

活性污泥工艺丝状污泥膨胀的预测维护^①

刘 坚¹ 袁凌玲² 黄道平² 刘乙奇² 于广平¹

(1. 广州中国科学院沈阳自动化研究所分所; 2. 华南理工大学自动化科学与工程学院)

摘 要 基于活性污泥工艺污水处理的稳定运行常常受到丝状污泥膨胀的影响。为了对丝状污泥膨胀进行有效的维护,利用峰度和典型相关分析的方法对传感器数据进行易于测量特征提取,同时使用基于ARMA模型的软测量仪表对难以测量的特征进行提取。这种交互信息被融入于故障检测、传播识别、因果关系分析、剩余使用寿命(RUL)预测和维护,从而实现污水处理丝状污泥膨胀维护的一体化解决方案。结果表明:笔者提出的方法能够检测、定位和跟踪丝状污泥膨胀,甚至预测污泥膨胀的寿命并建议维护策略。

关键词 故障诊断 活性污泥 丝状污泥膨胀 软测量

中图分类号 TP29 **文献标识码** A **文章编号** 1000-3932(2019)06-0458-06

活性污泥法(ASP)是从废水中去除有机物质最常用的工艺^[1]。ASP的稳定和可靠运行主要取决于生物反应器中的有效生物转化和二沉池中的正常泥水分离。二沉池中丝状细菌的过度生长是一个常见问题,常常导致污泥沉降性变差、操作性能差和处理成本高,这通常被称为丝状污泥膨胀。据考证,全球50%以上基于ASP的污水处理厂存在着严重的污泥膨胀问题^[2]。基于视情维护(Condition-based Maintenance)是一种常用的维护策略,可以实现系统早期故障的发现、实时诊断和故障预测^[3]。视情维护是改善污泥膨胀的一种行之有效的解决方案^[4]。丝状污泥膨胀主要是由二次澄清池中细菌的不平衡引起的,同时呈现出了漂移的特性。为了全局洞察丝状污泥膨胀的状态,全流程往往遍布多个传感器,这不可避免地增加了数据的量,同时也带来了数据的冗余和不相关。统计过程监控,包括主成分分析(PCA)、独立分量分析(ICA)和典型相关分析(CCA)^[5],常常用于识别和感知故障特征。通常来说,这些方法通过探索数据来建立经验模型,然后将导出的经验模型作为后续判断故障的参考坐标,此后通过Hotelling T^2 或平方预测误差(SPE)证明输入数据的期望过程行为。在这些方法中,CCA能够最大

化变量之间的相关性,从而识别具有内协方差强相关的故障^[6]。但是,由于基于CCA的过程监控是基于数据之间的相关性而非因果关系,因此CCA无法准确定位根原因。在发生丝状污泥膨胀的情况下,操作者难以在没有任何关于故障根原因信息的情况下做出决定。根原因诊断可通过执行工厂范围的振荡识别,可以在控制回路或操纵变量中识别出异常的振荡^[7]。全厂振荡的根原因诊断主要分为谱包络、邻接矩阵、因果分析、转移熵和贝叶斯网络^[8]。但是,所有根原因分析方法都对噪声较敏感,从而可能隐藏实际故障。尽管特征提取、CCA和Granger因果分析的结合可以对初期故障进行故障识别和根原因分析,但如何及时恢复污泥膨胀是困难且耗时的。因此,准确预测污泥膨胀的剩余使用寿命(RUL)是必不可少的一环,即如何确定剩余多少时间来纠正污泥膨胀。但是这在污泥膨胀管理中没有得到足够的重视^[9]。尽管故障检测、传播识别、因果定位和RUL预测已经获得了或多或少单独的研究,但少有文章专门阐述综合协调污泥膨胀管理。而且,每种方法都或多或少有一定的缺陷。通过整合的方式解决这些缺陷是行之有效的替代方法。这种整合方式可以充分利用所有可用信息克服个别方

