优化问题,用于故障检测与诊断。此外许多研究者还对上述方法提出了动态、非线性与非参数等角度的改进^[23-25]。

然而,上述基于 GLPP 的监测方法依然受到 GLPP 本身对数据全局与局部特征提取能力的限制而影响了其监测性能。因此,对 GLPP 本身特征提取能力的改进可以进一步提升基于 GLPP 的过程监测方法的监测性能。例如, GLPP 中均使用热核函数来定义数据的全局与局部权重矩阵,热核函数中的经验参数会对其特征提取及基于此的监测性能产生较大影响。Luo 等^[26]与 Cui 等^[27]分别提出了 GLPP 与 LNLE 方法的无参数化方法,使全局-局部特征提取方法的性能更加稳定,减少了参数选择对监测性能的影响。

在 GLPP 的实现步骤中,特征的描述和度量对

于其特征提取性能至关重要。GLPP 中对局部特征和全局特征的描述都基于欧氏距离的近邻和非近邻关系,这与流形学习的理论假设并不完全一致。考虑到流形定义仅涉及流形上的邻域和欧式空间同胚,即流形的局部可以用欧式度量来表示,而流形上的全局度量并不能直接采用欧氏距离准确度量。针对这个问题,经典的等距映射(Isometric Map, Isomap)方法在多尺度变换的基础上引入了测地距离(geodesic distance, GD)来代替欧氏距离,以度量数据的全局特征,从而有效提取了数据的全局特征^[28]。

受到上述全局-局部保留投影算法与等距映射方法的启发,本文提出了一种基于全局局部保留投影(GLPP)与测地距离的气化炉故障检测方法。主要创新之处在于:①利用 K 近邻方法构建了邻接图,在确定的邻域范围提取得到了数据的局部特征;②为了更好地描述非线性数据的距离关系,在构建的样本近邻关系基础上,进一步考虑流形的内在结构和曲率,引入了测地距离代替欧氏距离表征非近邻点的距离,提取到了数据真实的全局特征;③在提取的数据全局与局部特征基础上,构建了一个双目标优化函数,更好地完全保留了数据全局与局部特征的显式投影,提高了对过程的实时监测率。

1 预备知识

1.1 全局局部保留投影算法

全局局部保留投影算法可以同时提取数据的全局与局部特征^[22]。首先对于标准化后的数据集 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{m \times n}$, GLPP 方法旨在得到一个用于保留数据全局与局部特征的投影矩阵 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d] \in \mathbf{R}^{m \times d}$, 然后将 $\mathbf{X}$ 投影到 $\mathbf{Y} = \mathbf{A}^T \mathbf{X} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n] \in \mathbf{R}^{d \times n}$, 其目标函数如下:

$$\begin{aligned}
 J_{\text{GLPP}}(\mathbf{a}) &= \min_{\mathbf{a}} \{ \eta J_{\text{Local}}(\mathbf{a}) - (1 - \eta) J_{\text{Global}}(\mathbf{a}) \} \\
 &= \min_{\mathbf{a}} \frac{1}{2} \left\{ \eta \sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 W_{ij} - \right. \\
 &\quad \left. (1 - \eta) \sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 \bar{W}_{ij} \right\} \\
 &= \min_{\mathbf{a}} \frac{1}{2} \sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 R_{ij} \\
 &= \min_{\mathbf{a}} \left\{ \sum_i \mathbf{y}_i H_{ii} \mathbf{y}_i^T - \sum_{ij} \mathbf{y}_i R_{ij} \mathbf{y}_j^T \right\} \\
 &= \min_{\mathbf{a}} \left\{ \sum_i \mathbf{a}^T \mathbf{x}_i H_{ii} \mathbf{x}_i^T \mathbf{a} - \sum_{ij} \mathbf{a}^T \mathbf{x}_i R_{ij} \mathbf{x}_j^T \mathbf{a} \right\} \\
 &= \min_{\mathbf{a}} \mathbf{a}^T \mathbf{X} (\mathbf{H} - \mathbf{R}) \mathbf{X}^T \mathbf{a} \\
 &= \min_{\mathbf{a}} \mathbf{a}^T \mathbf{X} \mathbf{M} \mathbf{X}^T \mathbf{a}
 \end{aligned} \tag{1}$$

式中: $\mathbf{A}$ 为由投影向量 $\mathbf{a}$ 组成的投影矩阵; $R_{ij} =$

$\eta \mathbf{W}_{ij} - (1 - \eta) \bar{\mathbf{W}}_{ij}$, 为加权的权重矩阵; $H_{ii} = \sum_j R_{ij}$, 为度矩阵; $\mathbf{M} = \mathbf{H} - \mathbf{R}$, 为综合权重矩阵; η 为平衡局部信息和全局信息的重要参数; $\mathbf{W}$ 为数据近邻关系的权重矩阵, W_{ij} 为第 i 行第 j 列的元素, 表示数据点 $\mathbf{x}_i$ 和 $\mathbf{x}_j$ 之间的权重, 计算如公式(2)所示:

$$W_{ij} = \begin{cases} e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma_1}}, & \mathbf{x}_j \in \Omega(\mathbf{x}_i), \mathbf{x}_i \in \Omega(\mathbf{x}_j) \\ 0, & \text{其他} \end{cases} \quad (2)$$

