

# 基于传感物联网数据融合和线性支持向量算法的地表径流水质检测

卫荣浩<sup>1</sup>, 陈航<sup>1</sup>, 刘昊<sup>2</sup>, 徐飞<sup>3†</sup>

(1. 贵州大学继续教育学院, 贵州 贵阳 550025; 2. 都匀生态环境监测中心, 贵州 都匀 558000;

3. 贵州大学管理学院, 贵州 贵阳 550025)

**摘要:** 常规的地表径流水质检测方法主要结合遥感技术获取水体颜色, 并通过图像处理技术分析水体质量, 由于对水体等级的细化程度较低, 导致检测精度较差。为此, 提出了基于传感物联网数据融合和线性支持向量算法的地表径流水质检测。首先采用漂浮式传感器获取地表径流水体数据, 并通过对数据进行融合处理, 整理出水体数据与水质等级之间的映射关系。然后结合线性向量支持算法, 构建回归函数实现地表径流水质预测。最后结合预测残差阈值以及水体等级划分标准, 实现对水质等级的判断。在实验中, 对提出的方法进行了检测精度的检验。最终的测试结果表明, 采用提出的方法对地表径流水质进行检测时, 算法的RMSE值较低, 具备较为理想的检测精度。

**关键词:** 数据融合; 支持向量算法; 地表径流; 水质检测; 检测精度

中图分类号: TP181

文献标识码: A

## Water Quality Detection of Surface Runoff Based on Sensor-Internet-of-Things Data Fusion and Linear Support Vector Algorithm

WEI Ronghao<sup>1</sup>, CHEN Hang<sup>1</sup>, LIU Hao<sup>2</sup>, XU Fei<sup>3†</sup>

(1. School of Continuing Education, Guizhou University, Guiyang, Guizhou 550025, China;

2. Ecological Environment Monitoring Center of Duyun, Duyun, Guizhou 558000, China;

3. School of Management, Guizhou University, Guiyang, Guizhou 550025, China)

**Abstract:** The current conventional surface runoff water quality detection method mainly combines remote sensing technology to obtain the color of the water body and image processing technology to achieve the analysis of water quality, which results in poor detection accuracy due to the low degree of refinement of the water body grade. In this regard, surface runoff water quality detection based on sensing IoT data fusion and linear support vector algorithm is proposed. Firstly, floating sensors are used to acquire the surface runoff water body data, and the mapping relationship between the water quality data and the water quality grade is organized by fusion processing of the data. Then the regression function is constructed by combining the linear vector support algorithm to realize the prediction of surface runoff water quality. Finally, combining the prediction residual threshold and the water body grade classification standard, the judgment of water quality grade is realized. In the experiment, the proposed method was tested for detection accuracy. The final test results show that when the proposed method is used to detect surface runoff water quality, the algorithm has a lower RMSE value and has a more ideal detection accuracy.

**Key words:** data fusion; support vector algorithm; surface runoff; water quality detection; detection accuracy

收稿日期: 2023-08-02

基金项目: 贵州省2023年科技支撑一般项目(黔科合支撑[2023]一般108)