① 基金项目:国家自然科学基金项目(61873096 61673181 61533002);广州市科技计划项目(201804010256)。

作者简介:刘坚(1983-),工程师,从事水处理的研究。

通讯作者:刘乙奇(1983-),副教授,从事工业过程建模、诊断、优化和控制的研究, aulyq@scut.edu.cn。

法的局限性,从而支持更好的决策。笔者从系统的角度提出了一种新颖的混合框架,用于整合故障检测、因果分析、故障传播、*RUL* 预测和维护决策。对此,首先构建 CCA 模型和 Kurtosis(*KU*) 相结合的新的平方预测误差 *SPE*,即 SPE_{KU} ,并用于故障特征提取和故障检测。CCA 能够通过区分质量相关和质量无关的变量,然后探索它们的相关性以进行故障识别。与直接使用 *SPE* 进行故障检测的传统方法不同,*KU* 可用于 *SPE* 的故障特征提取,从而可以正确地放大和捕获异常 *SPE* 信号。Sludge Volume Index(*SVI*) 通常被选为污泥膨胀演变的监测性能指标。然而,由于信息单一且缺乏其他相关信息,如 *DO*(溶解氧)、营养物等,通常只能松散地建立丝状污泥膨胀的故障识别模型。因此,笔者在利用 SPE_{KU} 故障预警后,接着提出了多变量格兰杰因果关系(MVGC)方法并将更多信息吸收到故障识别中。通过实现 MVGC 方法,所提出框架能够捕获变量之间的空间和时间依赖性以及因果关系,从而通过图形方式定位污泥膨胀故障中最可疑的根原因。但是,如果不正确或多余的特征被包含入 MVGC 方法,根原因分析的性能会有很大偏差。因此,在运行 MVGC 之前 *KU* 被用于处理每个变量,旨在放大信号变化以促进 MVGC 有效分析。随后,实现了针对故障质量变量基于 ARMA 模型的多步预测,并实现了相关的 *RUL* 预测。鉴于 *RUL* 的结果,维护策略分为预防性维护和纠正性维护。通过引入预防性和纠正性维护的分类,可以及时有效地控制丝状污泥膨胀。

1 污泥膨胀的预测维护

笔者提出了一体化的解决方案:

a. 故障检测。选择变量训练和测试 CCA 模型,然后通过 *KU* 处理来自 CCA 模型的 *SPE*,一旦 SPE_{KU} 识别出早期故障,该过程将进入故障定位。如果新样本到达,它将返回故障检测。

b. 故障定位。一旦早期故障发生预警,*KU* 会再次被用于特征细化。细化后的特征可用于检验变量是否发生异变,按照特征随时间出现的先后序列组织变量,然后馈送到 MVGC 进行根原因分析。

c. *RUL* 预测。通过 ARMA 模型执行多步预测,一旦预测的 *SVI* 超过控制极限,就可以导出 *RUL*。

d. 维护。根据推导出的 *RUL* 和 MVGC 给出的故障传播概况支持后续的维护决策。

1.1 基于 CCA 的故障识别

CCA 是通过最大化变量之间的相关性来关联两个矢量的方法。假设收集 N 个数据样本, X_1 和 X_2 分别表示输入变量和输出变量,其中 $X_1 \in R^{d_1 \times N}$ 和 $X_2 \in R^{d_2 \times N}$, d_1 、 d_2 表示 X_1 和 X_2 的变量个数, $d = d_1 + d_2$ 。 X_1 和 X_2 去均值后,协方差可计算如下:

$$\begin{bmatrix} \Sigma_{X_1 X_1} & \Sigma_{X_1 X_2} \\ \Sigma_{X_2 X_1} & \Sigma_{X_2 X_2} \end{bmatrix} \approx \frac{1}{N} \begin{bmatrix} X_1 X_1^T & X_1 X_2^T \\ X_2 X_1^T & X_2 X_2^T \end{bmatrix} \quad (1)$$

设矩阵 $\Sigma_T = \Sigma_{X_1 X_1}^{-1/2} \Sigma_{X_1 X_2} \Sigma_{X_2 X_2}^{-1/2}$, 通过对矩阵进行奇异值分解(SVD)如下:

$$\Sigma_T = U \Lambda V^T \quad (2)$$

其中, U 和 V 是 SVD 分解的正交矩阵,而 Λ 是 SVD 分解的对角矩阵。定义 $J = \Sigma_{X_1 X_1}^{-1/2} U(:, 1:l)$ 和 $L = \Sigma_{X_2 X_2}^{-1/2} V(:, 1:l)$, 其中 l 是所选组分的个数, $l \leq \min(d_1, d_2)$ 。