式中: $\bar{\mathbf{W}}$ 为数据非近邻关系的权重矩阵, 矩阵每个元素 $\bar{W}_{ij}$ 的计算如公式(3)所示:

$$\bar{W}_{ij} = \begin{cases} e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma_2}}, & \mathbf{x}_j \notin \Omega(\mathbf{x}_i), \mathbf{x}_i \notin \Omega(\mathbf{x}_j) \\ 0, & \text{其他} \end{cases} \quad (3)$$

式中: $\Omega(\mathbf{x}_i)$ 为 $\mathbf{x}_i$ 的邻域即 $\mathbf{x}_i$ 最近的几个数据点的集合。与通常由 k 近邻方法确定; σ_1 与 σ_2 为经验常数。

为了避免小样本情况下 $\mathbf{XHX}^T$ 的奇异值问题, GLPP 给定了如公式(4)所示的约束以求解公式(1)的目标函数:

$$\begin{aligned} \mathbf{a}^T (\eta \mathbf{XHX}^T + (1 - \eta) \mathbf{I}) \mathbf{a} &= \mathbf{a}^T \\ \mathbf{Na} &= 1 \end{aligned} \quad (4)$$

式中: η 为权重参数, 用于平衡局部信息和全局信息的影响; $\mathbf{V}$ 为权重矩阵, 结合了度矩阵 $\mathbf{H}$ 和单位矩阵 $\mathbf{I}$ 。

由式(1)与式(4)构成的优化问题可以转化为求解如式(5)所示的广义特征值分解问题:

$$\mathbf{XMX}^T \mathbf{a} = \lambda \mathbf{Na} \quad (5)$$

通过求解矩阵 $\mathbf{M}$ 前 d 个最小的特征值 $\lambda_1 < \lambda_2 < \dots < \lambda_d$ 所对应的特征向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d$, 获得用以提取数据 $\mathbf{X}$ 全局与局部特征的投影矩阵 $\mathbf{A}$ 。

1.2 测地距离

测地距离是一种用于测量流形数据中点之间的距离的概念。与传统的欧氏距离不同, 测地距离考虑了流形的内在结构和曲率, 能更好地描述非线性数据的距离关系。通过测地距离, 可以捕捉到数据样本之间的全局关系, 提供更准确的特征表示。计算流形数据点之间的测地距离需要以下步骤。

步骤 1: 构建邻接图, 基于数据样本之间的距离关系, 构建邻接图表示流形的拓扑结构。对于数据集 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{m \times n}$, 如果样本点 $\mathbf{x}_i$ 是 $\mathbf{x}_j$ 基于欧氏距离最近的 k 个近邻点或者 $\mathbf{x}_j$ 属于 $\mathbf{x}_i$ 基

于欧氏距离最近的 k 个近邻点 (其中 k 由经验确定), 则表明两样本点之间具有近邻关系, 近邻图 G 中对应顶点有边连接。

步骤 2: 计算最短路径, 在邻接图上采用最短路径算法 (如 Dijkstra 算法), 计算数据样本点之间的最短路径, 即测地距离^[29]。

2 基于 GLPP 与 GD 的过程监测方法

2.1 GD-GLPP 算法

对于标准化的数据集 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{m \times n}$, 首先采用前文的 K 近邻邻域构造方法为数据的每个样本点根据近邻个数 K 确定邻域, 代入式(2)得到近邻权重矩阵 $\mathbf{W}$ 。基于上述邻域构建的近邻图, 引入测地距离来度量非近邻点之间的距离。Dijkstra 算法被用来计算近邻图上的最短路径距离以估计非近邻点间的测地距离。将非近邻点的测地距离代入式(3)得到非近邻权重矩阵 $\bar{\mathbf{W}}$ 。

将确定的近邻权重矩阵与非近邻权重矩阵代入式(1)。由于式(1)中 η 决定着两个子目标函数 $J_{\text{Local}}(\mathbf{a})$ 与 $J_{\text{Global}}(\mathbf{a})$ 的重要性, η 的确定会直接影响到算法对局部特征与全局特征保留的比例。本文子目标函数的规模定义为对应子目标函数拉普拉斯矩阵的谱半径, η 通过平衡两个子目标函数的规模来确定, 如式(6)所示:

$$\begin{aligned} \eta \rho(\bar{\mathbf{L}}) &= (1 - \eta) \rho(\mathbf{L}) \\ \eta &= \frac{\rho(\mathbf{L})}{\rho(\mathbf{L}) + \rho(\bar{\mathbf{L}})} \end{aligned} \quad (6)$$

