

DOI:10.19853/j.zgjsps.1000-4602.2022.09.009

# 基于LSSVM交互预测的供水管网爆管检测

吴以文<sup>1</sup>, 杜坤<sup>1</sup>, 吴汉清<sup>2</sup>, 罗雄武<sup>3</sup>, 宋志刚<sup>1</sup>, 黄乐烽<sup>1</sup>

(1. 昆明理工大学 建筑工程学院, 云南 昆明 650500; 2. 云南省城乡规划设计研究院, 云南 昆明 650228; 3. 昆明官房建筑设计有限公司, 云南 昆明 650051)

**摘要:** 随着我国越来越多的供水管网安装SCADA系统,基于数据预测的爆管检测方法越来越受到重视。传统基于数据预测的方法大多根据单点流量、压力的历史监测数据预测当前时刻监测值,当预测值与监测值的差值超过阈值时判定为爆管。然而,实践经验表明,监测数据的丢失与错误会严重影响单点预测结果,进而引起频繁误报与漏报。考虑到实际用水及监测数据的空间关联性(例如水压监测点布置距离越近其监测数据相关性越大),开展了基于最小二乘支持向量机(LSSVM)交互预测的供水管网爆管检测研究。对管网中不同位置监测数据构建多输入单输出LSSVM交互预测模型,选取1倍标准差为阈值进行爆管检测,并与传统的卡尔曼滤波爆管检测结果相对比。结果表明,LSSVM交互预测模型能降低数据丢失、数据错误对预测结果的影响,且对较小的爆管响应更加灵敏,进而有效地提高了基于数据预测的爆管检测性能。

**关键词:** 供水管网; 爆管; 最小二乘支持向量机(LSSVM); 交互预测

**中图分类号:** TU991 **文献标识码:** A **文章编号:** 1000-4602(2022)09-0058-06

## Water Supply Network Burst Detection Based on Least Squares Support Vector Machine Interactive Prediction

WU Yi-wen<sup>1</sup>, DU Kun<sup>1</sup>, WU Han-qing<sup>2</sup>, LUO Xiong-wu<sup>3</sup>, SONG Zhi-gang<sup>1</sup>, HUANG Le-feng<sup>1</sup>

(1. College of Civil Engineering and Architecture, Kunming University of Science and Technology, Kunming 650500, China; 2. Urban Planning & Design Institute of Yunnan Province, Kunming 650228, China; 3. Kunming Guanfang Architecture & Design Inc., Kunming 650051, China)

**Abstract:** With an increasing number of SCADA system installed in the water supply network, the method of pipe burst detection based on data prediction has been paid more and more attention. Most of the traditional methods based on data prediction predict the current monitoring value according to the historical monitoring data of single point flow and pressure. When the difference between the predicted value and the monitoring value exceeds the threshold, the pipe burst was identified. However, practical experience shows that the loss and error of monitoring data will seriously affect the single point prediction results, and then cause frequent false positives and missing reports. Considering the spatial correlation between actual water consumption and monitoring data (for example, the closer the distance of water pressure monitoring points is, the greater the correlation of monitoring data is), water supply network pipe

基金项目: 云南省重点领域科技计划项目(202003AC100001)

通信作者: 杜坤 E-mail: 765818261@qq.com

burst was detected based on least squares support vector machine (LSSVM) interactive prediction. A multi input and single output LSSVM interactive prediction model was constructed based on the monitoring data of different positions in the pipe network. A double standard deviation was selected as the threshold value for pipe burst detection, and the detection results were compared with those of traditional Kalman filtering. LSSVM interactive prediction model could reduce the influence of data loss and data error on the prediction results, and was more sensitive to small pipe burst, thus effectively improving the performance of pipe burst detection based on data prediction.