残差向量定义如下:

$$e = L^T X_2 - \Lambda_l J^T X_1 \quad (3)$$

平方预测误差可以表述为:

$$R = e^T e \quad (4)$$

为了正确识别异常行为,故障控制限定义如下:

$$R_{lim} = \xi \chi_{1-\alpha}^2(\tau) \quad (5)$$

其中 $\xi = \Sigma_R / 2\mu_R$, $\tau = 2\mu_R^2 / \Sigma_R$, Σ_R 和 μ_R 是 R 的均值和方差, α 是显著性水平。通常,CCA 能够充当过程监控模型。但是由于 CCA 对早期故障不敏感,因此通常会错过弱故障特征。

1.2 基于峰度分析的特征提取

峰度分析是捕获信号变化的典型方法之一。峰度分析适用于特征提取的原因在于能够测量信号分布中的拖尾,而这些是通过峰度参数的异变来体现的。因此,峰度分析可用于识别信号的异变。峰度值为 x_{iz} , 其中 $i = 1, 2, \dots, d$; $z = 1, 2, \dots, N_i, \dots, N$ 。 *KU* 的计算式为:

$$KU_i = \frac{\sum_{z=1}^{N_i} (x_{iz} - \bar{x}_i)^4}{(N_i - 1) \sigma_i^4} \quad (6)$$

其中, KU_i 表示第 i 个变量的 *KU* 值;分母中 σ_i^4 的存在能够放大不稳定变化并进一步强化提取故障特征; σ 是信号的标准偏差; \bar{x}_i 是 $[1, N_i]$ 区间中信号 x_i 的平均值:

$$\bar{x}_i = \frac{1}{N_i} \sum_{z=1}^{N_i} x_{iz} \quad (7)$$

$$\sigma_i = \sqrt{\frac{1}{N_i - 1} \sum_{z=1}^{N_i} (x_{iz} - \bar{x}_i)^2} \quad (8)$$

由于归一化,假设 KU 值遵循高斯分布。另外,令式(6)中的 $x_{iz} = SPE_i(t)$, SPE_{KU} 和 3σ 理论相结合可用于故障检测,可由下式计算:

$$SPE_{KU_i}(t) = \frac{KU_i(t) - \overline{KU_i}}{\sigma_{KU}} \leq 3\sigma_{KU} \quad (9)$$

其中 $\overline{KU_i}$ 和 σ_{KU} 表示 KU_i 的均值和标准方差。一旦 $SPE_{KU}(t)$ 超过控制极限 $3\sigma_{KU}$ 将发生故障状态预警。在本案例的研究中, KU 有两种用途:一种是用 SPE 的特征提取以增强故障检测,通常基于 CCA 的故障检测方法对初期故障不敏感,而峰度能够充分强化 SPE ; 另一种是提炼原始变量特征,然后提供给 MVGC,为 MVGC 提供初步序列信息。馈送给 MVGC 的特征定义如下:

$$x_{KU_i} = \frac{\sum_{z=1}^{N_i} (x_{iz} - \bar{x}_i)^4}{(N_i - 1) \sigma_i^4} \quad (10)$$

1.3 基于 MVGC 的根原因分析

由于基于 CCA 的过程监控是基于数据的相关性而非因果关系,因此 CCA 技术无法确切地定位根原因。而且,由于过程的闭环结构,污泥膨胀往往不限于局部,而是发生全局的传播。格兰杰因果关系检验则可通过确定一个时间序列是否有助于预测另一个时间序列来确定因果关系。此外,格兰杰因果关系检验是一种图形模型,可以以图形的方式来呈现因果关系。假设 x_1 和 x_2 是从 X_1 和 X_2 中选择两个变量。如果 x_1 通过一系列 t 检验或 F 检验定性 x_2 预测值统计意义上显著,则时间序列 x_1 被称为 x_2 的格兰杰根原因。假设时间滞后的两个时间序列 $x_1(t)$ 和 $x_2(t)$ 可构成如下矩阵:

$$I(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \quad (11)$$

为了定义从 $x_2(t)$ 到 $x_1(t)$ 的因果关系, $\mathcal{F}_{2 \rightarrow 1}$ 用于表示 $x_2(t)$ 的历史数据是否有助于预测 $x_1(t)$,且预测效果超过 $x_1(t)$ 的自回归预测精度。为了计算多变量因果关系, MVGC 中使用了向量自回归(VAR)模型。在 VAR 模型中:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \sum_{j=1}^p \begin{pmatrix} A_{11j} & A_{12j} \\ A_{21j} & A_{22j} \end{pmatrix} \begin{pmatrix} x_1(t-j) \\ x_2(t-j) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix} \quad (12)$$