作者简介: 卫荣浩(1977-), 男, 贵州榕江人, 硕士, 讲师, 研究方向: 物联网技术, 大数据与人工智能技术应用, 信息系统项目管理。

† 通信联系人, E-mail: fxu@gzu.edu.cn

当前针对地表径流水质检测技术主要有两种研究方向,分别为基于图像处理技术的水质检测以及基于模型预测技术的水质检测。其中,基于图像处理技术的水质检测指的是首先通过遥感技术对水质图像进行获取,提取水体图像中的灰度以及颜色特征,并与现有的水质图像分类标准进行对比,从而实现水体质量的检测。而基于模型预测技术的水质检测方法主要是通过结合不同的算法,构建出水质预测模型,并将模型预测结果与实际的水质测量结果进行对比,通过设定阈值对水质情况进行分析。例如,文献[1]通过采用单片机作为水质系统的主要硬件,结合图像处理技术对水体质量进行了有效检测。文献[2]通过对水体图像进行分割,并提取不同水体图像的 RGB 特征值,从而获得水质状况。文献[3]以 ARM 内核为硬件芯片,结合遥感技术设计了一种水质检测装置。文献[4]以远程无人海洋水质作为检测对象,通过结合 STM32 高性能芯片,对水质监测仪进行了硬件优化。文献[5]基于 STM32 的顺序注射水质在线监测系统,通过控制核心和温度控制算法实现了流路控制和实验温度条件的准确控制,并通过光强信号处理和计算得出测量结果。文献[6]汇编了可用的拟除虫菊酯使用数据和地表水采样数据,然后将 GIS 方法应用于动态模拟建模和使用限制缓冲分析。文献[7]基于连续投影算法(SPA)和粒子群优化-最小二乘支持向量机(PSO-LSSVM)建立了多元校正模型来提高可见-近红外光谱法检测水质 pH 值的精度和稳定性。对此,本文采用模型预测技术,通过对传感器采集到的水体数据进行融合处理,并结合模型预测结果,实现地表径流水体质量有效检测<sup>[8]</sup>。

## 1 基于传感器联网数据融合和线性支持向量算法的地表径流水质检测

### 1.1 基于传感器联网数据融合的地表径流传感数据处理

为对地表径流水质进行有效测量,本文通过在测点布设多个漂浮式传感器,通过对传感器采集到的水质数据进行计算,得到不同测试区域的区间概率分配,并结合水质异常阈值,实现地表径流水质检测<sup>[9]</sup>。为保证数据融合效果,本文选用的漂浮式传感器均配置了独立的操作系统,各监测节点相互

独立,同时安装通信芯片,支持无线网络通信功能,从而实现传感信息的有效传输。漂浮式传感器通过天线接收控制系统的操作指令,并采用锚和重物保持水体内部空间的受力平衡,使传感器能够漂浮在水体中,实现对水体的持续监测<sup>[10]</sup>。漂浮式传感器的工作原理如图 1 所示。

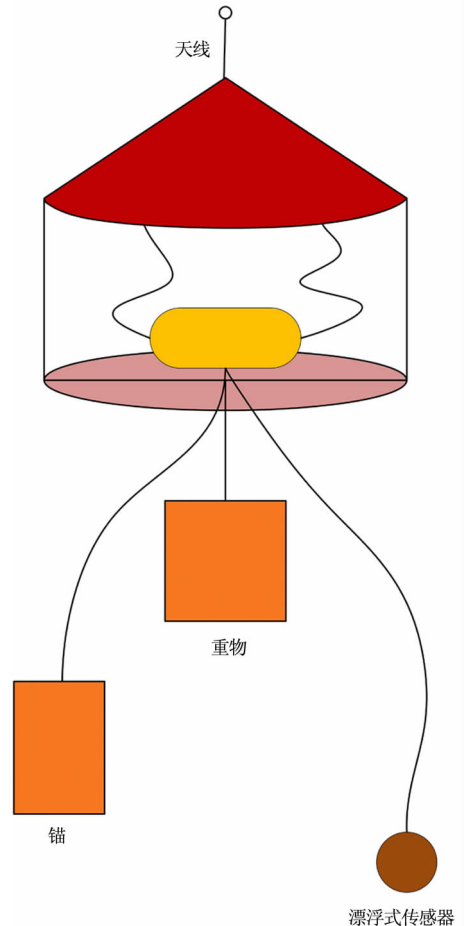


图 1 漂浮式传感器测量原理

基于上述漂浮式传感器所采集到的水质数据,对其进行多源数据融合处理。假设漂浮式传感器的总数为  $n$ ,则水质数据为  $[S]_k = \{[s]_1, [s]_2, \dots, [s]_i, \dots, [s]_n\}$ ,其中,  $[S]_k$  代表第  $k$  个漂浮式传感器所采集到的地表水质监测数据,  $[s]_i$  代表不同指标下的水质参数数据。由此可以得到水质数据与测量误差之间的关系表达式,具体如下所示。

$$[s]_i = [s_i - p_i, s_i + p_i] \quad (1)$$

其中,  $p_i$  代表传感器的测量误差。假设多传感器融合的水质判断框架为  $L = \{l_1, \dots, l_i, \dots, l_n\}$ ,其中,  $l_i$  代表水质分类等级<sup>[11]</sup>。结合传感器监测到的水质参数数据,可以对水质等级特征矩阵进行构建,具体矩阵表达式如下所示。

$$[F]_L = \begin{bmatrix} [f]l1_1 & [f]l1_2 & \cdots & [f]l1_m \\ [f]l2_1 & [f]l2_2 & \cdots & [f]l2_m \\ \vdots & \vdots & & \vdots \\ [f]lt_1 & [f]lt_2 & \cdots & [f]lt_m \end{bmatrix} \quad (2)$$