**Key words:** water supply network; pipe burst; LSSVM; interactive prediction

近年来,随着越来越多的供水管网安装SCADA系统,基于流量与压力监测数据的爆管检测方法受到了广泛关注。相较于传统检测方法,例如探地雷达法、热红外成像法、人工听音法,基于数据驱动的方法利用实时监测数据对爆管反应更加迅速,有利于缩短爆管修复时间。从数学原理上来说,基于数据驱动的方法可大致分为基于分类与基于预测两类。例如,Mounce等<sup>[1]</sup>利用人工神经网络(ANN)对管网水压监测信号分类实现了爆管检测。随后,Mounce等<sup>[2]</sup>将监测数据转化为向量格式并构建支持向量回归(SVR)模型识别异常数据。基于分类的爆管检测通常需要大量训练数据并耗费较长训练时间,这不利于爆管实时检测。针对该问题,Ye等<sup>[3]</sup>提出了基于卡尔曼滤波预测模型的爆管检测算法,其根据历史监测数据估计当前时刻监测值,当监测值与预测值的差值超出阈值时判断爆管发生。在Ye等的研究基础上,顾建强等<sup>[4]</sup>提出了基于自适应卡尔曼滤波的爆管检测算法。

虽然上述基于数据预测的爆管检测方法成功用于一些供水管网爆管识别,但都只考虑了单点监测数据,缺乏将多个监测数据结合起来的研究。而实践表明,监测数据的丢失与错误会严重影响单点预测效果,进而引起频繁误报与漏报。鉴于此,本研究提出了基于最小二乘支持向量机(LSSVM)交互预测的供水管网爆管检测方法,其根据管网中不同位置监测数据构建多输入单输出LSSVM交互预测模型,旨在降低数据丢失与错误对预测结果的影响,进而提高基于数据预测的爆管检测性能。

## 1 爆管预测模型

供水管网爆管会导致压力或流量监测信号异常波动,因此可根据历史压力或流量监测时间序列构建预测模型检测爆管。分别介绍了基于卡尔曼

滤波的预测模型、基于LSSVM的预测模型的基本原理,以及基于预测残差(预测值与测量值之差)的爆管检测方法。

### 1.1 基于卡尔曼滤波的预测模型

卡尔曼滤波的基本原理:在线性状态空间表示的基础上,对有噪声观测信号进行处理,以预测真实监测信号。Ye等<sup>[3]</sup>最早将卡尔曼滤波用于供水管网爆管检测,其将监测信号按用水周期分解得到不同时刻近似平稳的监测值时间序列,并以此为基础实现当前状态监测值的最优无偏估计,具体计算步骤如下:

#### ① 估计当前状态值及其方差

$$X_{i/(i-1)} = X_{(i-1)/(i-1)} \quad (1)$$

$$P_{i/(i-1)} = P_{(i-1)/(i-1)} + Q_i \quad (2)$$

式中: $X_{i/(i-1)}$ 为*i*时刻预测值; $X_{(i-1)/(i-1)}$ 为*i-1*时刻状态估计值; $P_{i/(i-1)}$ 为*i*时刻预测值的方差; $P_{(i-1)/(i-1)}$ 为*i-1*时刻状态变量方差; $Q_i$ 为状态过程噪声方差。

#### ② 计算*i*时刻的卡尔曼增益方程

$$G_i = \frac{P_{i/(i-1)}}{P_{i/(i-1)} + r_i} \quad (3)$$

式中: $r_i$ 为*i*时刻预测误差。

#### ③ 利用加权平均法修正状态估计值

$$X_{i/i} = X_{i/(i-1)} + G_i S_i \quad (4)$$

式中: $S_i$ 为新息序列(新更新的数据序列), $S_i = Z_i - X_{i/(i-1)}$ ,其中 $Z_i$ 为*i*时刻测量值。

#### ④ 更新*i*时刻状态变量方差

$$P_{i/i} = (1 - G_i) P_{i/(i-1)} \quad (5)$$

对*i+1*时刻状态进行滤波,其中 $Q_i = G_i^2 C_i$ , $R_i = C_i + P_{i/(i-1)}$ 。

信息标准差序列计算如下:

