

# 连续与离散变量协同分析的非平稳非高斯 工业过程异常检测

蔡美玲<sup>1</sup>, 罗 迪<sup>1,2</sup>, 肖敬日<sup>1</sup>, 李婧妍<sup>3</sup>, 刘金平<sup>1,2\*</sup>

(1. 湖南师范大学信息科学与工程学院, 湖南长沙 410081; 2. 湘江实验室, 湖南长沙 410205;  
3. 湖南师范大学数学与统计学院, 湖南长沙 410081)

**摘 要:** 工业过程数据涵盖离散和连续变量, 它们之间内在的统计分布特性是揭示系统运行状况的关键。然而, 现有的监测模型多聚焦于高斯假设下的连续过程变量, 忽略了离散变量、连续变量的多模分布特性以及数据中的噪声、离群点对模型的影响, 难以精准反映数据的真实分布特性, 对非高斯、非平稳过程的异常检测效果欠佳。因此, 本文提出一种鲁棒的工业过程异常检测方法——连续和离散变量协同分析的变分贝叶斯混合判别方法(Continuous and Discrete Variables-Concurrent Analysis-based Variational Bayesian Mixture Discriminant, CDVCA-VBMD)。通过构建基于变分贝叶斯推断的面向连续变量的混合学生分布模型与面向离散多变量的混合多项分布模型, 有效解决过程变量非高斯分布特性难以有效提取的问题, 实现了过程变量复杂相关性的高效处理与分析; 同时, 在过程监测中引入了持续学习的理念, 确保对非平稳时变工业过程异常的有效检测。以数值仿真系统和田纳西·伊斯曼过程为例, 进行了大量验证性和对比性实验。结果表明, CDVCA-VBMD 能准确估计非平稳工业过程多变量的混合多模分布特性, 并对数据中的噪声和离群点具有较强的鲁棒性, 从而实现工业异常的准确检测, 为非高斯复杂工业过程的长期、鲁棒监测提供了有力支持。

**关键词:** 混合学生分布; 混合多项分布; 非平稳过程; 非高斯过程; 变分贝叶斯推断; 持续学习

**基金项目:** 国家自然科学基金(No.62371187); 湘江实验室重大项目(No.22XJ01013)

**中图分类号:** TP273

**文献标识码:** A

**文章编号:** 0372-2112(2024)10-3291-10

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20240466

## Continuous and Discrete Variables-Concurrent Analysis-Based Nonstationary and Non-Gaussian Industrial Process Anomaly Detection

CAI Mei-ling<sup>1</sup>, LUO Di<sup>1,2</sup>, XIAO Jing-ri<sup>1</sup>, LI Jing-yan<sup>3</sup>, LIU Jin-ping<sup>1,2\*</sup>

(1. School of Information Science and Engineering, Hunan Normal University, Changsha, Hunan 410081, China;  
2. Xiangjiang Laboratory, Changsha, Hunan 410205, China;  
3. School of Mathematics and Statistics, Hunan Normal University, Changsha, Hunan 410081, China)

**Abstract:** Industrial process data encompasses continuous and discrete variables, whose underlying statistical characteristics are crucial for revealing operational conditions. However, current process monitoring models predominantly focus on continuous variables with Gaussian assumptions, which often overlook the significant effects of the multimodal distribution characteristics of process variables, as well as the noises and outliers in process data. These limitations hinder the models' ability to capture complex statistical characteristics, leading to low detection performance particularly in non-Gaussian and nonstationary processes. This article introduces a robust anomaly detection method termed continuous and discrete variables-concurrent analysis-based variational Bayesian mixture discriminator (CDVCA-VBMD). It models continuous variables with a mixed student's t-distribution and discrete variables with a mixed multinomial distribution based on variational Bayesian inference, which can adeptly manage and analyze the complex interdependencies between process variables and overcome the non-Gaussian nature of continuous variables effectively. Furthermore, CDVCA-VBMD incorporates continuous learning to ensure the effective detection of nonstationary industrial processes. Extensive validation and comparative ex-

periments were conducted on a numerical simulation system and the Tennessee Eastman (TE) process. The outcomes demonstrate that CDVCA-VBMD can accurately characterize the mixed multimodal distribution characteristics of time-varying industrial processes, facilitating accurate anomaly detection. Additionally, the method exhibits robustness against noise and outliers in process data, supporting long-term and reliable monitoring of complex and non-Gaussian industrial processes.

**Key words:** mixed student distribution; mixed multinomial distribution; nonstationary processes; non-Gaussian processes; variable Bayesian inference; continuous learning

**Foundation Item(s):** National Natural Science Foundation of China (No. 62371187); Major Program Project of Xiangjiang Laboratory (No.22XJ01013)

## 1 引言

以石化、冶金等为代表的过程工业在保障国家重大工程建设、带动国民经济发展等方面起着不可替代的作用。然而,这些制造过程多在高温高压、易燃易爆或有毒有害的极端环境下进行,其安全稳定生产面临巨大挑战。工业过程监测是保障工业过程安全稳定生产的关键,对推动“中国制造2025”、“十四五制造业高质量发展”至关重要<sup>[1]</sup>。数据驱动的过程监测因其不依赖于特定的物理模型,具有应用范围广、通用性强、实施成本低等优点,受到国内外研究者的广泛重视<sup>[2,3]</sup>。工业过程数据涵盖了连续型传感数据(比如温度、压力、流量等物理参数)和离散型控制信号<sup>[4]</sup>(如调度决策、设备状态、传感器超限信号等)。高效处理和分析过程数据潜在的统计特性,对于揭示过程生产状态至关重要<sup>[5]</sup>。因此,数据驱动的工业过程监测方法本质上可以归结为显式或隐式的多元统计监测方法。

然而,传统的监测方法往往倾向于在预处理阶段忽略离散变量<sup>[6]</sup>。尤其是当前主流的多元统计过程监测方法<sup>[7]</sup>多聚焦于分析连续变量,并预设了高斯分布的框架。这在很大程度上限制了数据的完整性利用。支持向量机和随机森林等方法虽然能同时处理两类数据,但它们通常是离散变量转化为连续变量进行处理,削弱了离散特征原有的丰富信息,难以保证过程异常检测的全面性和准确性。

离散变量的本质属性决定了其非高斯分布特性,同时工业过程生产原料、质量标准等的变化极易导致连续变量偏离高斯分布,展现出偏斜、多峰、重尾等多模态分布特性。这些特性增加了极端事件发生的风险,超出了正态分布模型所能预测的范围。此外,过程工业的内在复杂性、生产设备和传感器所处环境的恶劣性,常导致在某些操作点出现振荡行为,产生噪声、极值和离群数据,这些极值与离群点可能是指示罕见故障模式或异常工况的重要信息。采用相关离群点检测方法,简单地将它们剔除,虽可避免它们对检测模型的直接干扰,却也可能牺牲宝贵的诊断线索,降低模型的诊断精度和实用性。反之,若保留这些离群数据并直接应用于基于高斯分布假设的检测模型,则可能因模型对异常

值的过度敏感而导致模型过拟合,造成误报频发,最终削弱模型的泛化能力和实际应用效果。

高斯混合判别分析法(Gaussian Mixture Discriminant Analysis, GMDA)<sup>[8]</sup>利用混合高斯模型来增强过程监测模型对非高斯数据的辨识能力。例如,Huang等人<sup>[9]</sup>结合GMDA和统计过程控制图来进行非高斯过程故障诊断。然而,由于混合高斯模型(Gaussian Mixture Model, GMM)只能处理连续数据,该方法在分析和处理离散变量时力不从心。为突破该限制,有研究者探索将连续变量与离散变量分别进行建模的策略。例如,Wang等人<sup>[10]</sup>尝试融合连续和二值变量来提高故障诊断的性能,随后,他们进一步考虑了在多种生产模式下如何利用这两类变量进行异常检测<sup>[11]</sup>,充分展现了连续和离散变量融合分析在复杂工业过程监测中的潜力。

