徐小燕,霍冉冉,李惠平,等.基于可解释机器学习的供水管网三卤甲烷浓度预测及关键驱动因子识别模型 [J]. 中国环境科学, 2026,46(1):208-218.
Xu X Y, Huo R R, Li H P, et al. An interpretable machine learning-based model for prediction of THMs and identification of key factors in water supply networks [J]. China Environmental Science, 2026,46(1):208-218.

基于可解释机器学习的供水管网三卤甲烷浓度预测及关键驱动因子识别模型

徐小燕[1],霍冉冉[1],李惠平[2*],周佰勤[3],杜振齐[4],庞维海[2],周 婷[1],董晓晨[1],李亚仙[1] (1.苏州工业园区清源华衍水务有限公司,江苏 苏州 215021；2.同济大学环境科学与工程学院,上海 200092；3.哈尔滨工业大学(深圳)土木与环境工程学院,广东 深圳 518055；4.山东建筑大学市政与环境工程学院,山东 济南 250101)

摘要：以苏州某区域供水系统主干管(采样节点间距 2km)为研究对象,对干管的水温,余氯,浊度,pH 值,UV_{254} 和溶解性有机碳(DOC)和 THMs 等多项水质参数沿程变化开展了为期一年的系统性调查研究.空间尺度分析结果表明,供水管网末端 8~10km 处的THMs浓度处于相对较高水平;时间尺度分析结果表明,供水管网末端 THMs 的浓度在 6 月达到最高峰,浓度为 39.06μg/L.选择 4 种典型机器学习算法构建管网 THMs 预测模型,可解释的梯度提升树(GBDT)算法能够有效避免模型过拟合,对 4 种 THMs 预测精度的 R^2 分别达到 0.839,0.906,0.836 和 0.935.进一步应用 Shapley 附加解释(SHAP)对 GBDT 模型对全局解析发现,水温,UV_{254},DOC 和供水距离与 THMs 生成呈正相关关系,余氯与 THMs 呈显著负相关关系;SHAP 的局部解析表明,供水管网中余氯浓度低于 0.45,0.55 和 0.55mg/L 时可显著降低余氯对 TCM,BDCM 和 DBCM 的影响权重.基于本研究构建的数据驱动模型,可通过环境因子高效预测 THMs 的生成浓度,并实时控制余氯投加量防止夏季或管网末端的 THMs 过量生成.

关键词：供水管网；余氯；供水距离；THMs；机器学习；SHAP 分析
中图分类号：X832 **文献标识码**：A **文章编号**：1000-6923(2026)01-0208-11

An interpretable machine learning-based model for prediction of THMs and identification of key factors in water supply networks. XU Xiao-yan[1], HUO Ran-ran[1], LI Hui-ping[2*], ZHOU Bai-qin[3], DU Zhen-qi[4], PANG Wei-hai[2], ZHOU Ting[1], DONG Xiao-chen[1], LI Ya-xian[1] (1.Suzhou Industrial Park Qingyuan Hong Kong & China Water Co. Ltd., Suzhou 215021, China；2.School of Environmental Science and Engineering, Tongji University, Shanghai 200092, China；3.School of Civil and Environmental Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China；4.School of Municipal and Environmental Engineering, Shandong University of Architecture, Jinan 250101, China). *China Environmental Science*, 2026,46(1)：208~218

Abstract: A one-year study was carried out to investigate the changes in water quality parameters such as water temperature, residual chlorine, turbidity, pH, UV_{254}, dissolved organic carbon (DOC) and trihalomethanes(THMs) along the main pipe of a regional water supply system in Suzhou (2km sampling interval). The results of spatial scale analysis showed that the concentration of THMs at the end of the water supply pipe network at 8~10km was at a relatively high level; the results of time scale analysis showed that the concentration of THMs at the end of the water supply pipe network reaches the highest peak in June, with the concentration of 39.06 μg/L. Four typical machine learning algorithms were selected to construct the prediction model of THMs in the pipe network, and the interpretable gradient boosting tree (GBDT) algorithm was able to effectively avoid model overfitting, and the R^2 of the four THMs prediction accuracies reached 0.839, 0.906, 0.836, and 0.935, respectively. The further application of Shapley's additional explanation (SHAP) to the global analysis of the GBDT model revealed that the water temperature, UV_{254}, DOC, and the distance of the water supply were positively correlated with the generation of THMs, while the residual chlorine showed a significant negative correlation with the THMs. The local analysis of SHAP showed that residual chlorine concentration in the water supply network lower than 0.45, 0.55 and 0.55mg/L can significantly reduce the influence weights of residual chlorine on TCM, BDCM and DBCM. Based on the data-driven model constructed in this study, the generation concentration of THMs can be efficiently predicted by environmental factors, and the residual chlorine dosage can be controlled in real time to prevent the over-generation of THMs in summer or at the end of the pipe network.

Key words: water supply network；residual chlorine；distance to water supply；THMs；machine learning；SHAP analysis

氯消毒已成为饮用水处理厂(DWTP)最常用的消毒手段[1].然而,含氯消毒剂容易与饮用水供水系统(DWDS)中残留的消毒前体物发生反应,产生多种

收稿日期：2025-05-19
基金项目：应急管理部四川消防研究所基科费项目(20238814Z)
* 责任作者, 博士研究生, 2210525@tongji.edu.cn

消毒副产物(DBPs).三卤甲烷(THMs)为饮用水中率先被甄别出的 DBPs[2-3].毒理学与流行病学研究广泛证实了 THMs 的毒性与致癌性[4-5].因此,世界上多数国家在制定饮用水卫生标准时均将 THMs 纳入标准管控范围.在 DWDS 中,加氯工艺需遵循的重要原则是如何以最低的氯投加量实现最优的持续杀菌效果,同时减少消毒副产物的生成量.因此,建立精确的 THMs 预测模型,通过输入底物浓度,环境因子和反应时间等因素高效预测 THMs 的生成浓度,对 DWDS 的 THMs 风险控制至关重要.

近年来,机器学习(ML)作为一种大数据模型算法被广泛应用于饮用水处理及输配过程中,水质预警,智能配水,DBPs 生成量的预测,精准加氯等[6-7].有研究应用一般回归神经网络模型,通过流量和水温数据预测了 DWDS 游离余氯的变化,建立的 ML 模型能够有效预测饮用水中 9 种卤乙酸(准确率达到 75%~91%)[8].Li 等[9]发现,贝叶斯网络能够建立溶解性有机碳(DOC),总溶解性固体(TDS)与三卤甲烷生成浓度间的相关模型,预测精度最高可达 94%.Bowden 等[10]尝试选择 pH 值,温度,UV_{254} 和余氯等四个简单的水参数,建立了基于 LRM 和 ANNs 模型,用来预测自来水中的 THMs.相比于传统的机理模型,上述研究巧妙的利用机器学习算法构建了复杂输入因子与 THMs 预测的数据驱动模型,但基于机器学习的模型在应用中也存在局限性[11].即“黑箱”模型拟合无法像传统机理模型得出关键参数,进而判断每个影响因子的内在关系与权重.为解决 ML 的“黑箱”问题,相关研究应用 Shapley 附加解释(SHAP)或部分依赖图(PDP)对模型输入变量因子的 SHAP 贡献做出了分析,从而帮助我们进一步理解输入特征与输出特征之间的直接关系[12-13].例如,于林堂等利用可解释机器学习算法预测芬顿工艺中的污染物降解效率,并对关键影响因子做出识别[14].张煦等[15]使用可解释 ML 识别影响水库总磷浓度的关键因素,结果显示水位,浊度和出库流量对总磷浓度影响较大.在应用机器学习技术构建 DWDS 高效预测模型的过程中,建立模型在实际工程的应用能力以及模型的可解释性是需要重点评估的.目前,国内鲜有研究应用可解释 ML 算法对饮用水供水管网 THMs 预测模型,并应用可解释算法实现关键因子的识别以及输入变量与输出变量逻辑关系的梳理.