式中: $\rho(\cdot)$ 为谱半径; 两个子目标函数的拉普拉斯矩阵 $\mathbf{L}$ 与 $\bar{\mathbf{L}}$ 计算如下:

$$\begin{aligned} \mathbf{L} &= \mathbf{D} - \mathbf{W} \\ \mathbf{D}_{ii} &= \sum_j W_{ij} \end{aligned} \quad (7)$$

$$\begin{aligned} \bar{\mathbf{L}} &= \bar{\mathbf{D}} - \bar{\mathbf{W}} \\ \bar{\mathbf{D}}_{ii} &= \sum_j \bar{W}_{ij} \end{aligned} \quad (8)$$

式中: $\bar{\mathbf{D}}_{ii}$ 为度矩阵 $\mathbf{D}$ 的对角元素, 表示节点 i 的度数 C 节点 i 连接到的所有节点的权重之和。

之后类似于 GLPP 算法, 将确定的 η 与权重矩阵 $\mathbf{W}, \bar{\mathbf{W}}$ 代入公式(1)和式(4), 转化为求解广义特征值分解问题, 前 d 个最小的特征值对应特征向量组成了保留数据全局与局部特征的投影矩阵 $\mathbf{A}$ 。

2.2 基于 GD-GLPP 算法的过程监测方法

将标准化后的正常工况数据集 $\mathbf{X} = [\mathbf{x}_1,$

$\mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{m \times n}$ 作为训练数据, 利用 GD-GLPP 算法得到投影矩阵 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d] \in \mathbf{R}^{m \times d}$ (d 为主元个数):

$$\mathbf{y}_{\text{new}} = \mathbf{A}^T \mathbf{x}_{\text{new}} \quad (9)$$

式中: $\mathbf{y}_{\text{new}}$ 为新样本点相对于均值向量的差异向量。

实时数据 $\mathbf{x}_{\text{new}}$ 通过投影矩阵将其投影到特征空间。 T^2 统计量用于衡量特征空间的变化来实现对实时数据的监测, 计算如式(10)所示。

$$T_{\text{new}}^2 = \mathbf{y}_{\text{new}}^T \mathbf{S}^{-1} \mathbf{y}_{\text{new}} \quad (10)$$

式中: $\mathbf{S}$ 为训练数据 $\mathbf{X}$ 投影到特征空间的潜变量 $\mathbf{Y}$ 的协方差矩阵, 即 $\mathbf{S} = \mathbf{Y}^T \mathbf{Y} / (n-1)$; T_{new}^2 为新样本的 T^2 统计量, 其统计量控制限的计算如式(11)所示。

$$T_{\alpha}^2 = \frac{k(n^2-1)}{n(n-k)} F_{\alpha}(m, n-m) \quad (11)$$

式中: k 为变量的个数; n 为样本数量; $F_{\alpha}(m, n-m)$ 为 F 分布的临界值, 对应显著水平 α 和自由度 m 和 $n-m$; $\frac{k(n^2-1)}{n(n-k)}$ 为一个权重因子, 用于调整 F 分布的临界值以适应 T^2 统计量的计算。

因此, 本文所提过程监测方法监测流程包括离线数据建模与在线监测两部分, 如图 1 所示。其中步骤 1~步骤 9 为离线建模部分, 步骤 10~步骤 13 为在线监测部分, 具体步骤如下。

步骤 1: 对训练数据进行 Z-score 标准化处理, 得到训练数据样本的均值和方差。

步骤 2: 参照 1.2 节中的步骤 1, 对标准化后的数据利用 K 近邻方法构造邻域。

步骤 3: 参照 1.2 节中的步骤 2, 利用测地距离计算非近邻样本之间的距离。

步骤 4: 利用式(2)计算权重矩阵 $\mathbf{W}$ 。

步骤 5: 利用构建好的近邻图, 利用测地距离来度量非近邻点之间的距离。

步骤 6: 将非近邻点的测地距离代入式(3)得到非近邻权重矩阵 $\bar{\mathbf{W}}$ 。

步骤 7: 将确定的近邻权重矩阵与非近邻权重矩阵代入式(1), 利用式(6)确定参数 η 。

步骤 8: 将计算得到的参数 η 与权重矩阵 $\mathbf{W}$ 、 $\bar{\mathbf{W}}$ 代入式(1)和式(4), 将其转化为求解广义特征值分解问题, 构建 GLPP 模型。前 d 个最小的特征值对应特征向量组成了保留数据全局与局部特征的投影矩阵 $\mathbf{A}$ 。