在构建混合辨识模型过程中,期望最大化算法(Expectation-Maximization algorithm, EM)因其在处理含有隐变量的统计分布建模中的卓越性能而被广泛应用。然而,应用EM算法需要考虑如下挑战:(1)避免似然函数带来的奇异性;(2)克服局部最优的陷阱;(3)高迭代次数带来的计算成本。鉴于上述局限,特别是在处理具有非平稳特性的系统时(如原材料成分、强度或尺寸的实时波动),传统的基于GMDA的方法往往需要重新训练,无法及时捕捉统计特性的动态变化<sup>[12]</sup>。为解决上述问题,Pei等人<sup>[13]</sup>引入了变分推断(Variational Inference, VI)估计模型参数,虽在一定程度上缓解了这一问题,但其基于历史数据分析的局限性,使得系统在应对非平稳工况时缺乏自适应能力<sup>[14]</sup>。因而,亟需发展一种能持续学习<sup>[15]</sup>、动态更新的模型框架,自动调整参数配置,确保在面对时变工况检测需求和挑战时,仍能保持优异的检测性能,实现对非平稳过程的有效监测与响应<sup>[16,17]</sup>。

综上所述,当前数据驱动的非高斯、非平稳工业过程监测仍面临两大挑战:(1)如何协同分析连续和离散变量,实现非高斯过程的有效监测;(2)如何打造出可应对时变非平稳工业过程的强效监控系统。针对上述问题,本文提出一种具有持续学习能力的工业过程异常检测框架,即连续和离散变量协同分析的变分贝叶

斯混合判别方法(Continuous and Discrete Variables-Concurrent Analysis-based Variational Bayesian Mixture Discriminant, CDVCA-VBMD). 设计了面向连续型变量的混合学生分布模型和面向离散型变量的混合多项分布模型来刻画过程数据的统计分布特性, 推导了基于变分贝叶斯的统计模型参数优化学习方法, 实现混合变量统计分布特性的最优估计. 有效解决了过程数据非高斯特性难以有效表征的难题, 同时展现出对噪声和离群点的高度鲁棒性, 为复杂数据环境下连续和离散变量统计相关性的精确建模提供了有力工具. 此外, 引入持续学习的思想, 构建了基于自适应遗忘因子的时变过程异常检测方法, 确保监测性能的持续稳定, 为复杂工业过程的安全监控提供了强有力的技术支持.

## 2 基础理论

### 2.1 学生分布

学生分布亦称T分布, 对数据集中的离群点有更强的鲁棒性<sup>[18]</sup>. 单变量的概率密度函数为

$$S(x|\mu, \lambda, \nu) = \frac{\Gamma(\frac{\nu}{2} + \frac{1}{2})}{\Gamma(\frac{\nu}{2})} \left(\frac{\lambda}{\pi\nu}\right)^{\frac{1}{2}} \left[1 + \frac{\lambda(x-\mu)^2}{\nu}\right]^{-\frac{\nu}{2} - \frac{1}{2}} \quad (1)$$

T分布可以通过叠加不同精度的高斯分布获得, 即

$$\begin{aligned} S(x|\mu, A, \nu) &= \int_0^\infty N(x|\mu, (\eta A)^{-1}) G(\eta|\frac{\nu}{2}, \frac{\nu}{2}) d\eta \\ &= \frac{\Gamma(\frac{\nu}{2} + \frac{D}{2})}{\Gamma(\frac{\nu}{2})} \frac{|A|^{\frac{1}{2}}}{(\nu\pi)^{\frac{D}{2}}} \left[1 + \frac{\Delta^2}{\nu}\right]^{-\frac{\nu}{2} - \frac{D}{2}} \end{aligned} \quad (2)$$

其中,  $D$  为变量  $x$  的维度, 且  $\Delta = (x-\mu)^T A(x-\mu)$ .

### 2.2 多项分布

多项分布是二项分布在多元变量上的推广.  $d$  个状态的单一向量  $x$  的分布律为

$$C(x|\lambda) = \prod_{j=1}^d \lambda_j^{x_j} \quad (3)$$

其中,  $\lambda_j$  为变量处于  $j$  状态的概率, 需满足  $0 \leq \lambda_j \leq 1$  且  $\sum_j \lambda_j = 1$ ; 分量  $x^j \in \{0, 1\}$  且  $\sum_{j=1}^d x^j = 1$ .

通过  $M$  次实验, 获得状态  $j$  的次数为  $m_j$  的变量即为多项分布, 其分布律为

$$\text{Mult}(m_1, m_2, \dots, m_d | \lambda, M) = \binom{M}{m_1 m_2 \dots m_d} \prod_{j=1}^d \lambda_j^{m_j} \quad (4)$$

其中,  $\binom{M}{m_1 m_2 \dots m_d} = \frac{M!}{m_1! m_2! \dots m_d!}$ ,  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d)^T$ .

### 2.3 变分推断

变分推断(Variational Inference, VI)<sup>[19]</sup>是找一个易于处理的参数化分布  $q(\theta)$  来近似真实后验分布

$P(\theta|X)$ , 其中  $X = (x_1, \dots, x_N)$  为训练集,  $\theta = (\theta_1, \theta_2, \dots, \theta_K)$  代表其中的隐变量或者模型参数.

为估计出模型参数  $\theta$ , 引入度量变分分布与真实后验分布之间的差异的KL(Kullback-Leibler)散度, 即

$$\text{KL}(q||P) = \ln P(X) - \underbrace{\int q(\theta) \ln \frac{P(\theta, X)}{q(\theta)} d\theta}_{L(q)} \quad (5)$$

由式(5)可知, 由于  $\text{KL}(q||P) \geq 0$  且  $\ln P(X)$  为定值, 最小化  $\text{KL}(q||P)$  等价于最大化变分下界  $L(q)$ <sup>[20]</sup>. 借助于平均场理论<sup>[20]</sup>, 可得参数  $\theta_j$  最佳解的一般形式:

$$\ln \hat{q}(\theta_j) = E_{\theta_{\setminus j}}[\ln P(\theta, X)] + c \quad (6)$$

其中,  $\hat{q}(\theta_j)$  表示  $\theta_j$  的概率估计结果. 由于  $\hat{q}(\theta_j)$  和  $\hat{q}(\theta_i)$  相互耦合, 因此参数优化过程类似于EM算法.

## 3 提出的方法

本节详细介绍本文提出的异常检测方法的核心原理, 并给出相应的变分贝叶斯参数估计方法及在线持续性学习策略.

### 3.1 连续与离散变量的混合统计分布模型构建

CDVCA-VBMD是通过利用所收集的正常数据建立混合判别模型, 实现非高斯、非平稳工业过程异常的准确检测.

设  $X_c \in \mathbb{R}^{N_c \times d}$  是一组工况为  $c$  的  $N_c$  个独立同分布的  $d$  维自然数,  $c=n$  代表正常,  $c=a$  为异常;  $x_c = [x_c^{co}; x_c^{ca}] \in \mathbb{R}^d$  是  $X_c$  中的一个样本, 其中,  $x_c^{co} \in \mathbb{R}^{d_{co}}$  表示连续特征,  $x_c^{ca} \in \mathbb{R}^{d_{ca}}$  表示离散特征, 且  $d = d_{co} + d_{ca}$ .  $x_n$  的类条件分布概率可以采用混合分布模型进行定义, 即

$$P(x_n|n) = \sum_{k=1}^{K_n} \varphi_k P(x_n|n, k) \quad (7)$$

式(7)中,  $K_n$  为分量个数,  $\varphi_k$  为第  $k$  个分量的混合系数, 满足  $\sum_{k=1}^{K_n} \varphi_k = 1$ ,  $P(\cdot|n, k)$  为第  $k$  个分量的概率分布模型.