其中 p 是滞后观测值的最大步数或模型阶数, A 表示模型的系数, ε_1 、 ε_2 是每个时间序列的残差(预测误差)。残差协方差可计算如下:

$$\Sigma = \text{cov} \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (13)$$

$x_1(t)$ 的回归计算定义如下:

$$x_1(t) = \sum_{j=1}^p A_{11j} x_1(t-j) + \sum_{j=1}^p A_{12j} x_2(t-j) + \varepsilon_1(t) \quad (14)$$

剔除 $x_2(t)$ 变量的影响,重新计算 $x_1(t)$:

$$x_1(t) = \sum_{j=1}^p A_{11j} x_1(t-j) + \varepsilon_1(t) \quad (15)$$

定义 $\text{cov}(\varepsilon_1) = \Sigma_{11}$ 、 $\text{cov}(\varepsilon_1 \hat{\varepsilon}_1) = \Sigma'_{11}$, 则 $x_2(t)$ 到 $x_1(t)$ 的因果关系被定义为对数似然比:

$$\mathcal{F}_{2 \rightarrow 1} = \ln \frac{|\Sigma'_{11}|}{|\Sigma_{11}|} \quad (16)$$

从式(14)、(15)可以看出, Σ_{11} 表示包含了 x_2 的协方差,而 Σ'_{11} 表示没有包含 x_2 的协方差,式(16)则表示包含 x_2 是否可减小 ε_1 的方差的情况,且可称 x_2 是导致 x_1 的根原因。更多细节可见 MVGC 工具箱^[10]。

1.4 RUL 预测

故障预测的主要目的是通过评估组件或子系统的当前退化状态,估计当前状态到故障控制限所需要的时间,即剩余使用寿命(RUL)。RUL 可以表示为:

$$RUL = \inf\{t: x(t) \geq C | x(0) < C\} = t_f - t_0 | t_f > t_0 \quad (17)$$

式中 C —— 故障控制限;

t_0 —— 当前时刻;

t_f —— 随机失效时间。

为了预测 RUL,首先要合理地把握 SVI 的演化,但 SVI 是难以测量变量。ARMA 模型可作为长期预测变量的虚拟工具。在长期预测期间,需要收集历史和当前测量数据并作为第 1 步预测的起始点。然后,以迭代方式执行基于 ARMA 模型的多步预测。令 k 为预测的步骤,当 $1 < k < n_c$ 时迭代多步预测如下:

$$\hat{x}(t_0 + k | t_0) = \sum_{i=1}^{n_a} a_i \hat{x}(t_0 + k - i | t_0) + \sum_{i=k}^{n_c} c_i \hat{v}(t_0 + k - i) \quad (18)$$

其中 $\hat{v}(t)$ 和回归系数 a 、 c 可以通过递归扩展最小二乘法(RELS)估计。更多细节可以参考文献[11] n_a 和 n_c 分别是 a 和 c 的阶次。随着时间的演变,污泥膨胀的恶化通常变得更加严重。对于不同的阶段,必须采取不同的行动。为确保有

足够的时间进行连续污泥膨胀控制,考虑预防性维护和纠正性维护两种维护策略。对于大多数情况,预防性维护只能在一定程度上缓解污泥膨胀。如果越过纠正性维护边界,必须采取纠正措施。笔者按经验将预防性维护和纠正性维护之间的界限定义为 72mg/L 的 SVI ,而故障控制限则定义为 200mg/L 。当 SVI 超过 72mg/L ,根据污泥膨胀的严重程度可通过具体控制参数来调节。而在纠正性维护期间,必须进一步添加凝结剂或 FeSO_4 以杀死丝状细菌^[12]。若一旦 SVI 超过 200mg/L ,就必须关闭污水处理厂进行进一步的维护,以避免污水排入河流。

2 结果和讨论

2.1 基于氧化沟工艺的污水处理过程

活性污泥工艺主要用于去除废水中的有机物质和营养物。在本案例中利用改进的活性污泥法(即氧化沟工艺)的污水处理厂(中国北京)数据进行了相应的验证。这种改进的活性污泥工艺能够通过较长的污泥停留时间(SRT)来改善有机物的降解。其主要参数如下:平均进水流量约为 $170\,000\text{m}^3/\text{d}$,平均水力停留时间(HRT)为 16.5h , SRT 是 $15\sim 22$ 天。在许多污水处理厂中应用氧化沟工艺的主要原因是可以通过较低的堰溢流率、较少的污泥产量和更好的节能效果来实现总体污水处理性能。然而,由于比典型的活性污泥处理过程需要更大的土地面积,同时污泥膨胀事