步骤 9: 利用式(10)和式(11)计算统计量控制限。

步骤 10: 获得系统在线监测运行实时数据。

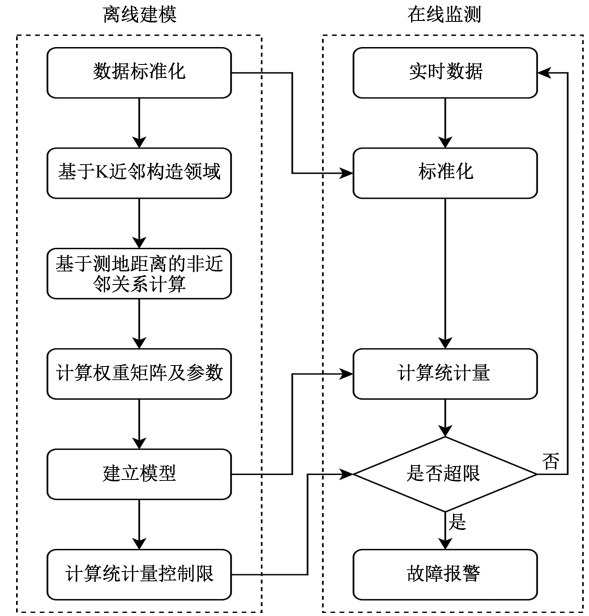


图 1 算法监测流程

步骤 11: 对实时数据按照离线建模步骤 1 中得到的均值与方差进行标准化。

步骤 12: 基于离线建立的 GLPP 模型对实时数据投影后计算统计量。

步骤 13: 判断统计量是否超出控制限, 如果超限则报警, 否则返回步骤 10。

3 案例分析

3.1 TE 过程故障检测

TE 过程由 Downs 和 Vogel, 即提出, 被广泛应用于过程监测领域各种监测方法监测效果的基准测试^[30]。TE 过程工艺流程如图 2 所示, 包括 5 个主要的单元, 即反应器、冷凝器、压缩机、汽提器和分离器。该过程包括 41 个测量变量和 12 个控制变量, 采样频率为 3 min。本文选择了其中包括 11 个控制变量 (XMV1-XMV11) 与 22 个测量变量 (XMEAS1-XMEAS22) 的 33 个变量作为监测变量。用于建模的训练数据包含 960 个样本。用于测试的 21 个故障数据集均包含 960 个样本, 故障从 161 个样本点引入。21 个故障的描述如表 1 所示。

将本文所提方法与 GLPP、PCA 在训练数据上建模。三种方法的主元个数通过 80% 的方差贡献率分别确定为 27、27、12。参考 Luo^[22], GD-GLPP 与 GLPP 的近邻个数 k 设定为 10。三种方法的统计量置信限均选择为 99%。为了对比三种监测方法的监测结果, 检测率 (fault detection rate, FDR) 与错报率 (false alarm rate, FAR) 被用来评估三种检测方法的监测效果, 其计算如式(16)和式(17)所示:

$$\text{FDR} = \frac{x \in \text{fault}; T^2(x) > \text{limit}}{\text{normal samples}} \times 100\% \quad (12)$$

$$\text{FAR} = \frac{x \in \text{normal}; T^2(x) > \text{limit}}{\text{normal samples}} \times 100\% \quad (13)$$

表 1 TE 过程故障说明

编号	故障描述	类型
1	A/C 进料流量比变化,组分 B 含量保持不变(管道 4)	阶跃
2	组分 B 含量发生变化,A/C 进料流量比阶跃不变(管道 4)	阶跃
3	物料 D 的温度发生变化(管道 2)	阶跃
4	反应器冷却水入口温度发生变化	阶跃
5	冷凝器冷却水入口温度发生变化	阶跃
6	物料 A 损失(管道 1)	阶跃
7	物料 C 压力损失(管道 4)	阶跃
8	物料 A、B、C 的组成发生变化(管道 4)	随机变量
9	物料 D 进料的温度发生变化(管道 2)	随机变量
10	物料 C 进料的温度发生变化(管道 2)	随机变量
11	反应器冷却水入口温度发生变化	随机变量
12	冷凝器冷却水入口温度发生变化	随机变量
13	反应动力学特性发生变化	缓慢漂移
14	反应器冷却水阀门	黏滞
15	冷凝器冷却水阀门黏滞	黏滞
16~20	未知	未知
21	管道 4 的阀门固定在稳态位置	恒定位置

式中: $x \in \text{fault}$ 为被检测为异常的样本集合; $x \in \text{normal}$ 为被检测为正常的样本集合; $T^2(x)$ 为样本 x T^2 统计量; limit 为 T^2 统计量的临界值,用于判断样本是否异常; normal samples 为实际正常的样本总数。