本文选取苏州市某区域供水干管开展系统研究,对 THMs 浓度及其他常见水质参数(如余氯,浊度,pH 值,UV_{254} 和 DOC)进行了为期一年的监测,进而探究 THMs 在 DWDS 生成的时空规律.同时,应用 ML 技术构建了能够快速和准确预测管网 THMs 生成的算法模型,实现 THMs 的精准与高效预测.为避免 ML 存在的“黑箱”问题,通过 SHAP 技术解析机器学习模型结果,识别底物浓度(余氯,UV_{254} 和 DOC),环境因子(浊度和 pH 值)和供水距离与 THMs 生成浓度间的内在关系.研究结果为城市供水管网 THMs 的精准预测及高效预警提供技术参考.

1　材料及方法

1.1　取样点选择及样品分析

选择苏州市某区域 X 水厂具有代表性的主干管为研究目标,该水厂采用常规+臭氧活性炭深度处理技术(具体工艺流程如图 1(a)),设计最大供水规模为 45 万 m^3/d,原水水质如表 1 所示.

表 1　水厂原水水质

Table 1　Raw water quality of DWTP

指标	单位	春	夏	秋	冬
水温	℃	18.0±4.7	28.7±1.9	21.2±5.7	9.6±2.7
pH 值	–	8.5±0.4	8.7±0.4	8.1±0.2	8.1±0.1
浊度	NTU	1.1±0.2	1.2±0.2	3.3±7.8	5.0±6.8
电导率	μS/cm	383±28	350±14	339±7	410±7
藻细胞	10^4cells/L	34±14	70±61	126±79	123±106
UV_{254}	cm^{-1}	0.048±0.004	0.046±0.008	0.049±0.003	0.048±0.004
DOC	mg/L	3.4±0.3	3.1±0.3	3.2±0.1	3.3±0.2
Br^-	mg/L	0.09±0.03	0.10±0.04	0.09±0.03	0.11±0.05

由于居民用水需求差异,其月均供水量在夏季达到最大,秋冬季节次之(图 1(c)),水温也呈现类似变化规律.该水厂服务的供水区域如图 1(b)所示.

本研究以出水作为参照点,2km 为步长取样至管网末端,其对应距离分别为 0,2,4,6,8 和 10km,对应编号分别为 X−0,X−2,X−4,X−6,X−8 和 X−10.从 1~12 月,每月取样分析上述 6 个点位的水质.采样过程中,为避免供水管网中水龄影响,先放水 5min,然后取 1L 水样加入适量抗坏血酸作为还原剂,冷藏于 0~4 ℃环境中转移至实验室开展分析.其中余氯和水温利用便携式余氯分析仪及便携式溶解氧测定仪在现场测试,其余指标在将样品带回水厂后立即完成测试.

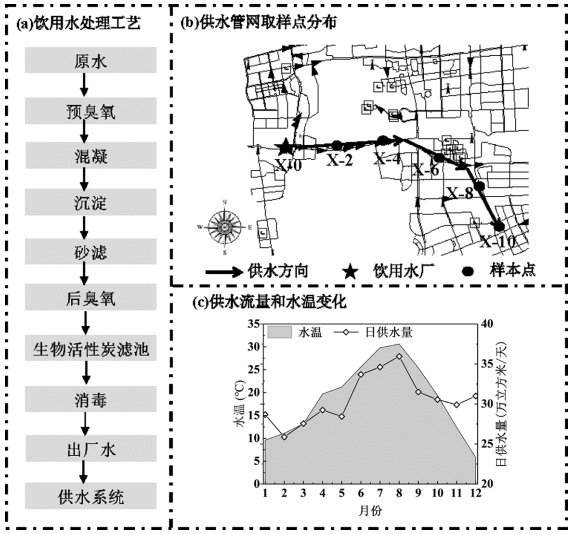


图 1 饮用水处理工艺、供水管网取样点分布及供水流量和水温变化

Fig.1 Drinking water treatment process, distribution of sampling pointsin the water supply network,and changes in water supply flow rate and water temperature

1.2 水质指标分析

DOC 测定采用总有机碳分析仪(O·I·Analytical Aurora Model 1030W);UV_{254} 测定采用紫外可见分光光度计(EVO300PC,Thermo);浊度测定采用 HACH TL2300 浊度仪;藻细胞浓度测定采用多功能水质分析仪(YSI-EXOSonde2);pH 采用酸度仪(PHSJ-3F,上海雷磁)测定;余氯采用 HACH-DR300 便携式余氯仪测定.

三卤甲烷测定依据顶空毛细管柱气相色谱法[16](GB/T 5750.10-20234.1)由 Agilent 7697A/8890 顶空气相色谱仪检测,ECD 电子捕获检测器,色谱柱为 HP-5(30m×0.32mm×0.25μm).取样时采用预先加入抗坏血酸的钳口顶空瓶,以中和管网水中余氯,并快速带回实验室完成分析.气相条件:进样口温度为 200℃;检测器温度为 250℃;分流比为 5:1;初始柱温为 50℃保持 1min,然后以 2.5℃/min 升至 67.5℃保持 0min,再以 15℃/min 升至 120℃,保持 0min;载气为高纯氮气.顶空条件:水样体积为 5mL,样品瓶容积为 20mL,平衡温度为 50℃,平衡时间为 60min.

1.3 指数函数模型

经过函数拟合效果评估,最终选择指数函数用于建立不同月份下供水距离与余氯和 THMs 浓度的生成模型,然后绘制时空分布热图,计算公式(1)所示:

$$y = y_0 + A_1 e^{\frac{-x}{t_1}} \tag{1}$$

式中:y_0 为截距;A_1 为指数系数;t_1 为修正系数.

1.4 机器学习算法

为建立物浓度(余氯,UV_{254} 和 DOC),环境因子(水温,pH 值)和供水距离对 THMs 的数据驱动模型,选择随机森林(RF)和梯度提升树(GBDT)两种机器学习算法进行建模,RF 和 GBDT 算法基于分类与回归树(CART)完成预测,通过构建二叉树将样本递归划分,分裂节点以样本最小方差为依据,上述算法的具体介绍及原理参见文献[17-19].通过对比拟合精度评估模型适用性,并选取最优算法建立供水系统的 THMs 预测模型.

对于回归任务的 RF 算法,最终预测结果如公式(2):

$$F(x) = \frac{1}{N}\sum_{i=1}^{N} h_i(x) \tag{2}$$

式中:N 为决策树个数;$h_i(x)$为每棵树的预测结果;$F(x)$为最终预测结果.

基于 GBDT 算法的计算过程如下:

首先设真实函数为 $F(x)$,假设一组基函数 $h(x)$的线性组合可无限逼近 $F(x)$,具体见公式(3).

$$F(x) \approx \sum_{i=1}^{N} h_i(x) \tag{3}$$

$h_i(x)$为其中一颗决策树,根据加法模型,由前往后逐步拟合残差,即对任一样本 x,在第 i 轮迭代中拟合模型所产生的残差,具体计算见公式(4)

$$r_{\text{in}} = -\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} F(x) = F_{i-1}(x_i) \tag{4}$$

式中:$L(y_i,F(x_i))$是损失函数,$F(x)$为样本 x_i 的真实标签值,$F_{i-1}(x_i)$是前 i-1 轮迭代得到的预测值.

利用新的决策树拟合上述残差,以降低残差数值,加入一个新的分段函数 $h_i(x)$,使得模型预测值更新为公式(5).

$$F_i(x)=F_{i-1}(x)+h_i(x) \tag{5}$$

1.5 机器学习结果验证与分析

为评估 RF 和 GBDT 算法建立的模型精度,采用十折交叉验证(10-fold CV)方法来调整超参数,训练数据被分为 10 个子集,在每个训练迭代中,一个子集被用作验证集,其他 9 个子集构成训练集,以提高模型的预测和泛化能力.模型性能由决定系数(R^2),平均平方误差(MSE)和平均绝对误差(MAE)评价,计算方法如公式 6,7 和 8 所示.根据计算原理,R^2 越高,

MSE 和 MAE 越低,则模型准确性就越高.

$$R^2=1-\frac{\sum_{n=1}^{N}(F(x)-F_i(x_i))^2}{\sum_{n=1}^{N}(F(x)-\overline{y})^2} \tag{6}$$

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(F(x)-F_i(x_i))^2 \tag{7}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|F(x)-F_i(x_i) \tag{8}$$

式中:$F(x)$、$F_i(x_i)$和$\overline{y}$分别是目标特征的预测值,实际值和平均值,n是任何特定实例的数据点,N是数据点的总数.