$$C_i = \frac{1}{M} \sum_{n=i-M+1}^i S_n^2 \quad (6)$$

式中:  $M$  为时域滚动窗口, 本研究中取  $M=5$ 。

## 1.2 基于LSSVM的预测模型

LSSVM是基于最小二乘函数和等式约束改进的SVM算法, 对小样本、非线性、高维数数据具有较强的预测能力。相较于卡尔曼滤波预测模型, LSSVM预测模型不依赖于数据平稳假设, 且能构建多输入单输出的预测模型, 能有效挖掘数据间的相关性。对给定训练样本集  $S = \{(x_i, y_i), i = 1, 2, \dots, n; x \in R^d, y \in R\}$ , 利用非线性映射函数  $\varphi(x)$  将样本从原始空间映射到高维特征空间, 则基于LSSVM的预测模型可以表示为:

$$y(x) = \omega^T \varphi(x) + b \quad (7)$$

式中:  $\omega$  为权重;  $b$  为偏差。

根据结构风险最小化原则, 寻找  $\omega$ 、 $b$  值最小化, 即有:

$$\min J(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i^2 \quad (8)$$

$$\text{s.t. } y_i = \omega^T \varphi(x_i) + b + \xi_i \quad (9)$$

式中:  $J(\omega, \xi)$  表示结构风险;  $\xi_i$  为模型对训练样本的预测误差;  $C$  为控制惩罚程度的正则化参数。采用拉格朗日求解得:

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i^2 - \sum_{i=1}^n \alpha_i [\omega^T \varphi(x_i) + b + \xi_i - y_i] \quad (10)$$

式中:  $\alpha$  为拉格朗日乘子,  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]$ 。

采用拉格朗日函数将上述优化问题转化为求解线性方程组, 并根据 Mercer 条件存在核函数  $K(x_i, x_j)$ , 使得:

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \quad (11)$$

径向基核函数具有较宽收敛域和较强泛化能力, 是较为理想的核函数, 所以本研究选择径向基核函数作为LSSVM核函数, 表达式为:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (12)$$

式中:  $\sigma$  为核函数参数, 表示径向基函数的宽度。

最终的LSSVM预测函数为:

$$y(x) = \sum_{i=1}^n \alpha_i K(x_i, x_j) + b \quad (13)$$

## 1.3 基于LSSVM的多输入单输出交互预测模型

结合置信区间与LSSVM预测判定数据错误, 其中置信区间法常被用于错误数据判断<sup>[5]</sup>。具体步骤为: ①基于历史压力监测数据, 利用LSSVM模型预测各压力监测点的下一时刻数据; ②构建置信水平为95%的置信区间, 同时检查各点监测值是否超出置信区间, 若仅有一个监测点数据超出置信区间, 则判定该监测点数据错误。上述判定的依据是实际中两个或两个以上监测点同时出现监测数据错误的可能性较小。置信水平为95%的置信区间为  $(y_i \pm 1.96 \sqrt{\text{Var}(R)})$ 。其中:  $y_i$  为  $i$  时刻预测值;  $R$  为预测数据和测量数据间的残差;  $\text{Var}(R)$  为  $R$  的方差。

由于供水管网大多呈环状布置且同一类型用户具有相似的用水习惯, 这使得管网中不同位置监测数据存在相关性, DMA分区管网中监测数据间的相关性尤为显著。图1给出了某DMA管网一周的流量与压力监测数据(1 m水柱产生的压力约为9.8 kPa)。可知, 用水量存在显著的周期变化, 而两个水压监测点数据则存在显著的同步变化规律, Wu等<sup>[6]</sup>将这一特征定义为不同位置监测数据之间随时间变化的相关性。采用Pearson相关系数<sup>[7]</sup>衡量数据间的相关程度, 其计算公式为:

$$\text{cor}(x_1, x_2) = \frac{\text{cov}(x_1, x_2)}{\sqrt{\text{Var}(x_1) \text{Var}(x_2)}} \quad (14)$$