假定离散变量和连续变量是相对独立的, 允许每种类型的变量具有其自然的统计分布, 则

$$P(x_n|n, k) = P(x_n^{co}|n, k) P(x_n^{ca}|n, k) \quad (8)$$

虽然假定它们是相互独立的, 但研究表明式(7)所示的混合分布模型仍可以捕获到连续和离散变量间的统计相关性<sup>[21]</sup>.

实际工业过程数据中的离散变量不一定是二值变量, 因而选用能有效表述多类别输出的多项分布进行建模. 同时, 对于连续变量而言, 学生分布能够灵活应对非高斯特性问题, 并提高对离群点的鲁棒性. 具体地, 在多元学生分布的积分形式式(2)中, 当  $\eta$  已知时可表示为

$$P(x_n^{co}|n, k) = N(x_n^{co} | \mu_{n,k}, (\eta_n A_{n,k})^{-1}) \quad (9)$$

$$P(\mathbf{x}_n^{ca}|n, k) = \prod_{j=1}^{d_{co}} C(x_n^{ca,j} | \lambda_{n,k,j}) \quad (10)$$

式(10)中,  $C(x_n^{ca,j} | \lambda_{n,k,j})$  是参数为  $\lambda_{n,k,j}$  的多项分布, 对应于有两个以上分类结果的离散特征.

将式(9)和式(10)代入式(8)的结果代入式(7), 即可得到一个对离散变量采用多项分布建模、连续变量采用学生分布建模的新型混合统计分布模型.

### 3.2 混合模型的变分贝叶斯参数估计

#### 3.2.1 参数先验

在 VI 中, 通常采用共轭先验来简化后验分布的计算. 基于式(7)可知, 模型的参数包括  $\lambda, \varphi, \mu, A, \eta$ , 其共轭先验为

$$P(\varphi) = D(\varphi | \alpha_0) \quad (11)$$

$$P(\lambda) = \prod_{k=1}^K \prod_{j=1}^{d_{co}} D(\lambda_{k,j} | \varepsilon_0) \quad (12)$$

$$P(\mu) = \prod_{k=1}^K N(\mu_k | m_0, (\beta_0 A_k)^{-1}) \quad (13)$$

$$P(A) = \prod_{k=1}^K W(A_k | w_0, u_0) \quad (14)$$

$$P(\eta | Z, v) = \prod_{k=1}^K \prod_{n=1}^N G(\eta_{nk} | \frac{v_k}{2}, \frac{v_k}{2})^{\tau_{nk}} \quad (15)$$

其中,  $D(\cdot)$  为 Dirichlet 分布,  $\alpha_0$  和  $\varepsilon_0$  分别为参数  $\varphi$  和  $\lambda$  分布的超参数.  $W(\cdot)$  为 Wishart 分布,  $w_0$  和  $u_0$  是尺度矩阵和自由度数量的超参数.  $G(\cdot)$  为伽马分布, 前后参数  $v_k/2$  分别为反尺度与形状参数, 决定曲线的陡峭程度和形状.

#### 3.2.2 变分贝叶斯参数估计

基于所提出的 CDVCA-VBMD 方法的假设, 若给定权重系数  $\varphi$  进行变分推断来更新模型的参数, 超参数的优化是一个迭代的过程, 因而该过程类似于 EM 算法.

参数估计的 E 步: 给定参数  $\varphi$ , 此时式(6)中的联合概率分布  $\ln P(\theta, X)$  改进为

$$\begin{aligned} \ln \{P(X, Z, \theta | \varphi)\} &= \ln \{P(X | \mu, A, \eta, Z, \lambda)\} \\ &+ \ln \{P(\mu)\} + \ln \{P(\lambda)\} \\ &+ \ln \{P(A)\} + \ln \{P(\eta | Z)\} + \ln \{P(Z | \varphi)\} \end{aligned} \quad (16)$$

其中,  $Z = (z_1, z_2, \dots, z_{N_c})$  代表每条数据代表的隐变量,  $z_n$  是一个“1-of-k”的二值向量, 给定  $\varphi$  时它的分布可以表示为

$$P(Z | \varphi) = \prod_{k=1}^K \prod_{n=1}^N \varphi_k^{\tau_{nk}} \quad (17)$$

式(16)中, 给定参数  $\theta$  和潜变量  $Z$  时,  $P(X | \mu, A, \eta, Z, \lambda)$  可以表示为

$$P(X | \mu, A, \eta, Z, \lambda) = \prod_{k=1}^K \prod_{n=1}^N \quad (18)$$

$$(N(\mathbf{x}_n^{co} | \mu_k, (\eta_{nk} A_k)^{-1}) \times \prod_{j=1}^{d_{co}} C(x_n^{ca,j} | \lambda_{k,j}))^{\tau_{nk}}$$

由式(6)可知, 使用 VI 估计模型参数时, 参数  $Z, \mu, A, \eta, \lambda$  的优化过程也类似于 EM 算法.

在变分 E 步, 将式(12)~(15)和式(17)~(18)代入式(16), 再将其结果代入式(6), 即可得到

$$q^*(Z) = \prod_{k=1}^K \prod_{n=1}^N r_{nk}^{\tau_{nk}} \quad (19)$$

其中,  $r_{nk} = \frac{\rho_{nk}}{\sum_{k=1}^K \rho_{nk}}$ , 且

$$\begin{aligned} \ln \rho_{nk} &= \ln \varphi_k + \frac{1}{2} E[\ln |A_k|] \\ &+ (\frac{d_{co} + v_k}{2} - 1) E[\ln \eta_{nk}] - \frac{d_{co}}{2} \ln 2\pi \\ &- \frac{1}{2} E_{\eta_{nk}} [\eta_{nk}] E_{\mu_k, A_k} [(\mathbf{x}_n - \mu_k)^T A_k (\mathbf{x}_n - \mu_k)] \\ &+ \frac{v_k}{2} \ln \frac{v_k}{2} - \ln \Gamma(\frac{v_k}{2}) - \frac{v_k}{2} E[\eta_{nk}] \end{aligned} \quad (20)$$

在变分 M 步, 通过固定  $r_{nk}$ , 将式(16)的结果代入式(6)以重新计算各个参数的分布, 得到

$$q^*(\mu_k | A_k) = N(\mu_k | m_k, (\beta_k A_k)^{-1}) \quad (21)$$

$$q^*(A_k) = W(A_k | w_k, u_k) \quad (22)$$

$$q^*(\eta) = \prod_{k=1}^K \prod_{n=1}^N G(\eta_{nk} | a_{nk}, b_{nk})^{\tau_{nk}} \quad (23)$$

$$q^*(\lambda) = \prod_{k=1}^K \prod_{j=1}^{d_{co}} D(\lambda_{k,j} | \varepsilon_{k,j}) \quad (24)$$