1.6　机器学习模型可解释性

为评估每个特征对 THMs 形成的贡献和重要性,本研究分配了 Shapley 附加解释(SHAP)值,以提高 GBDT 和 RF 算法结果的可靠性和有效性,主要计算公式如式(9).SHAP 值是指博弈论中的最优夏普利值,它为联盟成员之间的报酬分配提供了一种合理的方法,其特征值的子集代表联盟,子集的预测值代表联盟的报酬,具体介绍见文献[20].

$$g(z')=\varphi_0+\sum_{i=1}^{N}\varphi_i z_i' \tag{9}$$

式中: N 指特征值,φ_i 代表特征 i 的 SHAP 值,z'_i 指在 SHAP 计算过程中,简化的特征 i.

2　结果与讨论

2.1　常规水质变化

供水过程中常规水质指标变化如图 2 所示.根据图 2(a),1 月,2 月和 3 月 pH 值较高,10 月,11 月和 12 月 pH 值较低,其余月份介于中间.这可能受原水水质影响,根据表 1 可知春季和冬季原水的 pH 值分别为(8.5±0.4)和(8.1±0.1).春夏季节 pH 值高于秋冬季节的主要原因可能如下:苏州地区春夏季节降雨强度较大,雨水径流将碱性物质冲入水源地;夏季水温较高导致二氧化碳溶解度减小,导致原水 pH 值升高;夏季蓝藻繁殖消耗二氧化碳,导致 pH 值升高.根据图 2(b),浊度在 0.06~0.27NTU 之间变化,满足江苏省高品质饮用水标准的规定值(<0.5NTU)[21].由于原水经过常规+深度处理工艺净化,因此水厂出水浊度并未受原水浊度的季节性变化影响.相反,在供水过程中浊度指标的数值变化显著.在 X−0,X−2,X−4,X−6,X−8 和 X−10,浊度平均值分别为(0.10±0.02),(0.13±0.04),(0.13±0.02),(0.17±0.05),(0.15±0.06) 和(0.11±0.03)NTU,供水管网的浊度指标呈现先增高后降低的规律,在 X−6 和 X−8 区域波动最大.浊度升高的原因可能是由于输水过程中水力冲刷作用导致管壁生物膜,颗粒物等的释放.徐鹏等人也证实浑浊度受季节变化影响小,受管道和水力条件影响较大,尤其是供水交汇处的浊度指标波动最大[22].UV_{254} 与 THMs 生成潜能显著相关,该指标可作为 THMs 的指示参数[23].

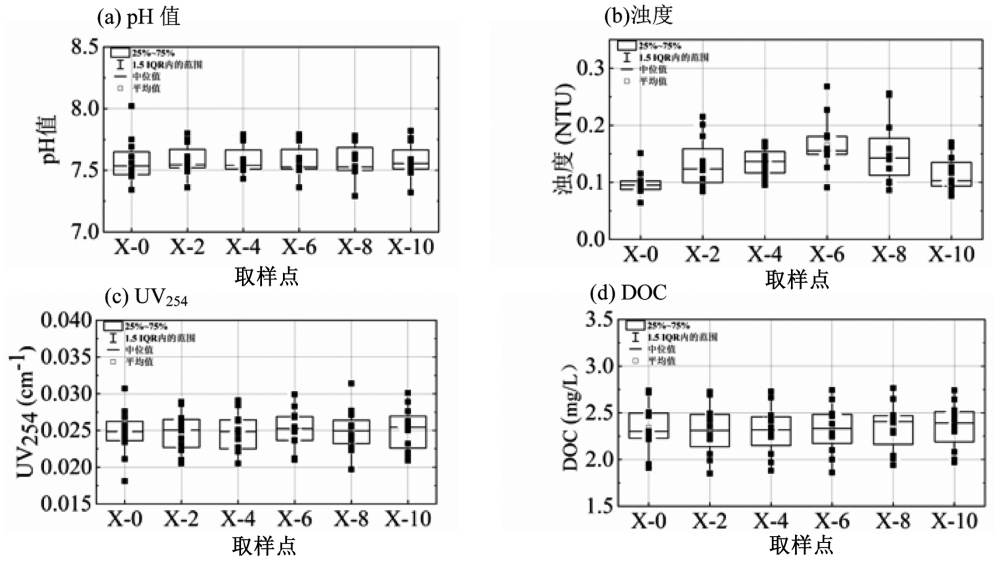


图 2　常规水质沿程变化

Fig.2　Changes in conventional water quality along the course

根据图 2(c),0~4km 的 UV_{254} 较为稳定,在 6km 处显著增加,这可能是由于余氯将部分大分子有机

物氧化为具有碳碳双键的不饱和烃类物质.图 2(d)为 DOC 在供水干管的变化规律,在 X-0,X-2,X-4,X-6,X-8 和 X-10 处 DOC 的平均浓度分别为(2.34±0.26),(2.32±0.27),(2.31±0.26),(2.33±0.27),(2.35±0.25)和(2.36±0.24)mg/L,供水末端的有机物浓度高于出厂水的,这可能与管壁生物代谢或者有机质脱落有关.相关研究也证明,球墨铸铁管壁易于形成生物膜,管壁生物膜对营养物质吸附利用,脱落释放会对管段主体水中有机物特性与浓度产生动态影响[24-25].

2.2 余氯及消毒副产物变化

余氯衰减和 THMs 的生成浓度是 DWDS 中水质安全评估的重要指标.依据水质标准规定,供水管网末梢的游离性余氯不得低于 0.05mg/L,THMs 的生成浓度不得高于 1.0[26].图 3(a)为余氯随供水距离增加的变化规律,X-0 的余氯浓度最高,X-10 达到最低.余氯衰减速率受季节影响较为显著,因此按季节对余氯的变化规律进行分析.X-0 的余氯平均浓度在春,夏,秋和冬季分别达到(0.69±0.04),(0.77±0.11),(0.79±0.04)和(0.70±0.02)mg/L;X-10 对应的余氯平均浓度分别为(0.23±0.06),(0.20±0.04),(0.31±0.10)和(0.39±0.04)mg/L,余氯的平均衰减量分别为 0.46,0.57,0.48 和 0.31mg/L.相关研究影响供水管网余氯衰减的主要因素有水温,水力停留时间,水质和管材等[27].当用水量上升时,水力停留时间较短,水量更替速度加快会导致氯浓度保持相对较高的水平.此外,水温会影响供水管网中细菌和病毒的繁殖速率以及水中 DOM 与氯的氧化反应,水温升高会导致耗氯量增加[28].