式中:  $\text{cov}(x_1, x_2)$  为  $x_1, x_2$  的协方差。

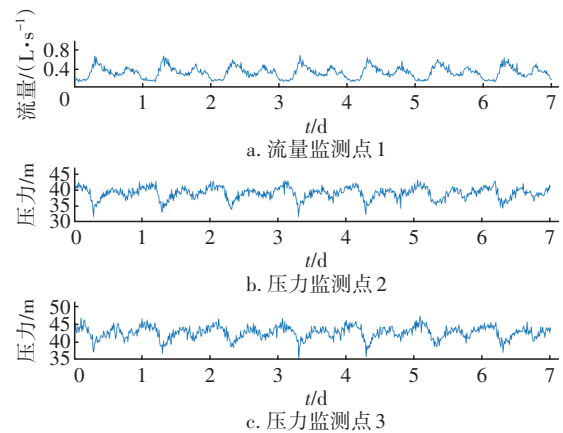


图1 监测数据

Fig.1 Monitoring data

当早晚用水高峰时, 流量监测值最大、水压监测值最小, 反之亦然, 因此可构建基于LSSVM的多

输入单输出交互预测模型。考虑到水压监测信号的同步变化规律,且水压监测点安装费用低,在实际工程中被更广泛应用,构建两输入一输出的水压监测值交互预测模型,如下所示:

$$\begin{cases} y(x_1) = \omega^T \varphi(x_2, x_3) + b \\ y(x_2) = \omega^T \varphi(x_1, x_3) + b \\ y(x_3) = \omega^T \varphi(x_1, x_2) + b \end{cases} \quad (15)$$

在实际应用中,当管网中布置水压监测点个数超过3个时,可预先通过空间关联性对不同位置监测点分组,再对具有强相关性的监测数据构建两输入一输出预测模型。这样当一个监测点数据出现丢失或错误时,采用其他监测点数据预测该监测点数据,形成管网中不同监测点数据间的交互预测,进而避免单点监测数据误差对预测结果的影响。

#### 1.4 基于预测残差的爆管检测

参考文献[8],采用基于预测残差的方法检测爆管,其具体步骤为:①基于历史压力监测数据,利用LSSVM模型交互预测压力监测点的下一时刻数据;②基于正常时段压力预测残差计算爆管检测阈值;③检查预测残差是否超出阈值,若小于阈值,则判定该时刻爆管。阈值 $T$ 计算公式如下:

$$T = \bar{R} + \left[ \frac{1}{N} \sum_{i=1}^N (R_i - \bar{R})^2 \right]^{1/2} \quad (16)$$

式中: $R_i$ 为预测值与测量值间的残差; $N$ 为测试集样本数,本研究中 $N = 14 \times 96 = 1344$ ;  $\bar{R}$ 为正常时段残差的均值,按式(17)计算。

$$\bar{R} = \frac{1}{N} \sum_{i=1}^N R_i \quad (17)$$

爆管检测成功率与误报率存在权衡关系,较小的爆管检测阈值能提高爆管检出率,但会导致较高的误报率,反之亦然。初次检测建议采用正常时段预测残差与1倍标准差之和作为阈值,实际中可根据具体情况调整阈值,当需要降低误报率时,可适当增大爆管检测阈值。

## 2 爆管检测效果分析

选取第二届水系统模拟与控制联合国际会议(CCWI/WDSA 2020)所提供的L-Town-C区管网2018年的SCADA监测数据,该数据可在www.ccwi-wdsa2020.com获取。如图2所示,供水管网有92个节点和3个压力监测点,每个监测点都以15 min为间隔,记录了140 d的瞬时流量、压力等数据。