件时有发生,特别是有显著的温度降低的情况下,丝状污泥膨胀的发生更加明显,所采集的数据均以一天为间隔进行采样,采样数据点共 213 个。从收集到数据的第 70 天起,由于进水温度突然降低,丝状污泥膨胀开始发生,污泥膨胀的持续时间超过了半年。天气、操作等不确定性的存在增加了丝状污泥膨胀的复杂性。因此,必须提出一种支持丝状污泥膨胀的框架。

2.2 故障检测

为了尽早发现早期故障,做到后期故障预防,并保证足够的时间进行后续维护,笔者首先用 CCA 进行初期故障检测。CCA 的潜在变量 1 和 2 分别为 $X_1 = [SRT, T, MLSS, SNO, COD]$ 和 $X_2 = [TP, BOD_5, TN, SVI]$ 。 X_1 是由易于测量的变量组成的向量,而 X_2 由污水处理厂难以测量的变量组成。213 个采样数据点中的 50 个样本用于故障检测模型训练,其余样本用于测试。通过 SPE 方法能够检测出故障大约发生在第 120 天。为了更好地提出方法,同时实现故障检测功能, KU 方法还可用于增强 SPE 信号从而改进故障特征,即 SPE_{KU} 方法。图 1 为故障检测结果,由于结合了 KU 方法和 SPE 方法,因此所提方法 SPE_{KU} 能够在测试装置的第 23 天识别出早期故障,这与传统 SPE 方法相比提前了 97 天,这主要在于 KU 方法能够放大残留误差并使故障检测更加灵敏。从训练数据中学习 SPE_{KU} 的控制极限,其显著性水平为 0.05。

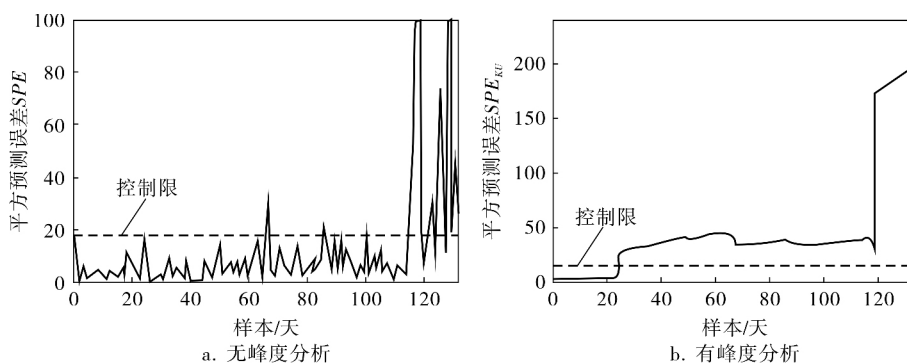


图 1 有无峰度分析情况下基于 SPE 的故障检测

2.3 故障识别

传统基于贡献图的故障识别方法通常由于拖尾效应,极易导致诊断错误。为解决这个问题,首先利用 KU 方法提取每个变量的特征。在变量集中, T 、 BOD_5 、 TP 、 COD 和 SNO 的变化是显而易见

的,但其他则不然。根据要发生的时间顺序依次重组所有变量。通过整理,变量 $[T, BOD_5, TP, COD, SNO]$ 表现出不同时间域的显著变化。由于彼此之间的时间延迟,可以合理地假设第 1 次发生异常事件是故障的来源。由于丝状污泥膨胀通常