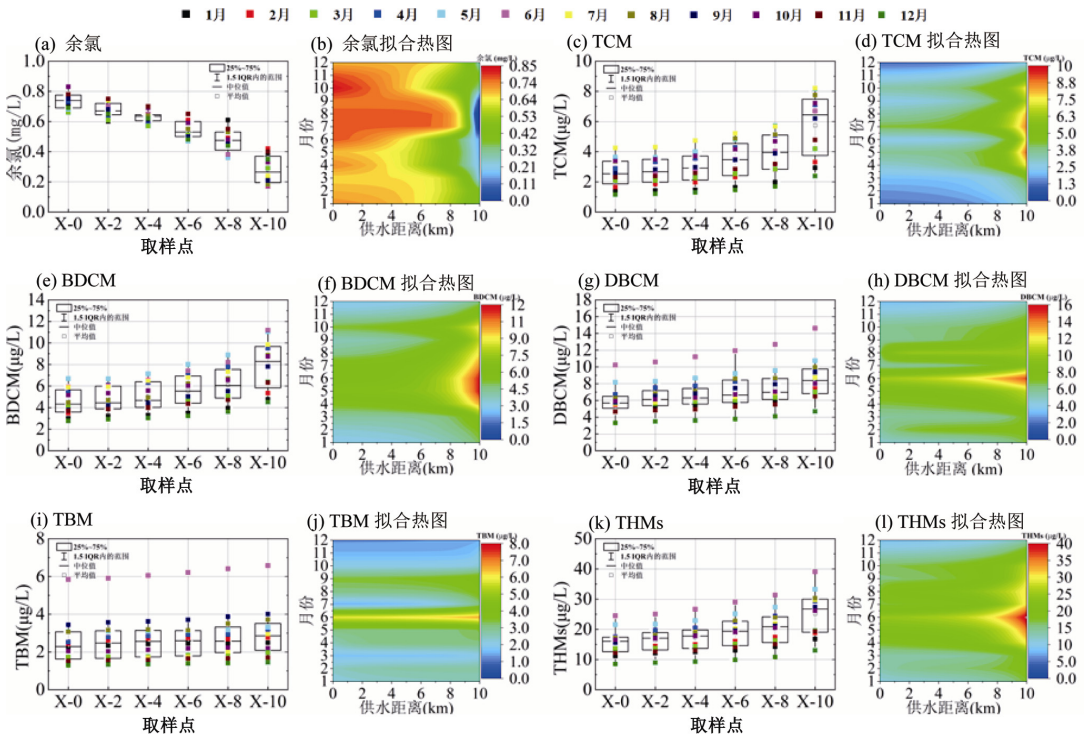


图 3 余氯和 THMs 延长变化及指数模型拟合热图

Fig.3 Prolonged changes in ECC and THMs and exponential model fit thermograms

THMs 浓度则随供水距离显著升高.根据图 3(c),X-0 的 TCM 最低,春,夏,秋和冬季 TCM 的平均浓度分别为(2.87±0.77),(3.37±0.87),(2.76±0.56)和(1.38±0.25)μg/L,而在 X-10 对应的 TCM 浓度分别达到(6.48±2.06),(7.55±0.77),(6.06±1.22)和(2.87±0.47)μg/L.由 X-0 至 X-10,TCM 的平均浓度增量分别为 3.61,4.18,3.31 和 1.49μg/L.在由水厂供水至管道末端的过程中,夏季 TCM 生成量最多,冬季最少.根据图 3(e),X-0 处 BDCM 的平均浓度分别为(5.43±1.24)(春),(5.42±0.85)(夏),(4.22±0.82)(秋)和(3.08±0.36)(冬)μg/L,对应季节管网末端(X-10)的 BDCM 平均浓度分别为(8.99±2.38),(9.96±1.17),(7.63±1.22)和(4.87±0.43)μg/L,BDCM 平均浓度增量分别为 3.55,4.54,3.41 和 1.79μg/L,BDCM 在夏季的

生成浓度仍为最高. 从 X−0 至 X−10,DBCM 的浓度增量分别为 2.28,3.73,2.50 和 1.75μg/L(图 3(g)).根据图 3(i),TBM 的生成量最低, 在 X−0 点位 TBM 的平均浓度分别为 (2.44±0.67)(春),(3.47±2.19)(夏),(2.34±0.99)(秋)和(1.96±0.60)(冬)μg/L,对应季节末端管段 (X−10) 的 TBM 浓度也仅为 (2.81±0.73),(4.41±1.91),(2.64±1.21)和(2.23±0.69)μg/L,TBM 浓度增量仅为 0.37,0.94,0.30 和 0.28μg/L,但夏季的浓度增量仍远高于其他季节.卢小艳等人对南方某河流原水中的 THMs 前体物季节性研究发现,夏季藻类繁殖会产生更高浓度的藻毒素,夏季 THMs 前体物浓度显著高于其他季节[29].

值得注意的是,DBCM 和 TBM 存在显著的异常点,6 月 DBCM 在 X−0,X−2,X−4,X−6,X−8 和 X−10 浓度分别为 10.21,10.59,11.20,11.92,12.69 和 14.60μg/L,对应月份和点位 TBM 的浓度分别为 5.84,5.91,6.05,6.21,6.41 和 6.57μg/L,该月份 DBCM 和 TBM 的浓度远高于其他月份,说明 6 月份管网末端的 DBCM 和 TBM 含量最高,风险最大.这可能与原水中溴离子浓度有关,6 月份水源水受藻类影响较大因此臭氧投加量较多,溴酸根的高反应性导致了卤素取代反应的发生.相关研究也表明[30],当溴离子含量达到 90μg/L 时,氯化消毒处理后生成的 THMs 中有 54%为溴氯混合型.根据表 1,本研究中夏季原水的溴离子浓度达到(100±40)μg/L,这是导致溴代三卤甲烷生成浓度高的主要原因.因此,夏季导致 THMs 生成浓度高于其他季节的主要原因有:(1)较高的水温加快了余氯与 THMs 前体物的反应速率;(2)夏季藻类生长释放易于生成 THMs 的前体有机物.

为进一步明晰余氯和 THMs 随季节和供水距离的时空生成规律,应用指数函数拟合的模型结果绘制余氯和 THMs 时空分布热图,拟合参数结果和精度如表 2 所示,拟合热图如图 3(b),(d),(f),(h),(j)和(l)所示.根据表 2,春夏秋冬季节余氯的拟合参数数值相差较小,而四种 THMs 的参数则随季节变化相差较大,尤其是夏季 TCM、BDCM、DBCM 和 TBM 的 y_0 分别达到 3.44,5.23,6.72 和 3.37,明显高于其他季节的 y_0,对应的 t_1 参数的绝对值也显著低于其他季节的.根据图 3(b),余氯的时空分布较为均匀,而 4 种 THMs 的时空分布差异较大,TCM、BDCM 和 DBCM 在 4~9 月的 X−10 区域出现浓度峰值,TBM 的浓度峰值则出现在 5~6 月的 X−10 区域.供水管网末端(8~10km)的 THMs 浓度需重点关注,且 4~9 月处于 THMs 浓度高峰,6 月份达到最高峰(THMs 浓度达到 39.06μg/L).根据时空分布热图可知:(1)时间分析上 THMs 生成浓度较高的月份在 6 月份左右;(2)空间分析表明管网末端 8~10km 范围的 THMs 风险较高,因此需重点关注对 6 月份管网末端的 THMs.

表 2　供水距离与余氯、THMs 的拟合参数

Table 2　Fitting parameters of water supply distance with residual chlorine and THMs

指标	季节	y_0	A_1	t_1	R^2
余氯	春	0.81	−0.12	−6.38	0.989
	夏	0.92	−0.15	−6.29	0.996
	秋	0.92	−0.14	−7.14	0.989
	冬	0.90	−0.21	−10.99	0.984
TCM	春	2.30	0.16	−9.09	0.933
	夏	3.44	0.05	−3.34	0.992
	秋	1.38	0.95	−35.41	0.993
	冬	2.23	0.27	3.03	0.979
BDCM	春	5.19	0.26	−3.78	0.992
	夏	5.23	0.23	−3.33	0.959
	秋	4.27	0.08	−3.65	0.972
	冬	2.78	0.33	−5.44	0.993
DBCM	春	6.28	0.60	−6.57	0.981
	夏	6.72	0.69	−5.46	0.978
	秋	4.94	0.55	−6.02	0.972
	冬	3.67	0.88	−9.56	0.964
TBM	春	2.72	0.16	−3.17	0.992
	夏	3.37	0.07	−2.43	0.977
	秋	2.80	0.04	−2.26	0.965
	冬	1.31	0.09	−3.47	0.998

2.3　机器学习模型建立

为阐明不同因素与 THMs 生成之间的关系,选择两种流行的 ML 模型—RF 和 GBDT 建立用于预测 THMs 的数据驱动模型.共收集到 142 组数据,输入特征包括供水距离,水温,余氯,浊度,pH 值,UV_{254} 和 DOC,输出为 TCM、BDCM、DBCM 和 TBM.模型的内置功能可计算出各变量的灵敏度系数,以便选择关键因素来优化模型.基于 RF 和 GBDT 算法建立模型的 THMs 实际值与预测值分布如图 4 所示.RF 算法对 TCM、BDCM、DBCM 和 TBM 的测试集 R^2 分别达到 0.804,0.788,0.781 和 0.795(图 4(a)).相比于 RF,GBDT 算法展现出更为优异的拟合性能,测试集 R^2 分别达到 0.839,0.906,0.836 和 0.935(图

4(c)).同样地,GBDT 算法的训练集 R^2 也显著高于 RF 的,但 RF 算法训练集 R^2 与测试集 R^2 的差值明显高于 GBDT 算法的,这说明 RF 算法相比较 GBDT 算法具有更高的过拟合风险.