理论上,应收集包括爆管发生时间、持续时长及爆管压力等实测数据来验证本研究所提出方法的可行性,但实际中爆管监测数据往往未被记录。因此,参考文献[9],在EPANET中构建该供水管网水力模型,以原始SCADA监测数据为基础采用仿真对比卡尔曼滤波与LSSVM的爆管检测性能。

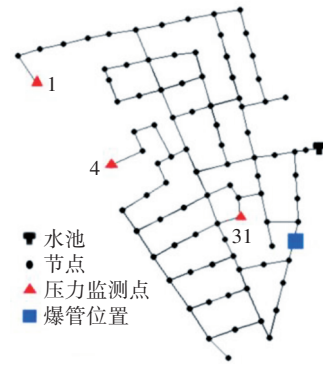


图2 供水管网拓扑结构

Fig.2 Topological structure of water supply network

#### 2.1 监测数据无异常时爆管检测效果对比

图3给出了基于卡尔曼滤波与LSSVM的爆管检测结果。

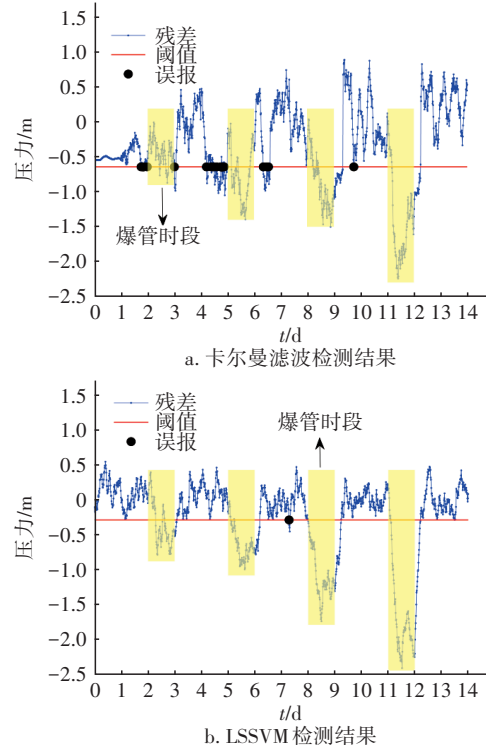


图3 数据无异常情况下基于卡尔曼滤波与LSSVM的爆管检测结果

Fig.3 Burst detection results based on Kalman filtering and LSSVM without abnormal data



为对比不同程度爆管检测性能,分别在第3、6、9、12天添加用水总量5%、10%、15%、20%的爆管。由图3可知,卡尔曼滤波与LSSVM法均能检测到4个爆管信号,但LSSVM法对较小爆管响应更灵敏且非爆管时段误报率更低,故LSSVM法整体检测效果更好。具体地,卡尔曼滤波与LSSVM爆管检测阈值分别为 $-0.53\text{ m}$ 和 $-0.31\text{ m}$ (如图3红线所示),对于LSSVM算法,5%的爆管导致压降超出阈值 $0.48\text{ m}$ ,而对于卡尔曼滤波算法,相同爆管导致的压降仅超出阈值 $0.16\text{ m}$ ,这表明LSSVM能更好地检测到较小爆管。

统计了数据无异常时卡尔曼滤波与LSSVM的爆管检出率(被检测出的爆管信号数与爆管信号总数的比值)和误报率(被检测出正常信号与正常信号总数的比值),结果如下:卡尔曼滤波与LSSVM的爆管检出率分别为59.11%和87.24%,误报率分别为16.67%和8.96%。与卡尔曼滤波相比,LSSVM爆管检出率提升了28.13%,误报率下降了7.71%,爆管检测性能显著提升。