其中,  $t$  代表水质分类等级的总个数,  $[f]lt_m$  代表水质特征。根据上述表达式,可以整理出水质数据与水质等级之间的映射关系,具体如图 2 所示<sup>[12]</sup>。

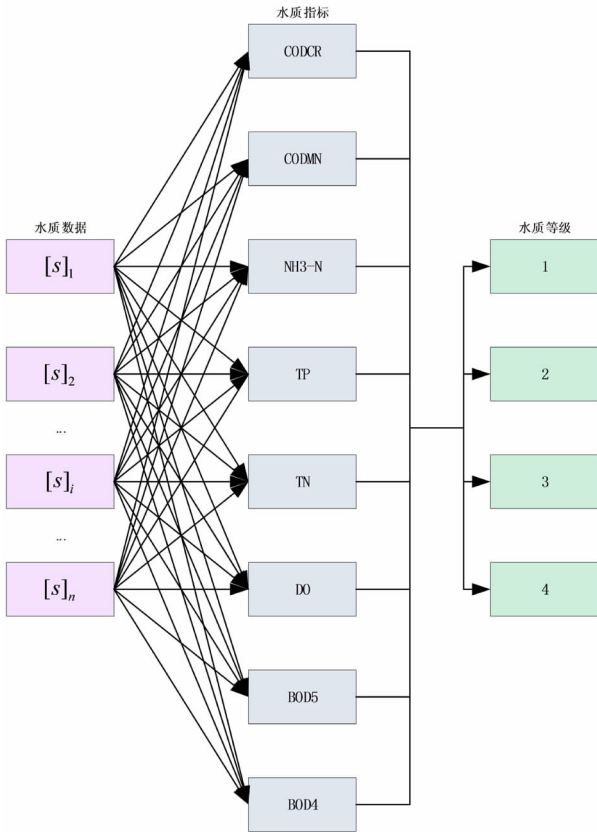


图 2 基于传感数据融合的水质数据与水质等级对应关系

通过上述水质等级与水质数据之间的映射关系,即可实现对水质数据的划分,从而完成对地表径流传感数据的有效预处理,为后续的水质检测提供帮助<sup>[13]</sup>。

### 1.2 基于线性支持向量算法的地表径流水质预测

基于上述融合处理得到的水质监测数据,本文结合线性支持向量算法,构建出水质预测模型,将水质预测数据与实际检测的数据进行对比,从而得到预测残差。通过判断预测残差是否超过设定的异常阈值,从而判定水质等级,实现水质检测<sup>[14]</sup>。因此,本文首先需要结合融合得到的水质参数数据,通过构建水质预测函数,结合约束条件,对地表径流水质进行预测。本文提出的水质检测整体流