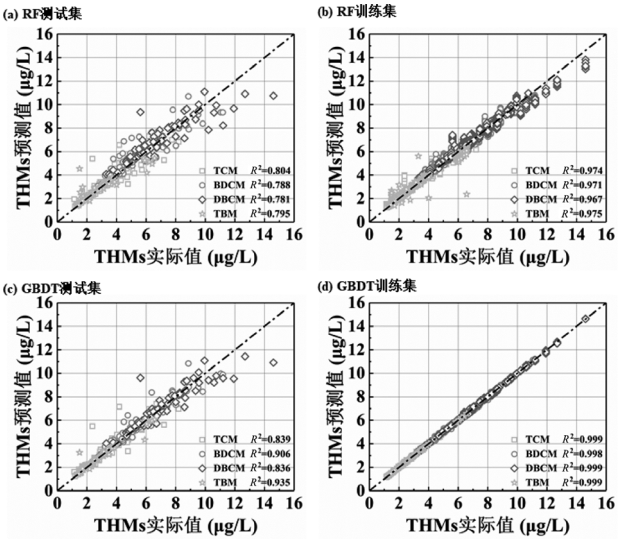


图 4 基于 RF 和 GBDT 建立的机器学习模型结果

Fig.4 Results of machine learning models built based on RF and GBDT

表 3 模型因子选择及测试 R^2、MSE 和 MAE 结果

Table 3 Model factor selection and test R^2, MSE and MAE results

模型算法	输入因子	目标结果	测试 R^2	MSE	MAE
RF 算法	水温,余氯,供水距离,pH 值	TCM	0.804	0.556	0.507
	水温,余氯,供水距离	BDCM	0.788	0.782	0.642
	水温,余氯,UV_{254},供水距离,pH 值	DBCM	0.781	1.032	0.662
	水温,DOC,余氯,pH 值,UV_{254}	TBM	0.795	0.320	0.291
GBDT 算法	水温,余氯,供水距离,pH 值	TCM	0.839	0.457	0.395
	水温,余氯,供水距离	BDCM	0.906	0.346	0.389
	水温,余氯,UV_{254},供水距离,pH 值	DBCM	0.836	0.772	0.553
	水温,DOC,余氯,pH 值,UV_{254}	TBM	0.935	0.103	0.159

此外,基于 GBDT 算法拟合模型的 MSE 和 MAE 数值显著低于 RF 算法的(表 3),此结果在其他相关研究也得以证明,GBDT 算法在避免过拟合方面展现出更加优异的性能[31-32].基于算法原理与思想,RF 算法是基于决策树的集成学习算法,通过随机选择样本自己和特征训练多颗决策树实现分类和回归功能,最终建立模型;而 GBDT 算法则是基于决策树的迭代优化算法,并通过对决策树反复迭代,对计算残差进行拟合,最终损失函数得到最小值或达到预定的迭代次数后建立预测模型[33-34].Li 等[17]在建立双水源供水系统的模型时,也发现基于 GBDT 建立的数据驱动模型对 THMs 的预测精度 R^2 达到 0.9 以上.本研究中,基于多颗决策树的残差拟合 GBDT 算法展现出更为优异的性能.

2.4 模型可解释性及驱动因素交互影响分析

尽管机器学习在建立多参数 THMs 预测模型方面展现出优异性能,但其存在的模型“黑箱”问题对模型解释造成困难.因此,本研究应用 SHAP 算法对预测性能最优异的 GBDT 模型做出解析,并对模型可解释性及驱动因素交互影响做详尽分析.从图 5 中可以看出,红点分布在正坐标轴上,蓝点分布在负坐标轴上,水温与 THMs 浓度之间存在显著的正相关关系.供水管网中的水温从 5.8℃(12 月)到 30.6℃(8 月)不等,证实温度对 DWDS 中 THMs 的形成起着非常重要的作用.作为关键反应底物的余氯与三卤甲烷的生成呈负相关(红色圆点分布在负轴上,蓝色圆点分布在正轴上),余氯作为造成 THMs 生成的关键反应底物,随着其浓度降低,THMs 生成浓度则越高.余氯与消毒副产物前体物间的反应效率受水温影响,水温越高反应活性越强,THMs 的生成速率越快.Wang 等[35]人对三卤甲烷与管道水温之间关系的研究也得出了类似的结论.供水距离与 TCM,BDCM 和 DBCM 的生成呈正相关关系,这主要是因为随着供水距离的增加余氯与消毒前体物的反应时间增加,供水时间等于供水距离/流速,因此在距离水厂越远距离的 THMs 浓度越高.王玲等人对成都市某区域生活饮用水中消毒副产物含量及影响因素分析的研究中发现,供水距离与 THMs 含量呈正相关[36].TBM 的生成效率和浓度则主要取决于水中溴离子的浓度和 DOC 浓度[37].一般而言,较低的 pH 值有利于三卤甲烷在饮用水输送过程中的形成,酸性环境(pH<5 时)下水中的氯主要以 HOCl 形式存在,更易与水中的 DOM 发生反应形成氯代消毒副产物,如三氯甲烷.也有研究指出,高 pH 值会导致高分子量 DBPs 的分解,进而导致 THMs 赋存浓度的增大[38].在本研究中,pH 值对 THMs 的生成影响较弱,呈现负相关关系,这主要是供水管网的 pH 值维持在 7.0~ 8.0,pH 值作为环境因子对 THMs 生成影响不如其他条件显著.