## 2.2 监测数据异常时爆管检测效果对比

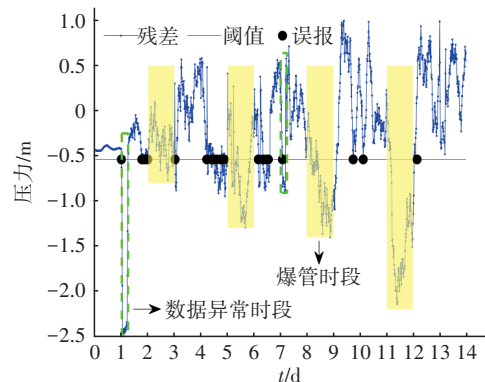
为对比分析数据异常对爆管检测结果的影响,探讨了数据丢失、数据错误两种异常情况下基于卡尔曼滤波与LSSVM的爆管检测结果,如图4所示。设置2次持续6 h的数据异常,其中在第2天设置数据丢失、第8天添加-5%的检测误差。在图4中用绿色虚线表示数据异常时段。

由图4可知,数据异常对卡尔曼滤波的爆管检测效果影响显著,而对LSSVM的检测效果影响较小。具体地,当数据丢失时,卡尔曼滤波出现了长达6 h的持续误报信号,而LSSVM仅出现5 min短时间误报;当数据错误时,卡尔曼滤波误报持续时间为4 h,而LSSVM未发生误报。

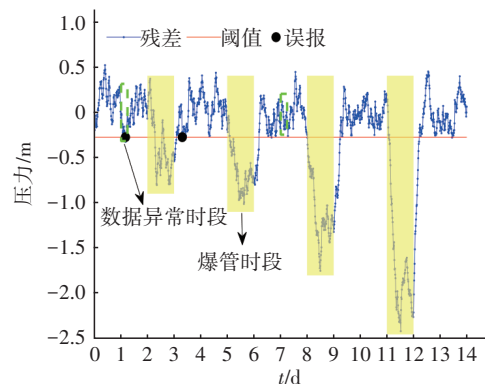
统计了数据异常情况下卡尔曼滤波与LSSVM的爆管检出率和误报率,结果如下:基于卡尔曼滤波与LSSVM的爆管检出率分别为56.20%和87.15%,误报率分别为22.92%和9.06%。对比数据正常时检测结果,数据出现异常时,卡尔曼滤波与LSSVM爆管检出率分别下降了2.91%和0.09%,误报率分别上升了6.25%和0.1%。综上所述,若数据出现异常,则对卡尔曼滤波的爆管检测效果影响更加显著。

为进一步解释数据异常对LSSVM爆管检测效

果影响较小的原因,图5给出了基于LSSVM异常数据的修正结果。可知,第2天和第8天分别出现了数据丢失和数据错误,修正后的数据符合正常压力变化的基本趋势。可见,该方法修正效果较好,并进一步降低了数据异常对爆管检测效果的影响。



a. 卡尔曼滤波检测结果



b. LSSVM检测结果

图4 数据异常情况下基于卡尔曼滤波与LSSVM的爆管检测结果

Fig.4 Burst detection results based on Kalman filtering and LSSVM with abnormal data

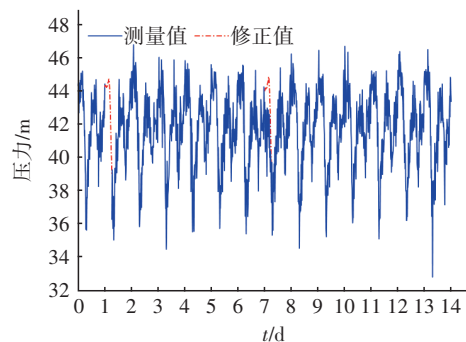


图5 基于LSSVM异常数据的修正结果

Fig.5 Correction results of abnormal data based on LSSVM

综上,数据异常对卡尔曼滤波爆管检测效果的影响显著,出现持续误报且误报率较高;对LSSVM

爆管检测效果的影响较小,误报率远低于卡尔曼滤波。这是由于爆管检测效果受预测精度影响,精度越高则检测效果越好。卡尔曼滤波通过对当前状态测量值与前一状态估计值进行加权平均估计当前状态值,对非线性样本预测精度较低,尤其存在异常数据样本时,预测精度会显著下降;而LSSVM通过径向基核函数将低维空间的非线性不可分问题转化为高维空间的线性可分问题,对非线性样本的预测精度高,引入松弛变量和惩罚因子剔除异常数据,可进一步提高预测精度。