其中,  $m_k = \beta_k^{-1} (\beta_0 m_0 + \sum_{n=1}^N E_{z_{nk}, \eta_{nk}} [z_{nk} \eta_{nk}] \mathbf{x}_n)$ ,  $b_{nk} =$

$$\begin{aligned} &\frac{r_{nk}}{2} (E_{\mu_k, A_k} [(\mathbf{x}_n - \mu_k)^T A_k (\mathbf{x}_n - \mu_k)] + v_k), \quad \beta_k = \beta_0 + \\ &\sum_{n=1}^N E_{z_{nk}, \eta_{nk}} [z_{nk} \eta_{nk}], \quad a_{nk} = \frac{r_{nk}}{2} (d_{co} + v_k), \quad w_k^{-1} = w_0^{-1} + \\ &\sum_{n=1}^N E_{z_{nk}, \eta_{nk}} [z_{nk} \eta_{nk}] (\mathbf{x}_n - \mu_k) (\mathbf{x}_n - \mu_k)^T + \beta_0 (\mu_k - m_0) (\mu_k - \\ &m_0)^T - \beta_k (\mu_k - m_k) (\mu_k - m_k)^T, \quad u_k = u_0 + N_k, \quad \varepsilon_{k,j} = \varepsilon_0 + N_k \bar{x}_k^{ca,j} \\ &\text{且 } N_k = \sum_{n=1}^N r_{nk}, \quad E[z_{nk}] = r_{nk}, \quad \bar{x}_k^{ca,j} = \frac{1}{N_k} \sum_{n=1}^N r_{nk} x_n^{ca,j}. \end{aligned}$$

由参数所服从的威沙特分布、伽马分布和狄利克雷分布的性质可得

$$E_{A_k} [\ln |A_k|] = \sum_{i=1}^{d_{co}} \Psi((v_k + 1 - i)/2) + d_{co} \ln 2 + \ln |w_k| \quad (25)$$

$$\begin{aligned} &E_{\mu_k, A_k} [(\mathbf{x}_n - \mu_k)^T A_k (\mathbf{x}_n - \mu_k)] \\ &= d_{co} \beta_k^{-1} + v_k (\mathbf{x}_n - m_k)^T w_k (\mathbf{x}_n - m_k) \end{aligned} \quad (26)$$

$$E_{\eta_{nk}} [\ln \eta_{nk}] = \Psi(a_{nk}) - \ln b_{nk} \quad (27)$$

$$E_{\eta_{nk}}[\eta_{nk}] = a_{nk}/b_{nk} \quad (28)$$

变分优化过程是迭代进行的,在变分E步将当前状态下参数概率分布性质的所得式(25)~(28)代入式(20)得到 $r_{nk}$ .变分M步中固定 $r_{nk}$ 的值,利用式(21)~(24)更新参数的变分分布.如此循环迭代至 $L(q)$ 的相对变化小于预设的阈值或参数收敛,得到参数 $\mathbf{Z}, \boldsymbol{\mu}, \mathbf{A}, \boldsymbol{\eta}, \boldsymbol{\lambda}$ 的最优近似分布.此时 $L(q)$ 为

$$L(q) = E_{\boldsymbol{\mu}, \mathbf{A}, \boldsymbol{\eta}, \mathbf{Z}}[\ln P(\mathbf{X}, \mathbf{Z}, \boldsymbol{\theta}|\boldsymbol{\varphi}) - \ln q(\boldsymbol{\theta})] \quad (29)$$

### 3.2.3 超参数优化

参数估计的M步,通过最大化 $L(q)$ 以重新估计参数 $\boldsymbol{\varphi}$ .

其中当考虑 $\boldsymbol{\varphi}$ 的先验分布式(11)时, $\boldsymbol{\varphi} = \operatorname{argmax}_{\boldsymbol{\varphi}} (\ln P(\boldsymbol{\varphi}) + L(q))$ ,故更新规则为

$$\varphi_k^i = \frac{N_k + \alpha_0}{N + K\alpha_0} \quad (30)$$

迭代执行变分E步与变分M步,当参数收敛时停止,得到最大化当前变分下界的最优参数估计.在实际参数估计中,一般预先设置较多的分量数和阈值如 $10^{-4}$ ,随着训练的进行,混合权重值小于阈值的分量被自适应丢弃,自动收敛到分量的实际数量.因此,CDVCA-VBMD可获得刻画实际数据分布的最优模型,详细的参数估计算法流程见算法1.

#### 算法1 混合分布模型的变分贝叶斯参数估计

输入:超参数: $\varepsilon_{\theta}, \alpha_{\theta}, \mathbf{m}_{\theta}, \boldsymbol{\beta}_{\theta}, \boldsymbol{\omega}_{\theta}, \mathbf{u}_{\theta}, \mathbf{v}_{\theta}, \boldsymbol{\varphi}$ ,训练集: $\mathbf{X}_n \in \mathbb{R}^{N_s \times d}$

输出:最优的参数估计分布

- 1: FOR 参数未收敛 DO
- 2: E步:给定参数 $\boldsymbol{\varphi}$ 的初始值;
- 3: FOR 变分下界式(29)的变化值大于预设的阈值 DO
- 4: 变分E步:由当前状态下模型参数概率分布的性质所得式(25)~(28)计算式(20),得到 $r_{nk}$ 和参数 $\mathbf{Z}$ 的分布;
- 5: 变分M步:由式(21)~(24)重新得到参数 $\boldsymbol{\mu}, \mathbf{A}, \boldsymbol{\eta}, \boldsymbol{\lambda}$ 的分布;
- 6: END FOR
- 7: M步:由式(30)更新参数 $\boldsymbol{\varphi}$ 的值;
- 8: 若混合分量的权重系数低于预设阈值,则相应的混合分量将被自适应丢弃,得到最优的混合分布模型
- 9: END FOR

### 3.3 异常检测

借助于训练集 $\mathbf{X}_n \in \mathbb{R}^{N_s \times d}$ 估计出正常类的先验概率 $P(n) = 1$ 和式(7)所定义的条件概率 $P(\mathbf{x}|n)$ ,则观测到新样本 $\mathbf{x}_{\text{new}}$ 的统计指标记为每个分量的似然度(Component Likelihood Indicator, CLI),定义如下:

$$\text{CLI}(\mathbf{x}) = -\ln(P(\mathbf{x}|n)P(n)) \quad (31)$$

样本的CLI值越小,该样本属于正常类的可能性就越大.为有效地监测出故障,经实验验证以CLI值的97.5%分位数作为阈值,若当前数据的CLI值高于阈

值,则发生了故障.

### 3.4 在线持续性学习策略

出于对现实工业场景中非平稳特性的考虑,本文引入了一种带自适应遗忘因子的面向复杂工业过程鲁棒检测的持续学习方法.其中遗忘因子 $\alpha$ 用于控制记忆衰减速度,使得模型能够在保留历史知识的同时,灵活地适应新的数据分布.

定义新数据 $\mathbf{x}_n$ 属于某一分量的概率为 $\delta_{nk} = \frac{\varphi_k S(\mathbf{x}_n^{co}|\boldsymbol{\mu}_k, \mathbf{A}_k, \eta_k)}{\sum_{j=1}^K \varphi_j S(\mathbf{x}_n^{co}|\boldsymbol{\mu}_j, \mathbf{A}_j, \eta_j)}$ ,则训练新任务时参数的更新策略为

$$\boldsymbol{\varphi}_{k, \text{new}} = \frac{\alpha}{N} \left( \sum_{n=1}^N \delta_{nk} \right) + \boldsymbol{\varphi}_{k, \text{old}} (1 - \alpha) \quad (32)$$

$$\boldsymbol{\mu}_{k, \text{new}} = \alpha \left( \sum_{n=1}^N \delta_{nk} \mathbf{x}_n^{co} \right) / \left( \sum_{n=1}^N \delta_{nk} \right) + \boldsymbol{\mu}_{k, \text{old}} (1 - \alpha) \quad (33)$$

$$\eta_{k, \text{new}} \mathbf{A}_{k, \text{new}} = \alpha \left( \sum_{n=1}^N \delta_{nk} (\mathbf{x}_n^{co} - \boldsymbol{\mu}_{k, \text{old}})(\mathbf{x}_n^{co} - \boldsymbol{\mu}_{k, \text{old}})^T \right) / \left( \sum_{n=1}^N \delta_{nk} \right) + \eta_{k, \text{old}} \mathbf{A}_{k, \text{old}} (1 - \alpha) \quad (34)$$

$$\boldsymbol{\lambda}_{j, \text{new}} = \alpha \left( \sum_{n=1}^N \delta_{nk} \mathbf{x}_n^{ca, j} \right) / \left( \sum_{n=1}^N \delta_{nk} \right) - (1 - \alpha) \boldsymbol{\lambda}_{j, \text{old}} \quad (35)$$

此外,根据新状态下数据的特性和模型性能反馈,能自动调整 $\alpha$ 以实现更加精细化的参数更新.