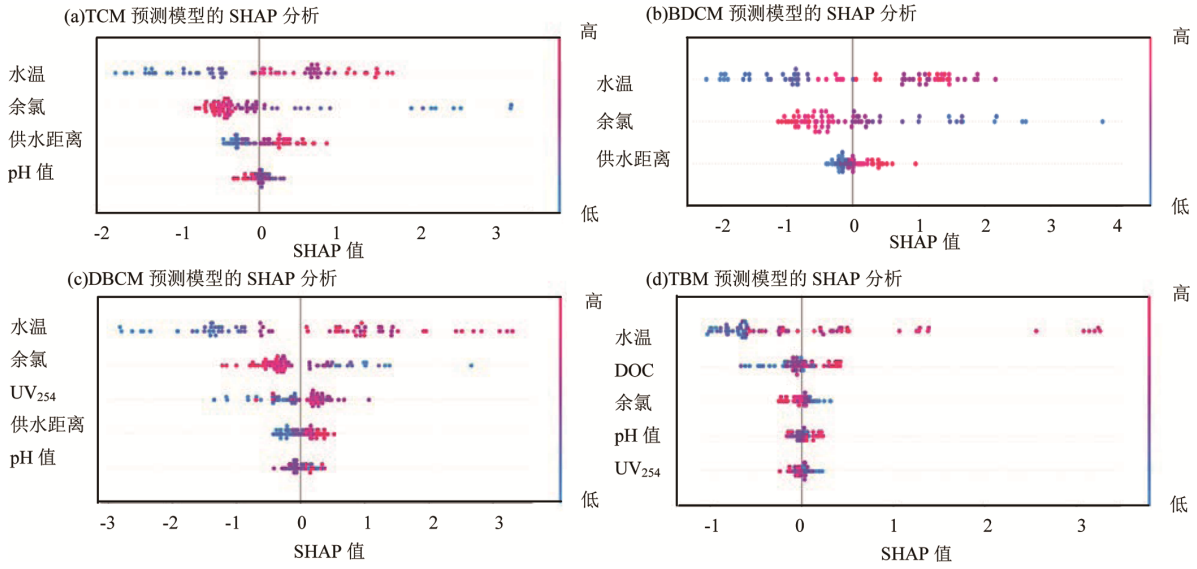


图 5　基于 SHAP 算法的 GBDT 模型结果解析

Fig.5　Analysis of GBDT model results based on SHAP algorithm

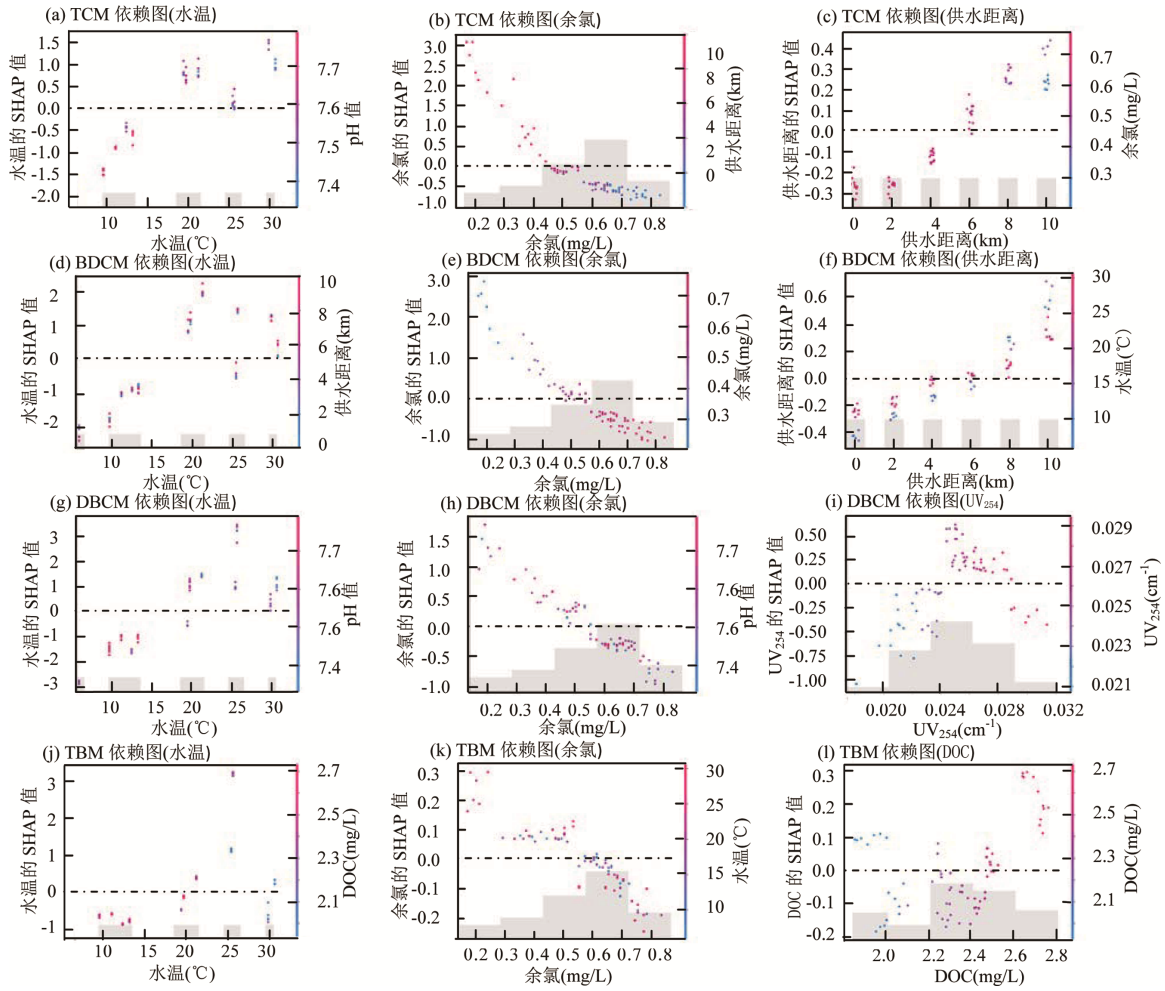


图 6　基于 SHAP 算法的部分依赖图

Fig.6　Partial dependency graphs based on SHAP algorithm

通过计算每个特征值对模型输出的贡献,进一步探究 GBDT 模型中各驱动因素对 THMs 浓度变化的局部影响(图 6).SHAP 值的零线(黑色虚线)用以区别关键因子对 TCM,BDCM,DBCM 和 TBM 所产生的正

面影响和负面影响.水温与 4 种 THMs 的 SHAP 值呈正相关.根据图 6(a)和(d),当水温大于 15℃时 TCM 和 BDCM 的 SHAP 值为正,而小于 15℃时对应的 SHAP 值则为负.DBCM 的水温分界线则为 20℃,说明 DBCM 的反应活性受温度影响更大(图 6(g)).上述结果表明,较高水温条件下该指标贡献较高,是影响模型结果的关键因子,而低温条件下其贡献降低.值得注意的是,水温对 TBM 的 SHAP 值并不呈单纯的正向贡献或者负向贡献,当水温低于 20℃或高于 30℃时,水温对 TBM 的 SHAP 值的影响为负向,而当水温介于 20~30℃时(尤其是在 25℃),水温对其 SHAP 为正向贡献(图 6(j)).由此可以推断,相比于水温,TBM 的生成速率受溴离子浓度影响更为显著.余氯与 4 种 THMs 的 SHAP 值呈显著负相贡献,当余氯浓度越高时,其对模型结果的负向贡献越大,反之则越小.根据图 6(b),当余氯浓度高于 0.45mg/L 时,余氯对 TCM 的 SHAP 值产生负向贡献,即余氯浓度降低, TCM 浓度快速升高.当余氯浓度小于 0.45mg/L 时,余氯的 SHAP 值为正,这证明具有活性前体物已基本参与反应,TCM 的生成活性受其他环境因子(尤其是水温)影响更多.根据图 6(e)和(h),对于 BDCM 和 DBCM,余氯的界限为 0.55mg/L.当余氯浓度小于 0.55mg/L 时,水温和供水距离成为影响两种 THMs 生成活性的关键因子.供水距离是影响 TCM 和 BDCM 的第三关键因子,余氯浓度随着供水距离的增加降低,此阶段(0~6km)THMs 生成速率最快,余氯消耗也最快.根据图 6(c)和(f),TCM 和 BDCM 对应的 SHAP 值在 0~2km 变化不明显,而在 2~10kmSHAP 值随着供水距离增加显著上升,并以 6km 作为分界线,6km 后供水距离对模型结果带来正向影响,且影响作用更大.基于 SHAP 的局部解析证实了余氯小于 0.45~0.55mg/L 时,余氯对 THMs 生成浓度的影响较低,这主要是由于供水管网中具有活性的 THMs 前体物已经与余氯发生反应,剩余 DOM 为活性较低的反应底物.根据图 6(i),UV_{254} 是仅次于水温和余氯影响 DBCM 生成的关键因子,当 UV_{254} 高于 0.024cm^{-1} 时,对 DBCM 的 SHAP 值产生正向影响,而低于 0.024cm^{-1} 时为负向影响,这说明 UV_{254} 的浓度限制(0.024cm^{-1})是能够评估 DBCM 生成浓度的关键指标之一.