程框图如图 3 所示。

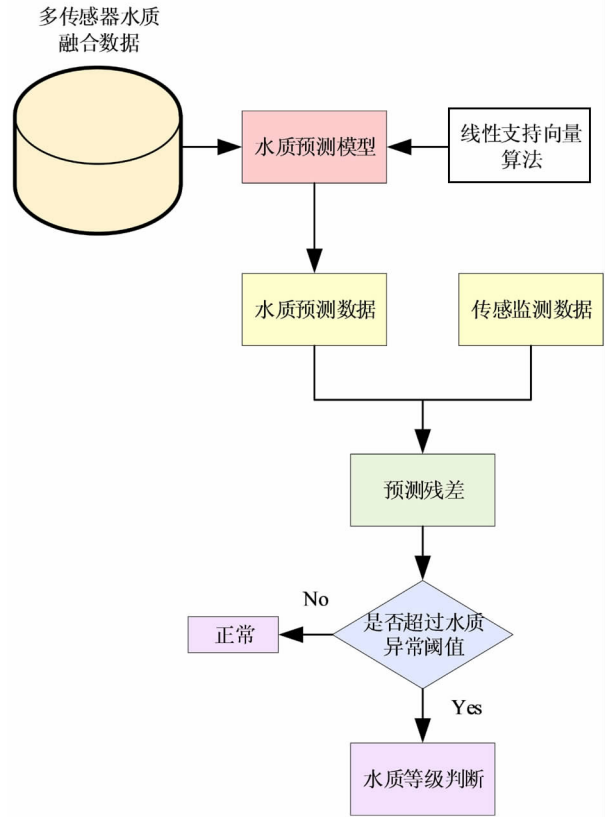


图 3 水质检测流程框图

针对水质数据的多源融合结果,本文结合支持向量算法,对回归函数进行构建,具体表达式如下所示。

$$f(x) = \omega^T \cdot \varphi(x) + b \quad (3)$$

其中,  $x = [x_1^1, \dots, x_i^d]$  代表样本数据集,  $\varphi(x)$  代表水质特征数据的非线性映射函数,  $\omega^T$  代表样本数据的边缘概率分布值,  $b$  代表水质异常判断阈值,  $f(x)$  代表预测值。针对上述构建出的预测函数,本文引入松弛变量对上述函数进行约束。具体约束表达式如(4)所示。

$$\begin{cases} f_i - \omega^T \cdot \varphi(x) - b \leq \theta + \xi_1 \\ -f_i + \omega^T \cdot \varphi(x) + b \leq c + \xi_2 \\ \xi_1 \geq 0, \xi_2 \geq 0, i = 1, 2, \dots, n \end{cases} \quad (4)$$

其中,  $\xi_1, \xi_2$  代表松弛变量,  $c$  代表正则化参数,  $\theta$  代表线性敏感损失函数。通过上述表达式,对构建出的水质预测函数进行约束,从而保证预测数据更贴近真实监测结果<sup>[15]</sup>。

通过上述步骤即可完成对于地表径流水质的有效预测,通过结合线性支持向量算法,构建出水质预测函数,并结合松弛变量,对函数进行约

束<sup>[16]</sup>。下面本文将结合水质异常判断阈值,将水质预测结果与传感器的实际检测结果进行对比,从而实现水质等级的有效判断。

### 1.3 地表径流水质等级判断

基于上述得到的水质预测结果,本文通过对水质异常阈值进行设定,判断预测数据与实测数据之间的预测残差是否超过了设定的异常阈值,结合等级分类指标,对地表径流水质进行检测。地表径流水质的预测残差计算公式如下所示。

$$\epsilon(t) = |y_t - y_t'| \quad (5)$$

其中,  $\epsilon(t)$  代表预测残差序列,  $y_t'$  代表采用线性支持向量算法预测得到的水质数据,  $y_t$  代表采用漂浮式传感器测量得到的水质数据<sup>[17]</sup>。为对预测残差序列的异常阈值进行合理设定,本文结合拉依达准则,对异常阈值的标准差进行分析。拉依达准则下的数据序列分布情况示意图如图4所示。

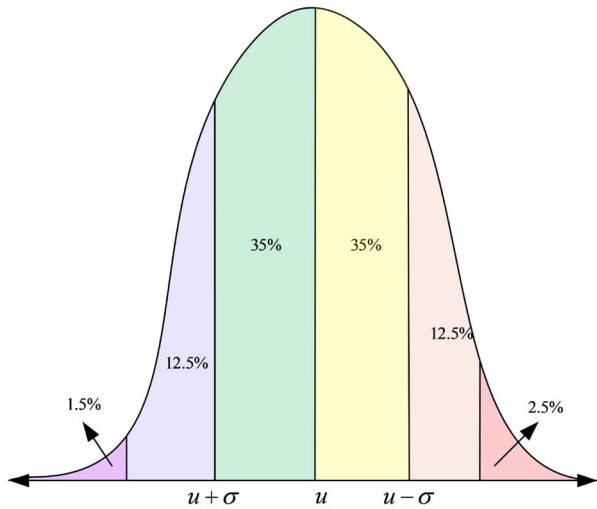


图4 拉依达准则下的数据序列分布情况示意图

根据上述依达拉准则示意图可以看出,残差数据的阈值取  $u \pm \sigma$  时,数据序列分布概率值较大。因此,本文将水质异常判断阈值的取值范围设为  $u \pm \sigma$ ,  $u$  和  $\sigma$  分别代表残差序列的均值和标准差,其具体计算公式如下所示。