由 SVI 表示,因此 SVI 作为目标变量并且位于表中的最终单元中。在通过 KU 特征提取进行可变

序列重组之后,执行 MVGC 用于因果关系分析(表1)。

表1 因果矩阵

因果关系 变量	变量	T	BOD_5	TP	COD	SNO	SRT	$MLSS$	TN	SVI
T		-	0.007 5	0.000 3	0.007 6	0.000 5	0.012 6	0.008 9	0.001 3	0.026 5
BOD_5		0.013 5	-	0.044 7	0.003 2	0.000 1	0.000 1	0.000 0	0.000 0	0.001 0
TP		0.000 8	0.017 2	-	0.000 1	0.000 1	0.000 1	0.008 5	0.002 3	0.011 1
COD		0.109 9	0.019 9	0.009 1	-	0.013 8	0.015 9	0.008 0	0.017 8	0.007 4
SNO		0.035 7	0.009 4	0.034 7	0.004 6	-	0.000 7	0.007 6	0.012 9	0.000 7
SRT		0.002 4	0.002 8	0.000 2	0.003 8	0.000 0	-	0.011 2	0.040 8	0.038 0
$MLSS$		0.023 6	0.029 1	0.044 4	0.000 0	0.007 1	0.033 1	-	0.019 0	0.013 1
TN		0.009 1	0.006 8	0.058 9	0.031 5	0.016 7	0.001 0	0.029 3	-	0.008 9
SVI		0.274 8	0.009 2	0.000 2	0.000 1	0.035 1	0.000 9	0.118 1	0.008 0	-

注: T ——温度; BOD_5 ——5 天生物需氧量; TP ——总氮; COD ——化学需氧量; SNO —— NO_2-N ; SRT ——污泥停留时间; $MLSS$ ——混合液悬浮固体浓度; TN ——输出总氮; SVI ——污泥沉降指数。

表1中列出的因果结果表明,加粗显示的单元格表明列变量对相应行变量的最大因果影响。例如,列中的变量 T 会影响行中的变量 COD 和 SVI 。在网络构建期间,由于 $\mathcal{F}_{SRT \rightarrow MLSS} > \mathcal{F}_{MLSS \rightarrow SRT}$ 较为显著,因此选择 $\mathcal{F}_{SRT \rightarrow MLSS}$ 。同时, $\mathcal{F}_{T \rightarrow SVI} > \mathcal{F}_{SVI \rightarrow T}$ 又较为显著,因此选择 $\mathcal{F}_{T \rightarrow SVI}$ 。图2表示了所开发的故障放大算法可以有效地以图形方式指示故障传播路径。

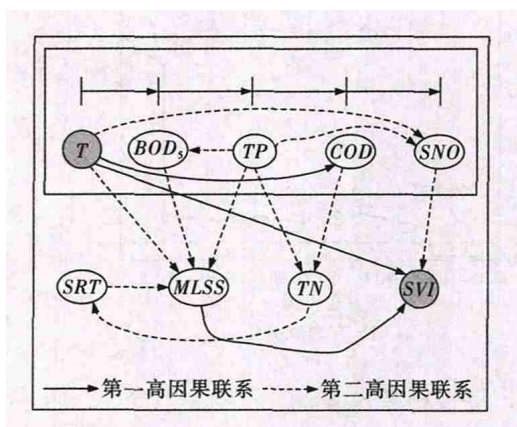


图2 基于峰度的故障传播路径

2.4 故障预测和 RUL 预测

为了进一步评估污泥膨胀状态,污泥膨胀剩余寿命 RUL 预测是不可或缺的一环。早期故障

预警之后,首先对 SVI 执行基于 ARMA 的多步预测。ARMA 模型的 AR 部分和 MA 部分的阶次分别选择 3 和 1。在本案例研究中,共得出了 36 步预测。图3中示意出了 9、18、27、36 步提前预测,表明所有预测结果达到了相应的精度。早期故障发生预警后,故障在第 32 天首次达到控制限值。也就是说,在本案例研究中, RUL 等于 32。

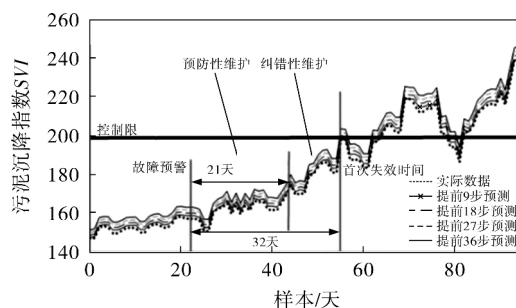


图3 基于 ARMA 的多步预测

3 结束语

提出了一种基于硬件测量和软测量混合预测维护框架,框架涵盖了故障检测、故障定位、故障识别、寿命预测和维护,从而有效地用于污泥膨胀控制和管理,并在氧化沟工艺的污水处理过程中得到了有效验证。在所提出方法中,具有峰度分