(1)经验分布函数(Empirical Distribution Function, EDF):当新数据到来时,分别计算新旧数据集的EDF,定义为

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \quad (36)$$

其中, $X_i$ 是样本集中的第*i*个观测值, $I$ 为指示函数,当 $X_i \leq x$ 时取值为1,否则为0.

(2)KS(Kolmogorov-Smirnov)测试:使用KS测试比较新数据与历史数据的分布,测试结果*P*值为

$$P = \sup_x |F_n(x) - F_m(x)| \quad (37)$$

(3) $\alpha$ 的调整策略:如果*P*值超过预设的阈值, $\alpha$ 增加0.05以给予新数据更多的权重,并更新模型的参数以适应当前实际数据分布.否则减小0.05以平滑输出,提高模型预测的稳定性.

### 3.5 总体流程

本文提出的基于CDVCA-VBMD的工业过程异常检测主要包括离线建模与在线检测两个阶段.

#### (1)离线建模

步骤1. 收集数据:收集训练数据 $\mathbf{X}_n \in \mathbb{R}^{N_s \times d}$ .

步骤2. 模型参数估计:用 $\mathbf{X}_n$ 基于CDVCA-VBMD

方法估计出式(7)所定义的正常类条件分布  $P(\cdot|n)$ .

步骤3. 设置控制限:使用统计指标 CLI 的 97.5% 分位数作为控制限.

#### (2) 异常检测与模型在线更新

步骤1. 获取新样本:收集新的样本  $\mathbf{x}_{\text{new}}$ , 由式(31) 计算出其统计指标  $\text{CLI}(\mathbf{x}_{\text{new}})$ .

步骤2. 异常检测:若  $\text{CLI}(\mathbf{x}_{\text{new}})$  大于控制限, 则将样本标记为异常, 否则样本正常.

步骤3. 触发模型更新:根据式(37)计算  $P$  值, 当  $P$  值大于阈值时, 触发参数更新, 以应对非平稳特性的影响.

## 4 实验验证与结果讨论

本节详细介绍所提方法在数值仿真系统与经典的田纳西·伊斯曼(Tennessee Eastman, TE)过程上的实验结果.

### 4.1 数值仿真系统实验结果分析

#### 4.1.1 模型介绍

受文献[22]的启发, 本文首先采用如下数值仿真实例进行实验验证:

$$\mathbf{y}^* = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} -0.1681 & 0.2870 & -0.2835 \\ 0.4354 & 0.3812 & 0.1455 \\ 0.0247 & -0.0235 & 0.4096 \\ -0.1173 & -0.1763 & 0.4382 \\ 0.0825 & 0.1398 & 0.3204 \\ -0.3825 & 0.1250 & 0.4836 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} + \varepsilon \quad (38)$$

式中  $y_i | i \in \{1, 2, \dots, 6\}$  代表模型的观测变量,  $s_i | i \in \{1, 2, 3\}$  为随机生成的服从高斯或伽马分布的潜变量,  $\varepsilon$  为高斯噪声. 在生成的数据中随机选择 30 个样本为其添加随机扰动形成离群点. 并随机选择三个变量进行 Z-Score 标准化处理, 计算其 0%、14.3%、28.6%、42.9%、57.1%、71.4% 和 100% 分位数以确定区间边界, 最后根据数据点的值分配相应的区间得到离散数据. 过程产生共 3 000 组样本作为初始训练集数据  $\mathbf{X}_{3000}^6$ .

引入故障信号构造仿真故障数据, 定义如下:

$$\mathbf{X}_f = \mathbf{X} + f\mathbf{e}_p \quad (39)$$

式(41)中,  $f$  为故障信号,  $p$  为可能发生故障的变量,  $\mathbf{e}_p = (0, 0, \dots, 1, \dots, 0)$  为故障方向向量.

在第 2 001 个样本处引入故障信号构成测试集, 记为 No.1~No.6. 其中 No.1 是在变量  $\mathbf{x}_1$  中引入服从  $\text{Uniform}(0, 2)$  的噪声故障; No.2 是在  $\mathbf{x}_2$  中引入幅值为 10 的阶跃故障; No.3 是在  $\mathbf{x}_3$  中引入幅值为 3 的正弦故障; No.4 是在  $\mathbf{x}_2$  中引入指数故障; No.5 是在  $\mathbf{x}_1$  和  $\mathbf{x}_2$  中分别引入幅值为 10 的阶跃故障和指数形式缓变故障; No.6 是在  $\mathbf{x}_2$  和  $\mathbf{x}_3$  中分别引入  $\text{Uniform}(0, 1)$  的噪声故障和幅值为 2 的正弦信号故障.

#### 4.1.2 实验设计与性能评价指标

为充分检验 CDVCA-VBMD 的有效性, 首先验证其鲁棒性以及完整类型数据集上的性能, 然后对方法中不同模块对系统整体性能的影响进行消融研究. 实验选用准确率(Accuracy, Acc)、召回率(Recall, Rec)、 $F_1$  值、误报率(False Positive Rate, FPR)和接收者操作特性曲线(ROC)评估模型的监测能力.

对比方法包括: 单类支持向量机(One-Class SVM, OCSVM)<sup>[23]</sup>、孤立森林(Isolation Forest, IF)<sup>[24]</sup>、自编码器(AutoEncoder, AE)<sup>[25]</sup>方法. 其中, OCSVM 通过找最优的分割超平面来最大化与正常点的距离, 从而隔离异常点; 在基于随机森林的 IF 中, 异常为分布稀疏且离密度高的群体较远的点; 基于神经网络的 AE 通过学习数据的有效编码以重构输入数据, 误差较大的点被视为异常.

#### 4.1.3 验证性实验

本文使用含有离群点的训练集连续数据进行模型鲁棒性验证实验. 图 1 为连续数据的散点图, 其中红色数据点为离群点, 其他数据点代表初始数据. 图 2 显示了采用混合学生模型与混合高斯模型的建模结果. 实验中采用了不同的初始  $K$  值来训练 GMM 和混合学生模型.

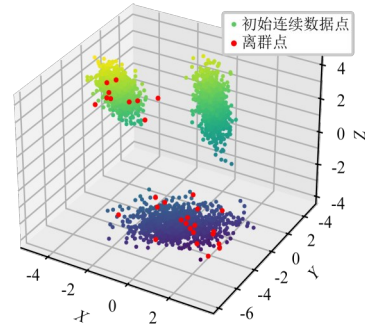


图 1 连续数据的散点图

从图 2(a)、图 2(b)可以看出, 由于 GMM 对于离群点较为敏感, 它在拟合正常数据分布和离群点之间做出妥协, 将离群数据点拟合到不同的分量中, 导致模型变得更加复杂, 拟合质量下降. 相比之下, 图 2(c)、图 2(d)表明, 混合学生分布模型对离群点具有较强的鲁棒性, 模型拟合效果更好. 实验表明, 混合学生分布提供了一种类似的高斯和非高斯分布之间过渡的模型, 更能适应数据的实际情况, 具有较好的鲁棒性与灵活性.