$$u = \frac{\sum_{i=1}^n \epsilon(t)}{n} \quad (6)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (\epsilon(t) - u)^2}{n}} \quad (7)$$

通过上述步骤即可完成对于地表径流水质异常阈值的设定,本文结合水质分类标准限制,对水质等级进行划分,具体划分标准如下表1所示。

表1 水质等级划分标准

水质指标	I类	II类	III类	IV类
COD <sub>Cr</sub>	2.5	2.8	2.9	3.2
COD <sub>Mn</sub>	6.0	6.8	7.2	7.4
NH <sub>3</sub> -N	1.0	1.2	1.4	1.6
TP	0.2	0.3	0.4	0.5
TN	1.0	1.1	1.5	1.8
DO	5.4	5.69	6.1	6.3
BOD <sub>5</sub>	4.1	4.2	4.5	4.8
BOD <sub>4</sub>	5.8	5.9	6.3	6.7

当计算得到的地表水质预测数据与实际监测数据之间的残差数据高于设定的异常阈值时,代表当前地表径流水质处于异常状态,结合表1中的水质等级,可以对其进行划分,从而明确当前地表径流的水质情况,从而实现水质检测<sup>[18]</sup>。

通过上述步骤即可完成对于地表径流水质等级的判断,通过引入残差序列阈值,结合水质等级划分标准,对地表径流水质等级进行判断,从而实现水质有效检测。将本节内容与上述提到的传感数据融合以及水质预测等相关内容进行结合,至此,基于传感物联网数据融合和线性支持向量算法的地表径流水质检测方法设计完成。

## 2 实验论证

为了证明本文提出的基于传感物联网数据融合和线性支持向量算法的地表径流水质检测方法的实际检测效果优于常规的地表径流水质检测方法,在理论部分的设计完成后,构建实验环节,对本文方法的调度效果进行检验。

### 2.1 实验说明

为验证本文提出的基于传感物联网数据融合和线性支持向量算法的地表径流水质检测在实际检测效果方面的有效性,本次实验选取了两种常规的地表径流水质检测方法作为对比对象,分别为基于神经网络的地表径流水质检测方法,以及基于机器学习的地表径流水质检测方法。通过构建实验平台,采用三种检测方法对同一组水质数据进行实时检测,对比不同检测方法下的实际检测效果。

### 2.2 实验准备

本次实验选用的水质监测数据来自多参数水质监测仪,通过采用水质监测仪对不同区域内的水体质量情况进行检测,构建出本文的实验数据集。

水质监测仪支持智能分析功能,可以对不同监测节点下的水质数据进行融合处理与分析,同时配备可视化显示功能,因此可以实时读取到不同指标下的水质参数数据。水质监测仪实物图如图 5 所示。



水质监测仪A



水质监测仪B

图 5 水质监测仪实物图

为采集到较为完整的地表径流水质数据,本文在研究区域附近布置了四个漂浮站,每个漂浮站周边均配置了 1~2 个集成站,用于对水质进行集成采样处理。研究区域监测点分布图如图 6 所示。

由于日常水质指标中缺乏足够的异常序列,因此为了模拟出较多异常水质数据,本次实验通过对污染物浓度进行配比,从而生成倒 U 形曲线,实现对地表径流水质异常情况的模拟,具体如图 7 所示。

为保证实验结果的可靠性,本次实验模拟了两种不同的异常叠加强度,分别为 0.2 和 0.8,并采用三种水质检测方法对水质数据进行检测,对比不同异常数据叠加强度下,三种方法的实际检测