三种监测方法在 TE 过程的检测率与错报率如表 2 所示。现有的研究已经证明了 TE 过程中故障 3、9 和 15 难以检测,GD-GLPP 的检测结果也出现了同样的情况,因此本文不再讨论这三个故障。从表 2 可以看出,本文所提 GD-GLPP 方法在除了故障 3、9、15 的 18 个故障中的 17 个取得了最高的检测率。相比于 GLPP 方法,GD-GLPP 方法在故障 1、10、11、16、17、18、19、20 与 21 共 9 个故障上拥有更高的检测率。最后本文计算了讨论的 18 个故障的平均检测率与错报率。相比于其他两种方法,GD-GLPP 方法拥有最高的平均检测率。同时 GD-GLPP 方法还拥有比 GLPP 方法更低的错报率。尽管 GD-GLPP 方法的平均错报率比 PCA 方法略高,但其平均检测率远高于 PCA 方法。因此,GD-GLPP 监测方法可以在实现对大多数故障优秀的检测率的同时保持较低的错报率。

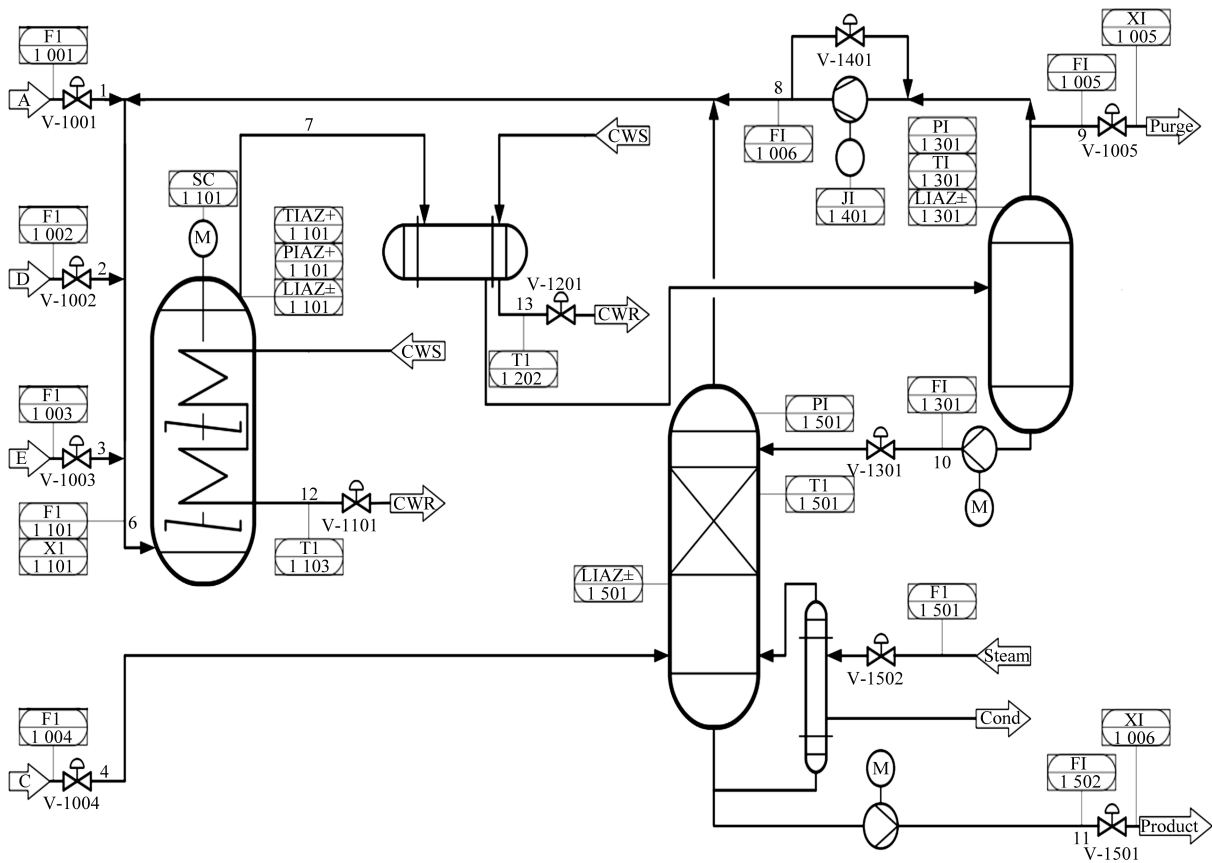


图 2 TE 过程工艺流程

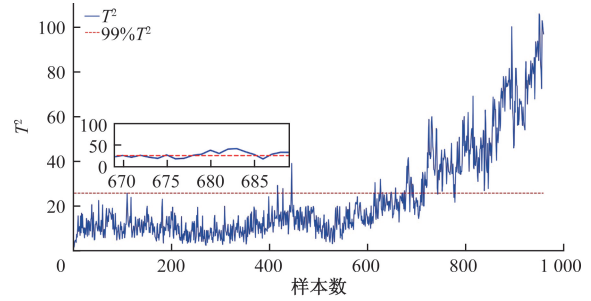
表 2 不同方法在 TE 过程的检测率与错报率

故障	PCA		GLPP		GD-GLPP	
	FDR/%	FAR/%	FDR/%	FAR/%	FDR/%	FAR/%
1	99.13	0.00	99.88	1.25	100.00	0.63
2	98.50	1.25	98.25	0.00	98.38	0.00
3	1.13	0.63	2.00	3.13	2.00	3.13
4	11.25	0.63	100.00	0.63	100.00	0.00
5	23.88	0.63	100.00	0.63	100.00	0.00
6	99.00	0.00	100.00	0.00	100.00	0.63
7	100.00	0.00	100.00	0.00	100.00	0.63
8	96.75	0.00	97.75	0.00	97.75	0.00
9	1.75	1.25	1.13	0.63	1.50	1.25
10	30.25	0.63	86.00	0.00	86.88	0.00
11	32.00	0.63	77.13	0.63	77.38	0.63
12	98.13	0.63	99.88	0.63	99.88	0.00
13	93.50	0.00	95.25	0.00	95.25	0.00
14	98.38	0.00	100.00	0.00	100.00	0.00
15	1.13	0.63	3.38	1.88	4.25	3.13