图 3(a)为离散数据散点图. 图 3(b)、图 3(c)显示了该训练集的混合多项分布模型的建模结果, 其中图 3(c)为建模得到的各类的概率. 由于多项分布是用来描述一系列独立试验的结果, 不会将单个数据点分配给特定的分量, 但为了可视化, 使用各类别的估计概率为数据点着色, 直观的分类散点图如图 3(b)所示.

图 4 显示了 CDVCA-VBMD 方法在 No.1 的连续、离

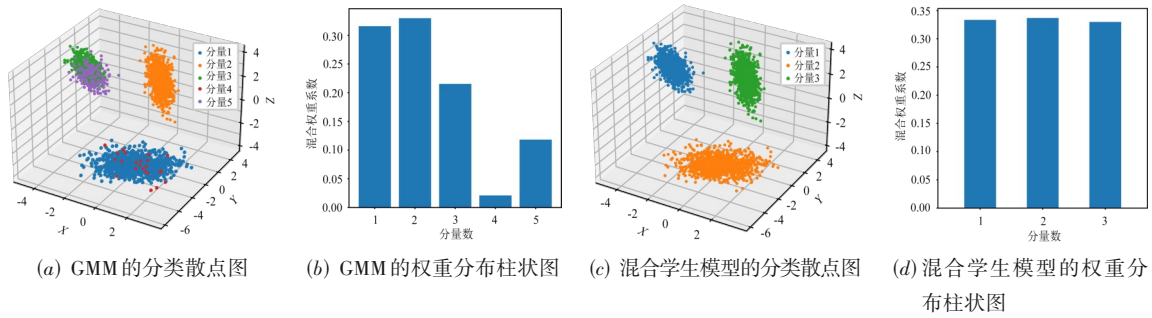


图2 两类模型在训练集连续数据上的拟合效果

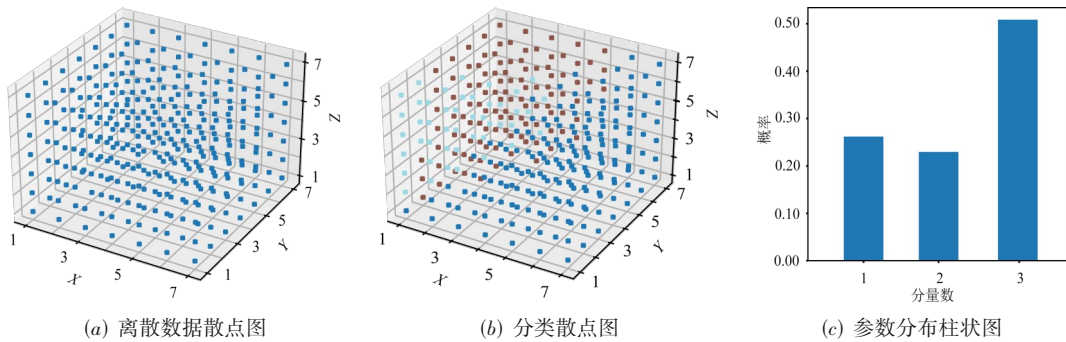


图3 混合多项分布模型在离散数据上的拟合效果

散变量和完整数据集上的 ROC 曲线.  $CDVCA-VBMD_{co}$  模式的 AUC 值相对  $CDVCA-VBMD_{all}$  模式下较低, 因为其忽视了离散变量所包含的判别信息. 而通过结合连续和离散变量,  $CDVCA-VBMD_{all}$  能更全面地捕捉数据的特性. 同时, 表 1 也显示  $CDVCA-VBMD_{all}$  下的  $F_1$  值最高. 实验结果表明同时考虑连续和离散变量可以提高模型的整体性能, 帮助识别出更多样化的异常情况. 上

述结果验证了同时考虑两类变量这一思想的有效性以及本文所提方法的鲁棒性.

表 1 三种模式下的  $F_1$  值与训练时间

	$CDVCA-VBMD_{co}$	$CDVCA-VBMD_{ca}$	$CDVCA-VBMD_{all}$
threshold	-4.20	-5.94	-7.32
$F_1$	0.95	0.96	0.98
Training time/s	0.23	0.29	0.62

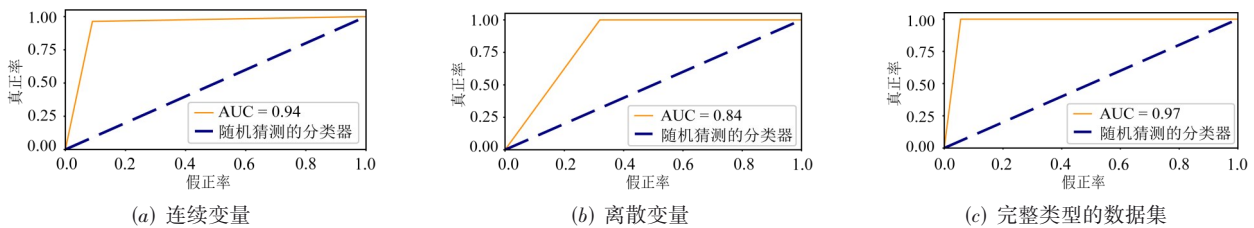


图4 本文方法在三种模式下的 ROC 曲线图

#### 4.1.4 消融实验

为探讨本文考虑到离散变量并对离散多变量建立混合多项分布模型、对连续变量建立混合学生模型对系统整体性能的影响, 在数值仿真数据集上进行了消融研究. 首先在传统的使用 GMM 对连续变量建模的基础上, 增加对离散变量建模模块, 命名为 GMM+C. 其次, 将 GMM 替换为混合学生模型, 命名为 ST+C.

由表 2 可以看出, 有效利用离散数据所包含的信息可以提高故障检测的准确度并降低误报率. 在此基础

上, 将 GMM 替换为混合学生模型对连续数据建模, 此时各个指标的值均得到了提高, 证明所提出的融合连续变量的混合学生模型和离散变量的混合多项分布的模型在故障检测方面的优势.

#### 4.1.5 对比实验

为进一步验证所提方法的有效性, 选取三种常用的无监督故障监测方法作为对比方法, 在完整数据集上进行实验. 各参数通过交叉验证进行优化, OCSVM 使用径向基函数核, AE 使用 2 个隐藏层, 每层 10 个节

表 2 在所有类型故障数据上的消融实验

Method	Acc	Rec	$F_1$	FPR
GMM	0.958	0.965	0.968	0.054
GMM+C	0.962	0.965	0.971	0.043
ST+C	0.968	0.971	0.977	0.037

点, 阈值设置为重构误差分布的95%分位数.

表3总结了这些方法在各种故障类型上的检测效果. 结果表明, CDVCA-VBMD始终保持较高的 $F_1$ 值和较低FPR, 显示了强大的故障检测能力. 为便于对比, 图5给出了CDVCA-VBMD与其他方法在本文方法检测效果稍差的No.3上的ROC曲线图, 其中CDVCA-VBMD的AUC值最大, 为0.95, 表明它有更突出的区分能力. 此外, 不难发现OCSVM方法平均误报率最低, 但其 $F_1$ 值远低于本文方法, 且AUC值只有0.88, 实验结果再次证明了本文所提方法的有效性.