性能。

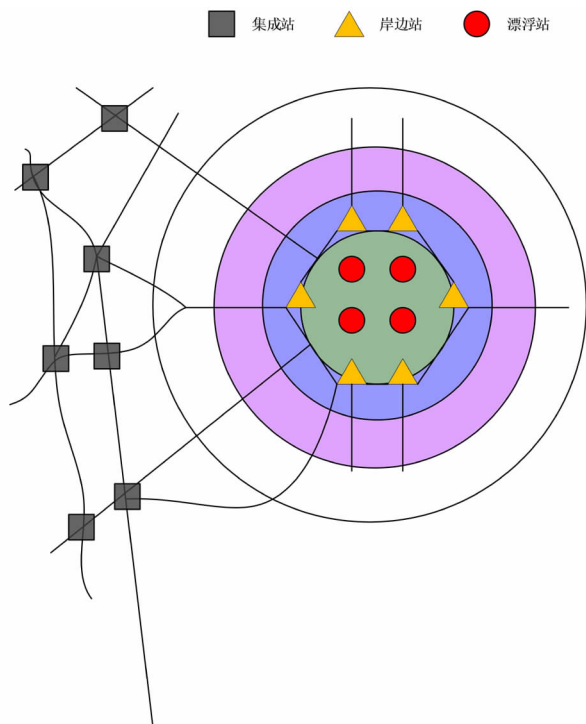


图 6 研究区域水质监测点位分布示意图

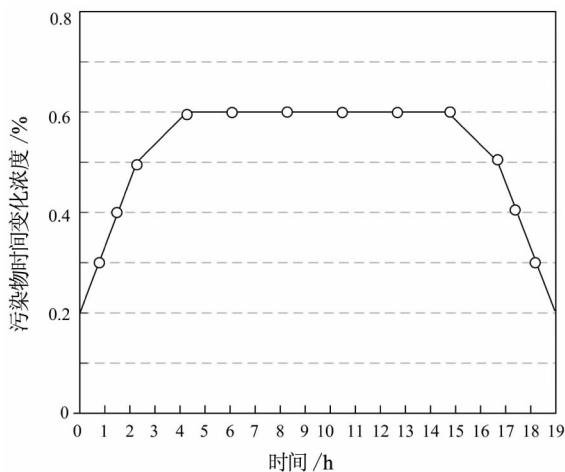


图 7 水质异常情况模拟

### 2.3 检测精度对比结果

本次对比实验选取的对比指标为不同检测方法的检测精度,具体衡量指标为算法的均方根误差,该值越低,代表算法的检测效果越好,具体实验结果如图 8 和图 9 所示。

通过上述实验结果可以看出,异常叠加强度对于算法的检测精度有着一定的影响。通过数值上的对比可以明显看出,本文提出的基于传感物联网数据融合和线性支持向量算法的地表径流水质检测方法在检测精度上明显高于两种常规的水质检测方法,算法的均方根误差值更低。