### 3 结论与建议

① 监测数据无异常时,LSSVM算法整体检测效果优于卡尔曼滤波法,爆管检出率由59.11%提升至87.24%,误报率由16.67%降至8.96%。此外,LSSVM法对较小的爆管响应更加灵敏,5%的爆管导致压降超出阈值0.48 m,而卡尔曼滤波法仅超出阈值0.16 m。

② 当数据出现异常时,卡尔曼滤波爆管检测效果显著下降,出现了共计10 h的持续误报信号且误报率高达22.92%;而对LSSVM爆管检测效果的影响较小,仅出现5 min短时间误报,且误报率为9.06%,远低于卡尔曼滤波。

③ LSSVM交互预测模型有效地降低了数据丢失、数据错误对预测结果的影响,且对较小的爆管响应更加灵敏,进而显著地提升了基于数据预测的爆管检测性能。值得说明的是,针对本研究中的DMA管网,笔者建议将Pearson相关系数大于0.8作为交互预测启动条件。但对实际非DMA管网,由于管网拓扑结构差异、监测点数量与布局不同,交互预测的启动条件还需要进一步验证。

### 参考文献:

- [1] MOUNCE S R, MACHELL J. Burst detection using hydraulic data from water distribution systems with artificial neural networks [J]. Urban Water Journal, 2006, 3(1):21-31.
- [2] MOUNCE S R, MOUNCE R B, BOXALL J B. Novelty detection for time series data analysis in water

distribution systems using support vector machines [J]. Journal of Hydroinformatics, 2010, 13(4):672-686.

- [3] YE G L, FENNER R. Kalman filtering of hydraulic measurements for burst detection in water distribution systems [J]. Journal of Pipeline Systems Engineering and Practice, 2011, 2(1):14-22.
- [4] 顾建强,刘海星,张朝,等. 基于自适应卡尔曼滤波的用水量预测和爆管诊断[J]. 中国给水排水, 2019, 35(7):62-67.  
GU Jianqiang, LIU Haixing, ZHANG Chao, et al. Flow prediction and burst diagnosis based on adaptive Kalman filter [J]. China Water & Wastewater, 2019, 35(7):62-67 (in Chinese).
- [5] 杨浩. 基于供水管网离线流量数据处理定位漏失研究[D]. 哈尔滨:哈尔滨工业大学, 2016.  
YANG Hao. Study on Leakage Localization of Water Distribution Network Processing Off-line Flow Data [D]. Harbin: Harbin Institute of Technology, 2016 (in Chinese).
- [6] WU Y P, LIU S M, SMITH K, et al. Using correlation between data from multiple monitoring sensors to detect bursts in water distribution systems [J]. Journal of Water Resources Planning and Management, 2017, 144 (2): 04017084.
- [7] DE A, TFM B, GJSC B. On relationships between the Pearson and the distance correlation coefficients [J]. Statistics & Probability Letters, 2020, 169:108960.
- [8] YE G L, FENNER R. Study of burst alarming and data sampling frequency in water distribution networks [J]. Journal of Water Resources Planning and Management, 2013. DOI: 10.1061/(ASCE)WR.1943-5452.0000394.
- [9] YUAN X H, CHEN C, YUAN Y B, et al. Short-term wind power prediction based on LSSVM-GSA model [J]. Energy Conversion and Management, 2015, 101: 393-401.

作者简介:吴以文(1995-),男,江苏徐州人,硕士研究生,主要研究方向为供水管网爆管检测。

E-mail: 3435636264@qq.com

收稿日期:2021-01-15

修回日期:2021-04-26

(编辑:任莹莹)