2.5 展望与应用前景

以往研究建立的 THMs 生成动力学模型主要基于实验室数据,在各项因素可控的前提下建立的基于反应时间和生成浓度的 THMs 动力学模型.但实际饮用水供水过程中难以准确计算水龄,推断反应时间的长短,因此难以在实际工程应用.在本研究中,基于可解释机器学习算法的应用为供水管网多因子预测模型的建立提供技术支持.结合以往文献研究结果[24],通过对拟合性能最优 GBDT 模型的 SHAP 分析对模型结果做出解释,梳理了对模型结果起到正面或者负面影响的关键因子.如图 7 所示,将供水距离,水温,余氯,浊度,pH 值,UV_{254} 和 DOC 中的关键因子作为输入因子可实现对供水管网不同位置 TCM,BDCM,DBCM 和 TBM 的高效预测.模型识别到水温,余氯和供水距离是影响供水管网中 THMs 生成的关键参数,夏季在管网末端属于 THMs 的高风险区域.根据 SHAP 解析可知,保证供水管网中余氯浓度小于 0.45,0.55 和 0.55mg/L 时可显著降低余氯对 TCM,BDCM 和 DBCM 的影响比重;当 UV_{254} 低于 0.024cm^{-1} 时,可显著降低 DBCM 的生成浓度.模型识别高温季节是造成 THMs 生成浓度过高的关键因子.针对本研究的目标区域,6~8 月被识别为该区域的水质高风险期,8~10km 为 THMs 生成浓度最高的区域范围.在此期间应加强水处理工艺的运维和采样检测工作,将采集到的数据输入模型可快速预测 THMs 的生成,根据预测结果灵活调整供水压力和加氯量等工况参数,保证供水管网水质安全.此外,本研究建立的 THMs 随供水距离的生成模型则有效避免了水龄计算,通过建立基于空间位置和复合环境因子的预测模型,可有效预测供水系统不同位置 THMs 的浓度,进而为水质分析提供便利.毫无疑问,这项工作能够减少实际供水管网采样工作,预警 THMs 生成风险.此外,根据 SHAP 的局部解析证明 UV_{254} 和 Br 可作为控制 DBCM 和 TBM 的导向指标,这为水厂精准加氯提供参考依据.

当然,本研究也存在如下不足和缺陷:(1)模型建立所收集的数据量较少,模型准确度和泛化能力有待提升.(2)模型预测边界有限,无法准确预测供水距离大于 10km 处的 THMs.(3)模型数据基于苏州工业园区某水厂供水系统,其模型迁移应用效能有待验证.然而,ML 的另一个重要优势是自适应迭代学习,通过积累水厂每年获取的新数据并将其输入模型,ML 算法可以不断迭代和学习,从而提高 ML 模型的泛化能

力和准确性.在后续研究,我们将致力于提高模型的预测精度,扩大模型的边界范围和应用条件.

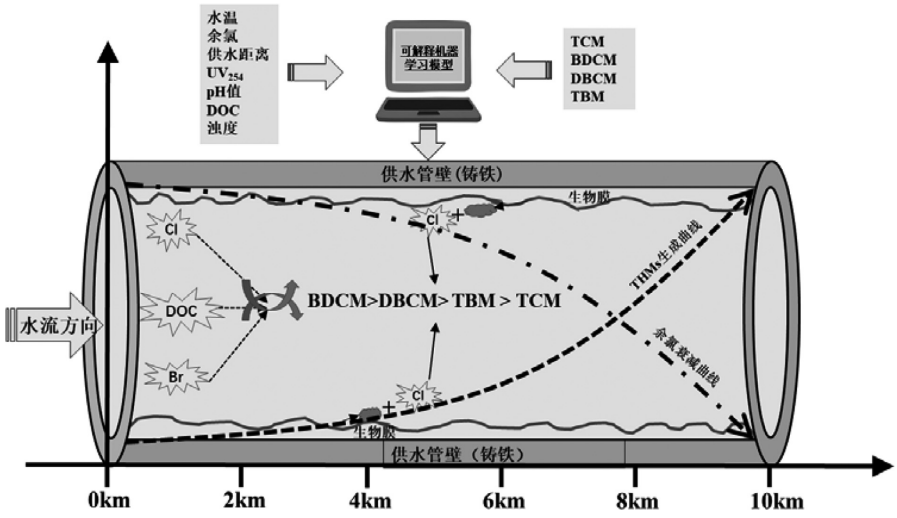


图 7　基于可解释机器学习算法建立供水管网 THMs 生成模型及关键因子识别

Fig.7　Generation model of THMs in water supply network based on interpretable machine learning algorithm and identification of key factors

3　结论

3.1　空间尺度分析表明,供水管网末端 8~10km 处的 THMs 浓度处于高风险水平;时间尺度分析表明,供水管网末端 THMs 的浓度在 6 月达到最高峰,浓度为 39.06μg/L.

3.2　机器学习可辅助构建多个环境因子与 THMs 浓度的数据驱动模型,其中梯度提升树(GBDT)算法能够有效避免模型过拟合,训练集和测试集的 R^2, MSE 和 MAE 结果更接近,并对 4 种 TCM, BDCM, DBCM 和 TBM 的预测精度 R^2 分别达到 0.839,0.906, 0.836 和 0.935.

3.3　Shapley 附加解释(SHAP)对 GBDT 模型的全局解析发现,水温,供水距离,UV_{254} 和 DOC 与 THMs 生成呈正相关关系,余氯与 THMs 呈显著负相关关系,此结果与先验知识一致,SHAP 能够有效解决传统 ML 的“黑箱”问题.

3.4　SHAP 的局部解析表明,供水管网中余氯浓度小于 0.45,0.55 和 0.55mg/L 时可显著降低余氯对 TCM,BDCM 和 DBCM 的 SHAP 值,因此水厂可基于管网余氯监测结果,反馈控制出厂水余氯投加量.此外,UV_{254} 的浓度限制($0.024cm^{-1}$)是能够评估 DBCM 生成浓度的关键指标之一,在进水 UV_{254} 较高时,水厂应严格控制该指标结果,进而减少 DBCM 生成浓度.

参考文献：

[1] 张金松,卢小艳.饮用水消毒工艺及副产物控制技术发展 [J]. 给水排水, 2016,52(9):1−3.
Zhang J S, Lu X Y. Drinking water disinfection process and by−product control technology development [J]. Water Supply and Drainage, 2016,52(9):1−3.

[2] 陈　超,王　玉,谢宇煊,等.饮用水水质标准升级带来的消毒副产物挑战与对策 [J]. 给水排水, 2024,60(9):7−14.
Chen C, Wang Y, Xie Y X, et al. Challenges and countermeasures of disinfection by−products brought by upgrading drinking water quality standards [J]. Water Supply and Drainage, 2024,60(9):7−14.

[3] 李惠平.纳滤膜在高品质饮用水处理中的应用研究 [D]. 兰州:兰州交通大学, 2020.
Li H P. Research on the application of nanofiltration membrane in high quality drinking water treatment [D]. Lanzhou: Lanzhou Jiaotong University, 2020.

[4] 任洋洋,钱海雷,裴赛峰,等.应用中断时间序列方法评估干预措施对陆家嘴水厂出厂水三卤甲烷的影响 [J]. 环境与职业医学, 2024,41(4):420−424.
Ren Y Y, Qian H L, Pei S F, et al. Application of an interrupted time series approach to assess the effects of interventions on trihalomethanes in Lujiazui water plant effluent [J]. Environmental and Occupational Medicine, 2024,41(4):420−424.

[5] 王　荀,周玉潇,韦慧慰,等.深圳市龙岗区生活饮用水中三卤甲烷健康风险评估 [J]. 华南预防医学, 2024,50(7):594−598.
Wang X, Zhou Y X, Wei H W, et al. Health risk assessment of trihalomethanes in drinking water in Longgang District, Shenzhen [J]. South China Preventive Medicine, 2024,50(7):594−598.

[6] Brown D, Bridgeman J, West J R.Predicting chlorine decay and THM formation in water supply systems [J]. Reviews in Environmental Science & Bio/technology, 2011,10(1):79−99.

[7] Huang R, Ma C, Ma J,et al.Machine learning in natural and engineered water systems [J]. Water Research, 2021,(21):117666.

[8] Bowden G J, Nixon J B, Dandy G C, et al. Forecasting chlorine residuals in a water distribution system using a general regression neural network [J]. Mathematical & Computer Modelling, 2003,44 (5/6):469−484.