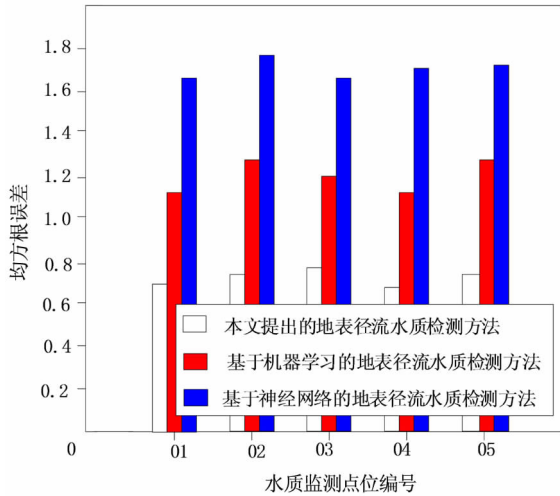


图8 异常叠加强度为 0.2 时的水质检测精度对比结果

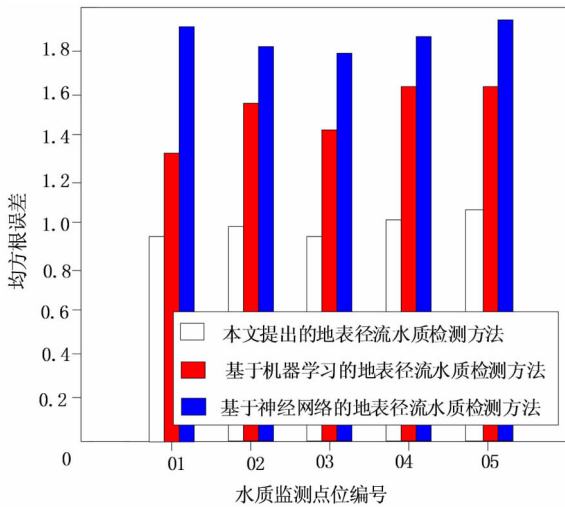


图9 异常叠加强度为 0.8 时的水质检测精度对比结果

### 3 结论

针对常规的地表径流水质检测方法在检测精度方面不够理想的问题,通过结合线性向量支持算法,提出了一种新型的水质检测方法。实验结果表明,本文方法具有更低的均方根误差值,具备较高的准确性和可靠性。在今后的研究工作中,需要继续深入研究水质的分类标准,以便更准确地对不同水质进行预测和检测。调整水质预测模型的权重分布情况,以提高模型的稳定性和适应性。

### 参考文献

- [1] 任丰兰,操政,姚星.基于单片机的水质检查系统设计[J].机电工程技术,2023,52(6):166-169+248.
- [2] 彭子康,张芹,陈世航,等.基于机器学习和图像处理的水质检测方法[J].自动化应用,2023,64(10):188-191.
- [3] 梁桃华,钟子龙,纪丁愈,等.基于 ARM 内核的水质检测装置设计研究[J].大众科技,2023,25(5):1-4.
- [4] 王天钊,钱振明.基于 STM32 的远程无人海洋水质检测系统设计[J].科学技术创新,2023(8):213-216.
- [5] 李文,马俊源,毛潭,等.基于 STM32 的顺序注射水质检测系统设计[J].现代电子技术,2022,45(6):44-48.
- [6] DEMARS C, WANG R Y, GRIENEISEN M L, et al. Assessment of pyrethroid contamination and potential mitigation strategies in california central Coast surface waters[J]. Journal of Environmental Management, 2021, 278(Pt. 1): 111507.
- [7] 李登册,李丽娜,张认成.基于 SPA 和 PSO-LSSVM 的可见-近红外光谱水质 pH 值检测[J].激光与光电子学进展,2023, 60(1):0130002.
- [8] 邝怀秀.微生物检测技术在水质环境监测中应用的质量控制[J].皮革制作与环保科技,2023,4(4):51-53.
- [9] 于银山,张慧,卫晋松,等.基于分光光度法和神经网络的湖泊水质检测方法研究[J].电子器件,2023,46(1):275-281.
- [10] 于夕娟.饮用水水质检测中检出限和测定下限的评价方法与合理性判定[J].食品安全导刊,2023(5):157-159.
- [11] 朱红生,费晓昕,吴述园,等.基于无人船的地表水水质检测系统设计与实现[J].科技与创新,2023(2):1-4+7.
- [12] 冯琳芬,钟彩月,谢雪云.提高污水水质检测准确性及稳定性的策略分析[J].皮革制作与环保科技,2023,4(1):95-97.
- [13] 韦莹.关于废水水质检测化验误差分析与数据处理探讨[J].皮革制作与环保科技,2022,3(19):8-9+12.
- [14] 黄恋媚.关于提高污水水质检测准确性及稳定性的策略探讨[J].皮革制作与环保科技,2022,3(19):10-12.
- [15] 陈宇,邓文清,邱婷婷,等.质量控制图的绘制及在水质常规污染物检测中的应用[J].计量与测试技术,2022,49(9):17-19+22.
- [16] 费晓昕,吴述园,朱红生,等.面向地表水水质检测的无人船上位机软件设计与实现[J].计算机时代,2022(9):49-52.
- [17] 唐自霖,易艺,付金松,等.多传感器融合的水质检测分流控制系统设计[J].电子设计工程,2022,30(17):50-54.
- [18] 周昌海,杨焯,马计划,等.基于 DSP 的无人船水质检测系统设计[J].科技与创新,2022(17):24-26.