16	13.75	0.63	88.75	0.00	90.63	1.25
17	76.13	0.63	97.00	2.50	97.13	1.875
18	89.38	0.00	90.00	0.00	90.13	0.63
19	7.50	0.00	93.25	1.25	94.00	0.63
20	31.38	0.00	90.63	0.63	91.00	0.63
21	34.38	0.00	45.13	2.50	54.38	2.50
平均	62.96	0.31	92.16	0.59	92.93	0.56

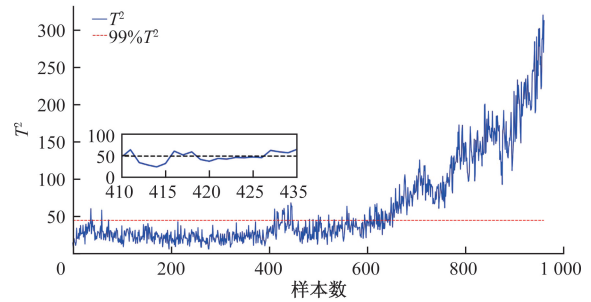
三种监测方法以故障 21 为例的监测结果如图 3 所示。从图 3 中可以看出三种方法均能在故障发生后检测到。其中图 3(a)中 PCA 监测方法在 679 样本点检测到故障,故障检测延迟较大。图 3(b)中 GLPP 监测方法与图 3(c)中 GD-GLPP 监测方法均在 417 样本点检测到故障。相比于 PCA 监测方法, GLPP 与 GD-GLPP 监测方法提前了 262 个样本点。同时相比于 GLPP 监测方法, GD-GLPP 监测方法拥有更高的检测率。故障 21 的监测结果说明 GD-GLPP 方法对数据全局与局部特征准确的提取有效提高了对故障的监测效果。

3.2 气化炉故障检测

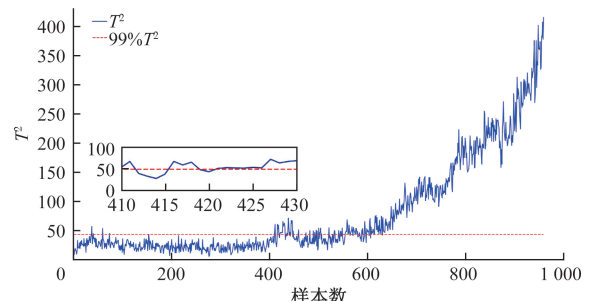
进一步该方法被应用于某企业一气化炉炉膛高温热电偶失效检测,采集包含一次热电偶失效记录,采样间隔为 1 min 的运行数据共 5 760 组。基于工艺知识与现场操作人员经验确定输入变量,包括入口煤浆流量、入口煤浆压力、入炉激冷水流量、入口氧气流量、入口氧气压力、入口氧气温度、气化炉压力、气化炉出口工艺气温度、入炉激冷水温度、炉内反应温度等共 22 个变量。选择采集数据中的前 4 760 组作为建模数据,后 1 000 组作为测试数据。相关参数的设置与 TE 过程中建模参数保持一致。



(a) PCA 监测结果



(b) GLPP 监测结果



(c) GD-GLPP 监测结果

图 3 三种方法对故障 21 的监测

与直接采用 PCA 模型、GLPP 模型对比的结果如图 4 所示。可以看出, GD-GLPP 监测方法在图中 275 样本超限报警, 相比于 PCA 与 GLPP 模型分别提前了 60 min 与 103 min。与该企业操作人员确认, 并与操作日志中对比, 在 3 月 13 日 15:30 气化炉内一高温热电偶失效, 所提方法可提 2 h 23 min 检测到该仪表的故障, 为操作人员提供及时的决策支持, 验证了该方法的有效性。