表 3 不同方法在多种故障类型上的 $F_1$ 值与误报率FPR

Fault No.	OCSVM $F_1$ /FPR	IF $F_1$ /FPR	AE $F_1$ /FPR	CDVCA-VBMD $F_1$ /FPR
1	0.87/0.007	0.78/0.081	0.96/0.60	0.97/0.03
2	0.95/0.00	0.83/0.00	0.79/0.00	0.98/0.00
3	0.95/0.00	0.85/0.00	0.93/0.28	0.95/0.15
4	0.94/0.00	0.83/0.00	0.92/0.32	0.98/0.00
5	0.95/0.00	0.83/0.00	0.98/0.07	0.98/0.00
6	0.94/0.00	0.81/0.00	0.95/0.19	0.97/0.04
Average	0.93/0.001	0.83/0.014	0.92/0.243	0.97/0.036

## 4.2 TE过程实验结果分析

TE过程是基于一个真实工业过程的仿真, 为评价过程监控方法提供了一个现实的工业过程. 该过程有52个连续变量, 先排除对故障监测无影响的2个常量. 鉴于变量的分布在正常情况下是接近高斯的, 为构建合适的实验场景使用3-sigma规则获取离散变量. 具体地, 选择其中28个变量, 将每个变量的上下界分别定义

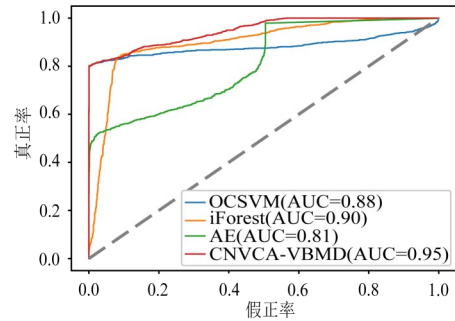


图 5 各类方法在No.3上的ROC曲线图

为均值加减3倍的标准差, 根据变量值在边界内、大于上界或小于下界将该变量转换为0、1和2. 由此, 每个样本中便有28个离散变量和22个连续变量.

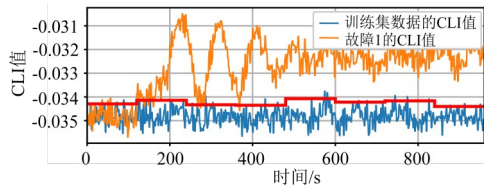
使用25 h仿真下获得的正常工况的样本<sup>[26]</sup>为训练集的主要部分, 可将48 h仿真下获得的正常样本加入训练集以增加正常数据的多样性. 并以48 h仿真下获得的带有故障的数据作为测试集.

表4给出了各个方法在不同类型故障数据上的 $F_1$ 与FPR值, 其中CDVCA-VBMD的 $F_1$ 值在第3、9和15类故障上低于AE方法, 但整体上CDVCA-VBMD的 $F_1$ 值较高, 且与其他方法的差值较为明显. 从FPR的角度来看, 尽管IF方法的平均FPR值低于本文方法, 但差值极小, 而本文方法的 $F_1$ 值却远高于IF. 总体上再次验证了CDVCA-VBMD在故障监测问题上的有效性.

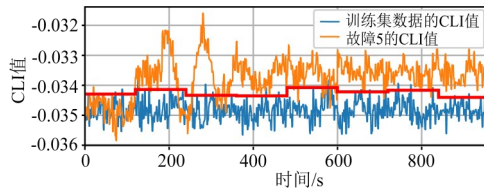
随着时间的推移运行仿真可以得到不同故障类型随时间变化的数据集. 图6给出了CDVCA-VBMD方法在其中两种故障数据集上的检测效果图. 可以看出当条件变化时触发了模型更灵活地更新, 而不需要重新进行训练, 此时模型的阈值(红色折线)也可能会发生改变以适应新的数据分布, 利用更新后的模型能够更精确地检测故障, 展现了模型对数据分布变化的动态适应能力.

表 4 CDVCA-VBMD与其他方法在不同故障类型数据上的 $F_1$ 值与误报率

Fault No.	OCSVM $F_1$ /FPR	IF $F_1$ /FPR	AE $F_1$ /FPR	CDVCA-VBMD $F_1$ /FPR	Fault No.	OCSVM $F_1$ /FPR	IF $F_1$ /FPR	AE $F_1$ /FPR	CDVCA-VBMD $F_1$ /FPR
1	0.58/0.04	0.95/0.00	0.75/0.48	0.99/0.05	12	0.75/0.07	0.97/0.01	0.65/0.43	0.98/0.01
2	0.90/0.03	0.97/0.00	0.90/0.46	0.98/0.05	13	0.89/0.01	0.96/0.01	0.57/0.31	0.97/0.05
3	0.21/0.08	0.08/0.02	0.63/0.36	0.29/0.13	14	0.17/0.07	0.06/0.00	0.63/0.52	0.99/0.08
4	0.21/0.09	0.07/0.00	0.65/0.43	0.99/0.05	15	0.21/0.05	0.09/0.00	0.62/0.51	0.33/0.07
5	0.36/0.09	0.41/0.00	0.63/0.43	0.99/0.05	16	0.28/0.11	0.24/0.10	0.61/0.48	0.94/0.23
6	0.96/0.03	0.98/0.04	0.92/0.56	0.99/0.03	17	0.23/0.05	0.31/0.00	0.66/0.45	0.96/0.08
7	0.44/0.06	0.63/0.00	0.66/0.48	0.99/0.05	18	0.83/0.05	0.93/0.00	0.86/0.55	0.95/0.06
8	0.76/0.05	0.97/0.00	0.65/0.49	0.98/0.05	19	0.20/0.08	0.02/0.00	0.66/0.36	0.97/0.05
9	0.23/0.08	0.05/0.06	0.65/0.45	0.27/0.23	20	0.26/0.03	0.43/0.00	0.54/0.05	0.93/0.06
10	0.32/0.05	0.53/0.00	0.64/0.53	0.93/0.05	21	0.32/0.03	0.27/0.00	0.36/0.52	0.76/0.21
11	0.20/0.07	0.19/0.01	0.65/0.51	0.84/0.09	Average	0.45/0.06	0.52/0.01	0.66/0.44	0.86/0.08



(a) 故障1的检测效果图



(b) 故障5的检测效果图

图6 两类故障的检测效果图

## 5 结论

本文提出了一种针对非平稳非高斯复杂工业过程的鲁棒的故障监测方法——连续与离散变量协同分析的变分贝叶斯混合判别方法. 构建了连续与离散变量的有限混合模型, 推导了基于变分贝叶斯的混合模型参数最优估计方法, 实现了两类数据在概率框架下的有效融合, 增强了模型对复杂数据的适应能力. 由于所提方法内置了对离群点效应的鲁棒处理机制, 有效缓解了异常值对参数估计精度的负面影响, 同时, 该方法集成了在线学习机制, 使其能持续地从实时数据中汲取新知, 从而有效适应各种复杂多变的工业场景. 采用数值仿真系统和经典的TE过程进行了实验验证, 结果表明所提方法具有充分利用连续与离散两类数据信息优势方面的能力, 显著提升了故障监测的精度与效率.

## 参考文献

- [1] ALDRINI J, CHIHI I, SIDHOM L. Fault diagnosis and self-healing for smart manufacturing: A review[J]. *Journal of Intelligent Manufacturing*, 2024, 35(6): 2441-2473.
- [2] MELO A, CÂMARA M M, PINTO J C. Data-driven process monitoring and fault diagnosis: A comprehensive survey[J]. *Processes*, 2024, 12(2): 251.
- [3] 杨明, 周斯维, 谷健. 基于数据驱动的城市居民出行OD预测[J]. *长沙理工大学学报(自然科学版)*, 2022, 19(1): 73-80.  
YANG M, ZHOU S W, GU J. Traffic OD prediction of urban residents based on data driven[J]. *Journal of Changsha University of Science & Technology (Natural Science)*, 2022, 19(1): 73-80. (in Chinese)
- [4] 柴蓉, 陈米铃, 李锦红. 基于效用优化的星地融合网络联合用户关联及资源块调度算法[J]. *电子学报*, 2023, 51(12): 3483-3495.