析和CCA的结合能够尽可能早地捕获污泥膨胀特征。随后,MVGC可有效地诊断故障的根本原因且显示系统故障的传播路径。最后,通过基于ARMA模型软测量推导出36步的提前预测,可相应地提出适当的维护策略。该方案可为污泥膨胀管理和维护提供指导,也可以扩展到轻工、化工及材料腐蚀等众多过程。

参 考 文 献

- [1] Ng W J, Ong S L, Hossain F. An Algorithmic Approach for System-specific Modelling of Activated Sludge Bulking in an SBR [J]. *Environmental Modelling & Software*, 2000, 15(2): 199 ~ 210.
- [2] Guo J, Zhang L, Chen W, et al. The Regulation and Control Strategies of a Sequencing Batch Reactor for Simultaneous Nitrification and Denitrification at Different Temperatures [J]. *Bioresource Technology*, 2013, 133: 59 ~ 67.
- [3] Allison A E F, Dickson M E, Fisher K T, et al. Dilemmas of Modelling and Decision-making in Environmental Research [J]. *Environmental Modelling & Software*, 2018, 99: 147 ~ 155.
- [4] Montserrat A, Bosch L, Kiser M A, et al. Using Data from Monitoring Combined Sewer Overflows to Assess, Improve, and Maintain Combined Sewer Systems [J]. *Science of the Total Environment*, 2015, 505: 1053 ~ 1061.
- [5] Lee J M, Qin S J, Lee I B. Fault Detection and Diagnosis Based on Modified Independent Component Analysis [J]. *AIChE Journal*, 2006, 52(10): 3501 ~ 3514.
- [6] Chen Z, Ding S X, Zhang K, et al. Canonical Correlation Analysis-based Fault Detection Methods with Application to Alumina Evaporation Process [J]. *Control Engineering Practice*, 2016, 46: 51 ~ 58.
- [7] Zhu J, Ge Z, Song Z. Distributed Parallel PCA for Modeling and Monitoring of Large-Scale Plant-Wide Processes with Big Data [J]. *IEEE Transactions on Industrial Informatics*, 2017, 13(4): 1877 ~ 1885.
- [8] Wang J, Yang F, Chen T, et al. An Overview of Industrial Alarm Systems: Main Causes for Alarm Overloading, Research Status, and Open Problems [J]. *IEEE Transactions on Automation Science and Engineering*, 2016, 13(2): 1045 ~ 1061.
- [9] 周东华, 魏慕恒, 司小胜. 工业过程异常检测、寿命预测与维修决策的研究进展 [J]. *自动化学报*, 2013, 39(6): 711 ~ 722.
- [10] Barnett L, Seth A K. The MVGC Multivariate Granger Causality Toolbox: A New Approach to Granger-causal Inference [J]. *Journal of Neuroscience Methods*, 2014, 223: 50 ~ 68.
- [11] Chen J, Ganigue R, Liu Y, et al. Real-Time Multi-step Prediction of Sewer Flow for Online Chemical Dosing Control [J]. *Journal of Environmental Engineering*, 2014, 140(11): 401 ~ 437.
- [12] 游佳, 高晨晨, 陈轶, 等. 微丝菌引发的污泥膨胀对污水处理厂效能的影响分析 [J]. *给水排水*, 2018, 54(8): 57 ~ 60.

(收稿日期: 2019-01-24, 修回日期: 2019-04-24)

Prediction and Maintenance of Filamentous Sludge Bulking in Activated Sludge Process

LIU Jian¹, YUAN Ling-ling², HUANG Dao-ping², LIU Yi-qi², YU Guang-ping¹

(1. *Shenyang Institute of Automation, Guangzhou Chinese Academy of Science;*

2. *School of Automation Science and Engineering, South China University of Technology*)

Abstract The filamentous sludge bulking influences the activated sludge process' stable operation often. For purpose of effectively maintaining the filamentous sludge bulking, making use of the kurtosis and canonical correlation analysis to extract the easy-to-measure features from sensor data was implemented, including having ARMA model-based soft sensor adopted to extract difficult-to-measure features. Having this interactive information integrated into fault diagnosis, propagation identification, causality analysis, residual service life (*RUL*) prediction and maintenance can achieve an integrated solution for filamentous sludge bulking maintenance of wastewater treatment. The results show that, the proposed method can detect, locate and track filamentous sludge bulking, even predict the life of sludge bulking to suggest maintenance strategies.

Key words fault diagnosis, activated sludge, filamentous sludge bulking, soft measurement