4 结论

本文针对气化炉易发生的仪表故障, 提出了一种基于全局局部特征保留投影与测地距离的故障监测方法, 很好地解决了高温、高压、强腐蚀工作环境下的气化炉仪表测量故障检测的难题, 在 TE 过程与气化炉 2 个案例进行了验证分析。案例结果表明, 相比于传统的基于 PCA 与 GLPP 的监测方法, 本文所提出的监测方法由于更为准确的全局特征的保留, 不论在解决传统 TE 过程故障检测还是实际工程气化炉仪表测量故障检测难题方面都展现

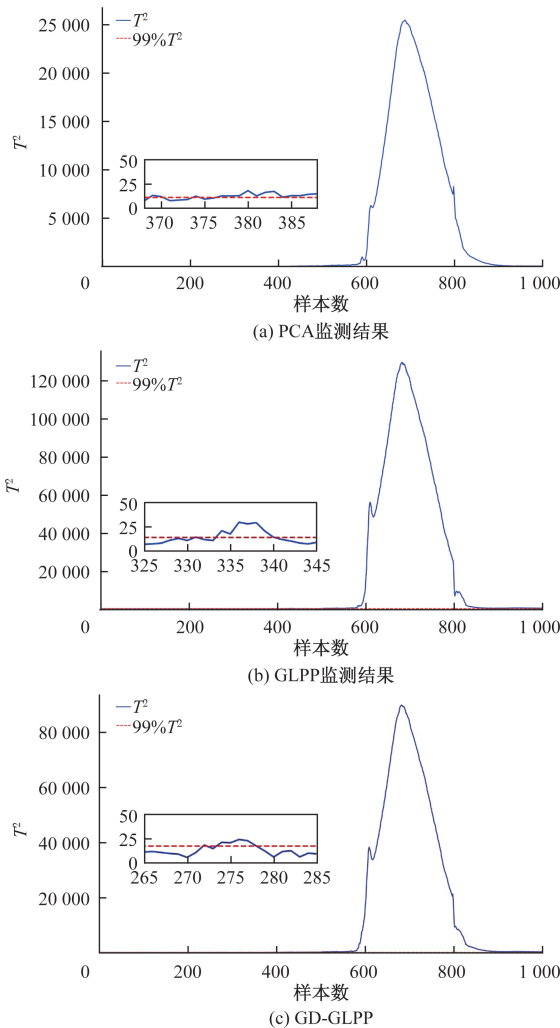


图4 三种方法对气化炉高温热电偶故障的监测

了出色的效果。

本文所提出的方法在实际工业生产中,针对关键设备开展预测性维护具有重要的意义和实用性。该方法同样可以进一步推广应用于其他类似的工业生产过程中,如化工、能源等领域。在未来,结合人工智能技术的进一步优化,还有更大的发展空间和应用潜力。

参考文献

- [1] 乔元,王建立. 现代煤制油化工安全管理体系的探索与实践[J]. 中国安全科学学报, 2021, 31(S1): 34-38.
- [2] GUO Q, LI S, GONG Y, et al. Application of qualitative trend analysis in fault diagnosis of entrained-flow coal-water slurry gasifier [J]. Control Engineering Practice, 2021, 112: 104835.
- [3] ZHANG J, HOU J AND ZHANG Z. Real-time identification of out-of-control and instability in process parameter for gasification process; Integrated application of control chart and kalman filter [J]. Energy, 2022,

238: 121845.

- [4] AGUADO R, CASTELEIRO-ROCA J L, VERA D, et al. A hybrid intelligent model to predict the hydrogen concentration in the producer gas from a downdraft gasifier[J]. International Journal of Hydrogen Energy, 2022, 47(48): 20755-20770.
- [5] 彭开香,马亮,张凯. 复杂工业过程质量相关的故障检测与诊断技术综述[J]. 自动化学报, 2017, 43(3): 349-365.
- [6] 王晓慧,王延江,邓晓刚,等. 基于加权深度支持向量数据描述的工业过程故障检测[J]. 化工学报, 2021, 72(11): 5707-5716.
- [7] GUO Q, LI S, GONG Y, et al. Application of qualitative trend analysis in fault diagnosis of entrained-flow coal-water slurry gasifier [J]. Control Engineering Practice, 2021, 112: 104835.
- [8] CALLÉN M S, MARTÍNEZ I, GRASA G, et al. Principal component analysis and partial least square regression models to understand sorption-enhanced biomass gasification[J]. Biomass Conversion and Biorefinery, 2024, 14: 2091-2111.
- [9] YIN S, DING S X, XIE X, et al. A review on basic data-driven approaches for industrial process monitoring[J]. IEEE Transactions on Industrial Electronics, 2014, 61(11): 6418-6428.
- [10] 张谦谦,王文标,郝友维. 基于动态加权差分主成分分析的故障检测[J]. 科学技术与工程, 2023, 23(36): 15522-15529.
- [11] 郭金玉,赵文君,李元. 基于特征值变化的工业过程实时故障检测[J]. 计算机应用与软件, 2023, 40(6): 330-336.
- [12] 宋冰,郑城风,侍洪波,等. 基于 VAE-Occa 的质量相关故障检测方法研究[J]. 化工学报, 2023, 74(4): 1630-1638.
- [13] YU J. Adaptive hidden Markov model-based online learning framework for bearing faulty detection and performance degradation monitoring[J]. Mechanical Systems and Signal Processing, 2017, 83: 149-162.
- [14] SONG B, TAN S, SHI H. Process monitoring via enhanced neighborhood preserving embedding[J]. Control Engineering Practice, 2016, 50: 48-56.
- [15] SHA X, LUO C, DIAON. A robust fault detection method based on neighborhood preserving embedding [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
- [16] HE F, XU J. A novel process monitoring and fault detection approach based on statistics locality preserving projections[J]. Journal of Process Control, 2016, 37: 46-57.
- [17] YANG J, ZHANG M, SHI H, et al. Dynamic learning on the manifold with constrained time information and its application for dynamic process monitoring[J]. Chemometrics and Intelligent Laboratory Systems, 2017, 167:

- 179-189.
- [18] 徐静,王振雷,王昕. 基于DGLPP-SVDD算法的化工过程故障检测[J]. 控制工程, 2021, 28 (12): 2462-2468.
- [19] 宋冰,郭涛,侍洪波,等. 基于双子空间并行回归的化工过程质量相关故障检测方法[J]. 化工学报, 2023, 74 (11): 4600-4610.
- [20] ZHANG M, GE Z, SONG Z, et al. Global-local structure analysis model and its application for fault detection and identification[J]. Industrial & Engineering Chemistry Research, 2011, 50(11): 6837-6848.
- [21] YU J. Local and global principal component analysis for process monitoring [J]. Journal of Process Control, 2012, 22(7): 1358-1373.
- [22] LUOL. Process monitoring with global-local preserving projections[J]. Industrial & Engineering Chemistry Research, 2014, 53(18): 7696-7705.
- [23] ZHAN C, LI S, YANG Y. Improved process monitoring based on global-local manifold analysis and statistical local approach for industrial process[J]. Journal of Process Control, 2019, 75: 107-119.
- [24] LUO L, BAO S, MAO J, et al. Nonlinear process monitoring using data-dependent kernel global-local preserving projections[J]. Industrial & Engineering Chemistry Research, 2015, 54(44): 11126-11138.
- [25] WANG Y, LI S, LIU C, et al. Multiscale feature fusion and semi-supervised temporal-spatial learning for performance monitoring in the flotation industrial process [J]. IEEE Transactions on Cybernetics, 2023, 54(2): 974-987.
- [26] LUO L, BAO S, MAO J, et al. Nonlocal and local structure preserving projection and its application to fault detection[J]. Chemometrics and Intelligent Laboratory Systems, 2016, 157: 177-188.
- [27] CUI P, WANG X, YANG Y. Nonparametric manifold learning approach for improved process monitoring[J]. The Canadian Journal of Chemical Engineering, 2022, 100(1): 67-89.
- [28] XU K K, LI H X, LIU Z. ISOMAP-based spatiotemporal modeling for lithiumion battery thermal process[J]. IEEE Transactions on Industrial Informatics, 2017, 14 (2): 569-577.
- [29] DIJKSTRA E W. A note on two problems in connexion with graphs[J]. Numerische Mathematik, 1959, 1(1): 269-271.
- [30] DOWNS J J, VOGEL E F. A plant-wide industrial process control problem[J]. Computers & Chemical Engineering, 1993, 17(3): 245-255.

Global-local Preserving Projection and Geodesic Distance Based Gasifier Fault Detection

ZHUANG Jia¹, WEI Wei², ZHU Shuben², LI Yang³, BAO Tao¹,

WANG Wei¹, CHANG Xueding¹, WANG Cunsong⁴

(1. Energy Group Coal to Liquid Branch, Ningxia Coal Industry Co. Ltd., Lingwu 750411, Ningxia, China;

2. Zhejiang Supcon Software Co. Ltd., Hangzhou 310053, China; 3. SUPCON Technology Co. Ltd., Hangzhou 310053, China;

4. Institute of Intelligent Manufacturing, Nanjing Tech University, Nanjing 210009, China)

Abstract: Gasifiers operating in high-temperature, high-pressure, and highly corrosive working environments are prone to instrument measurement failures. The failure affects production processes such as coal-to-liquids and coal-to-methanol, and even endangers personnel safety. In order to solve the above problems, a global local preserving projection and geodesic distance gasifier based fault detection method were proposed in this paper. Firstly, the GLPP algorithm was adopted to extract the local features of the data determined by the sample neighborhood. Then, a geodesic distance measurement sample's non-neighbor relationship-based data global feature extraction method was proposed. Further, the extracted global features were used to construct corresponding statistics for fault detection. Finally, the effectiveness and feasibility of the proposed method were verified through two cases which were Tennessee Eastman (TE) and a real gasifier.

Keywords: fault detection; gasifier; geodesic distance; global-local preserving projection