- CHAI R, CHEN M L, LI J H. Utility optimization-based joint user association and resource allocation algorithm for integrated satellite-terrestrial network[J]. *Acta Electronica Sinica*, 2023, 51(12): 3483-3495. (in Chinese)
- [5] FENG L J, ZHAO C H. Fault description based attribute transfer for zero-sample industrial fault diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(3): 1852-1862.
- [6] WANG M, ZHOU D H, CHEN M Y, et al. Anomaly detection in the fan system of a thermal power plant monitored by continuous and two-valued variables[J]. *Control Engineering Practice*, 2020, 102: 104522.
- [7] CHEN J H, ZHAO C H. Exponential stationary subspace analysis for stationary feature analytics and adaptive non-stationary process monitoring[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(12): 8345-8356.
- [8] HASTIE T, TIBSHIRANI R. Discriminant analysis by Gaussian mixtures[J]. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996, 58(1): 155-176.
- [9] HUANG C C, CHEN T, YAO Y. Mixture discriminant monitoring: A hybrid method for statistical process monitoring and fault diagnosis/isolation[J]. *Industrial & Engineering Chemistry Research*, 2013, 52(31): 10720-10731.
- [10] WANG M, SHENG L, ZHOU D H, et al. A feature weighted mixed naive Bayes model for monitoring anomalies in the fan system of a thermal power plant[J]. *CAA Journal of Automatica Sinica*, 2022, 9(4): 719-727.
- [11] WANG M, ZHOU D H, CHEN M Y. Adjustable multi-mode monitoring with hybrid variables and its application in a thermal power plant[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(2): 1425-1435.
- [12] 李友云, 陈佳, 王硕, 等. 沥青混合料劈裂疲劳损伤力学特性多尺度研究[J]. *长沙理工大学学报(自然科学版)*, 2022, 19(1): 1-11, 80.  
LI Y Y, CHEN J, WANG S, et al. Multi-scale research on mechanical characteristics of splitting fatigue damage of asphalt mixture[J]. *Journal of Changsha University of Science & Technology (Natural Science)*, 2022, 19(1): 1-11, 80. (in Chinese)
- [13] 裴洪, 司小胜, 胡昌华, 等. 零寿命标签下退化设备剩余寿命预测方法[J]. *电子学报*, 2023, 51(7): 1939-1948.  
PEI H, SI X S, HU C H, et al. Prognostic method of remaining useful life for degraded equipment under zero life label[J]. *Acta Electronica Sinica*, 2023, 51(7): 1939-1948. (in Chinese)
- [14] 刘金平, 王杰, 唐朝晖, 等. 基于自适应滑窗递归稀疏主成分分析的工业过程故障监测[J]. *电子学报*, 2020, 48(9): 1795-1803.  
LIU J P, WANG J, TANG Z H, et al. Industrial process

- fault monitoring based on adaptive sliding window-recursive sparse principal component analysis[J]. *Acta Electronica Sinica*, 2020, 48(9): 1795-1803. (in Chinese)
- [15] ZHAO C H. Perspectives on nonstationary process monitoring in the era of industrial artificial intelligence[J]. *Journal of Process Control*, 2022, 116: 255-272.
- [16] 蒋晓玲, 吴映波, 陈蒙, 等. 基于跨域结构保持投影的异构在线多源迁移学习方法[J]. *电子学报*, 2023, 51(8): 1983-1994.
- JIANG X L, WU Y B, CHEN M, et al. Heterogeneous online multi-source transfer learning with cross-domain structure preserving projection[J]. *Acta Electronica Sinica*, 2023, 51(8): 1983-1994. (in Chinese)
- [17] 柴铮, 汪嘉业, 赵春晖, 等. 面向工业监控典型监督任务的深度迁移学习方法: 现状、挑战与展望[J]. *中国科学: 信息科学*, 2023, 53(5): 821-840.
- CHAI Z, WANG J Y, ZHAO C H, et al. Deep transfer learning methods for typical supervised tasks in industrial monitoring: State-of-the-art, challenges, and perspectives [J]. *Scientia Sinica (Informationis)*, 2023, 53(5): 821-840. (in Chinese)
- [18] DENG H Q, LI X Y. Anomaly detection via reverse distillation from one-class embedding[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 9737-9746.
- [19] BLEI D M, KUCUKELBIR A, MCAULIFFE J D. Variational inference: A review for statisticians[J]. *Journal of the American Statistical Association*, 2017, 112(518): 859-877.
- [20] BISHOP C M. *Pattern Recognition and Machine Learning* [M]. New York: Springer, 2006.
- [21] CHEN J H, ZHAO C H, DING J L. A flexible probabilistic framework with concurrent analysis of continuous and categorical data for industrial fault detection and diagnosis [J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(10): 10578-10590.
- [22] ALCALA C F, QIN S JOE. Analysis and generalization of fault diagnosis methods for process monitoring[J]. *Journal of Process Control*, 2011, 21(3): 322-330.
- [23] SIREGAR S M, PURWANTO Y, WIBOWO S A. Enhancing network anomaly detection with optimized one-class SVM (OCSVM)[C]//2023 3rd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA). Piscataway: IEEE, 2023: 84-88.
- [24] AHMED M R, MYO T, BAROOMI B AL. Outlier detection to secure wireless sensor networks based on iForest [C]//2022 10th International Symposium on Digital Forensics and Security (ISDFS). Piscataway: IEEE, 2022: 1-6.
- [25] MICHELUCCI U. An introduction to autoencoders[EB/

OL]. (2022-01-11)[2024-04-12]. <https://doi.org/10.48550/arXiv.2201.03898>.

- [26] REINARTZ C, KULAHCI M, RAVN O. An extended Tennessee Eastman simulation dataset for fault-detection and decision support systems[J]. *Computers & Chemical Engineering*, 2021, 149: 107281.

#### 作者简介



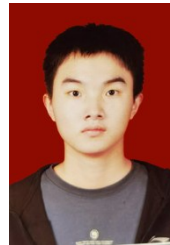
蔡美玲 女, 1982年3月出生于湖南省岳阳市. 现为湖南师范大学信息科学与工程学院副教授. 主持、参与国家和省部级科研课题10余项. 在国内外发表学术论文20余篇.

E-mail: Cai.Meiling@hunnu.edu.cn



罗迪 女, 2001年9月出生于河南省南阳市. 现为湖南师范大学软件工程专业硕士研究生, 主要研究方向为模式识别、工业故障检测与诊断. 中国电子学会会员编号: E190158141A.

E-mail: wsld@hunnu.edu.cn



肖敬日 男, 2002年4月出生于湖南省邵阳市. 现为湖南师范大学软件工程专业硕士研究生. 主要研究方向为智能故障诊断.

E-mail: 201930222056@hunnu.edu.cn



李婧妍 女, 1998年2月出生于河南省安阳市. 目前在湖南师范大学数学与统计学院攻读博士学位. 主要研究方向为工业过程智能监测.

E-mail: jingyanlili@163.com



刘金平 男, 1983年9月出生于湖南省邵阳市. 现为湖南师范大学信息科学与工程学院教授、博士生导师. 主持、参与国家和省部级科研课题10余项. 发表国际高水平SCI期刊论文60余篇, 中文权威期刊论文20余篇, 申请国家发明专利40余项, 已授权20项.

E-mail: lip@phunnu.edu.cn