[9] Li R A, Mcdonald J A, Sathasivan A, et al. A multivariate Bayesian network analysis of water quality factors influencing trihalomethanes

formation in drinking water distribution systems [J]. Water Research, 2021,(190):116712.
[10] Bowden G J, Nixon J B, Dandy G C, et al. Forecasting chlorine residuals in a water distribution system using a general regression neural network [J]. Mathematical & Computer Modelling, 2003,44 (5/6):469−484.
[11] Mahya P, Johannes Fürnkranz. An empirical comparison of interpretable models to post−hoc explanations [J]. AI, 2023,4(2):426−436.
[12] Li L B, Qiao J D, Yu G, et al. Interpretable tree−based ensemble model for predicting beach water quality, Water Research, 2022,(211):118078.
[13] Yuan X Z, Suvarna M, Low S, et al. Applied machine learning for prediction of CO_2 Adsorption on biomass waste−derived porous carbons, Environmental Science&Technology, 2021,(55):11925–11936.
[14] 于林堂,陈咚咚,陶翠翠,等.基于可解释性机器学习的芬顿工艺降解有机污染物速率的研究 [J]. 中国环境科学, 2025,45(8):4294-4302.
Yu L T, Chen D D, Tao C C, et al. Research on the degradation rate of organic pollutants in fenton process based on explainable machine learning [J]. China Environmental Science, 2025,45(8):4294-4302.
[15] 张 煦,黄 霞,张晓彤,等.基于可解释机器学习方法的水库总磷预测研究 [J]. 水生态学杂志, 2025-05-23:1−14.
Zhang X, Huang X, Zhang X T, et al. Prediction of total phosphorus in reservoir based on interpretable machine learning methods [J]. Journal of Hydroecology, 2025-05-23:1−14.
[16] GB5750-2023 生活饮用水标准检验方法 [S].
GB5750-2023 Standard examination methods for drinking water [S].
[17] Li H P, Zhou B Q, Xu X Y, et al. The insightful water quality analysis and predictive model establishment via machine learning in dual−source drinking water distribution system [J]. Environmental Research, 2024,(250):118474.
[18] 陈雨桐.集成学习算法之随机森林与梯度提升决策树的分析比较 [J]. 电脑知识与技术, 2021,17(15):32−34.
Chen Y T. Analyzing and comparing random forest and gradient boosting decision tree of integrated learning algorithms [J]. Computer Knowledge and Technology, 2021,17(15):32−34.
[19] 周杰英,贺鹏飞,邱荣发,等.融合随机森林和梯度提升树的入侵检测研究 [J]. 软件学报, 2021,32(10):3254−3265.
Zhou J Y, He P F, Qiu R F, et al. Research on intrusion detection by Fusing Random Forest and Gradient Boosting Tree [J]. Journal of Software, 2021,32(10):3254−3265.
[20] Baptista M L, Goebel K, Henriques E M P. Relation between prognostics predictor evaluation metrics and local interpretability SHAP values [J]. Artificial Intelligence, 2022,(306):103667.
[21] DB32/T 4786−2024 城镇供水服务质量标准 [S].
DB32/T 4786−2024 Quality standards for urban water supply services [S].
[22] 徐 鹏,李忠群,程战利,等.水源水消毒副产物生成势与 UV_{254}/DOC 的相关性研究 [J]. 环境科学学报, 2018,38(8):3021−3026.
Xu P, Li Z Q, Cheng Z L, et al. Correlation study between the generation potential of disinfection by−products and UV254/DOC in water source water [J]. Journal of Environmental Science, 2018,38(8):3021−3026.
[23] Lin H, Dai Q, Zheng L, et al. Radial basis function artificial neural network able to accurately predict disinfection by−product levels in tap water: Taking haloacetic acids as a case study [J]. Chemosphere, 2020,248:125999.
[24] 张心悦,周克梅,刘 煜,等.供水管材对管壁生物膜特性和水质稳定性的影响 [J]. 中国给水排水, 2022,38(9):44−51.
Zhang X Y, Zhou K M, Liu Y, et al. Effect of Water Supply Pipe Material on Biofilm Characteristics and Water Quality Stability [J]. China Water & Wastewater, 2022,38(9):44−51.
[25] Li H, Huo R, Xu X, et al. Applied gradient boosting decision tree algorithms for accurate prediction of trihalomethanes: A case study in dual−sources drinking water distribution system in metropolitan [J]. Journal of Water Process Engineering, 2023,(56):104416.
[26] GB 5749−2022 生活饮用水卫生标准 [S].
GB 5749−2022 Hygienic standard for drinking water [S].
[27] 韦运明,吕 竹,黄昭铭.城市供水管网余氯影响因素研究 [J]. 供水技术, 2022,16(6):10−14.
Wei Y M, Lv Z, Huang Z M. Study on the influencing factors of residual chlorine in urban water supply network [J]. Water Supply Technology, 2022,16(6):10−14.
[28] 季京宣.H 市双水厂管网供水分界带定位及水质改善研究 [D]. 杭州:浙江大学, 2022.
Ji J X, H. Research on the localization of water supply demarcation zone and water quality improvement of dual waterworks pipe network in H city [D]. Hangzhou: Zhejiang University, 2022.
[29] 卢小艳,蔡广强,刘丽君,等.南方某河流原水中消毒副产物前体物的季节性变化 [J]. 中国给水排水, 2015,31(11):51−55.
Lu X Y, Cai G Q, Liu L J, et al. Seasonal variation of disinfection by−product precursors in raw water of a southern river [J]. China Water Supply and Drainage, 2015,31(11):51−55.
[30] 邹 琳,周圣东,顾新春,等.加氯对不同工艺出水消毒副产物生成的研究 [J]. 中国给水排水, 2019,35(5):1−7.
Zou L, Zhou S D, Gu X C, et al. Study of chlorination on the generation of disinfection by−products in different process effluents [J]. China Water Supply and Drainage, 2019,35(5):1−7.
[31] Mohamed M, Mounir F, Ahmed E A. Machine learning methods for predicting argania spinosa crop yield and leaf area index: A combined drought index approach from multisource remote sensing data [J]. AgriEngineering, 2024,6(3):2283−2305.
[32] 郑 震,冷东梅,郑 莺,等.山仔水库环境因子与藻类群落生长的相互作用机制探究及应对建议 [J]. 灌溉排水学报, 2023,42(11):131−139.
Zheng Z, Leng D M, Zheng Y, et al. Exploration of the interaction mechanism between environmental factors and algal community growth in Shanzai Reservoir and suggestions for response [J]. Journal of Irrigation and Drainage, 2023,42(11):131−139.
[33] 成 佩,孟 勇.集成学习方法的应用与比较 [J]. 统计与决策, 2023, 39(23):28−33.
Cheng P, Meng Y. Application and comparison of integrated learning methods [J]. Statistics and Decision Making, 2023,39(23):28−33.
[34] 郭衍昊,窦 杰,向子林,等.基于优化负样本采样策略的梯度提升决策树与随机森林的汶川同震滑坡易发性评价 [J]. 地质科技通报, 2024,43(3):251−265.
Guo Y H, Du J, Xiang Z L, et al. Evaluation of Wenchuan co−seismic landslide susceptibility based on gradient boosting decision tree and random forest with optimized negative sample sampling strategy [J]. Geoscience and Technology Bulletin, 2024,43(3):251−265.
[35] 王晋宇,陈玲瑚,赵 辰,等.饮用水中三卤甲烷生成影响因素的初步研究 [J]. 净水技术, 2009,28(6):30−34.
Wang J Y, Chen L H, Zhao C, et al. Primary investigation on influencing factors of trihalomethanes generation in drinking water, water purification technology, 2009,28(6):30–34.
[36] 王 玲,幸薇洁,张翔飞,等.成都市龙泉驿区生活饮用水中消毒副产物含量及其影响因素分析 [J]. 预防医学情报杂志, 2025,41(6):813−818.
Wang L, Xing W J, Zhang X F, et al. Analysis of disinfection by-products in drinking water and their influencing factors in Longquanyi District, Chengdu [J]. Journal of Preventive Medicine Intelligence, 2025,41(6):813−818.
[37] Dong F L, Pang Z, Yu J Q, et al. Spatio-temporal variability of halogenated disinfection by-products in a large-scale two-source water distribution system with enhanced chlorination [J]. Journal of Hazardous Materials, 2022,423(Feb.5 Pt.A):127113.1−127113.11.
[38] Fang C, Yang W Y, Lu N N, et al. Alkaline chlorination of drinking water: A trade-off between genotoxicity control and trihalomethane formation [J]. Water Research: A Journal of the International Water Association, 2023,246(Nov.1):1.1−1.10.

作者简介：徐小燕(1979−),女,江苏如东人,工程师,学士,主要研究方向为水质检测及水处理技术.xuxiaoyan@heisino.com.