

黎相荣, 钟丽霞, 孙娇, 等. 基于 Patch-Transformer 模型的供水管网余氯衰减智能预测的应用[J]. 净水技术, 2025, 44(10): 64–72, 143.

LI X R, ZHONG L X, SUN J, et al. Application of water supply network in residual chlorine decay intelligent prediction based on Patch-Transformer model[J]. Water Purification Technology, 2025, 44(10): 64–72, 143.

基于 Patch-Transformer 模型的供水管网余氯衰减智能预测的应用

黎相荣*, 钟丽霞, 孙娇, 李昆朋

(重庆市渝南自来水有限公司, 重庆 401320)

摘要 【目的】供水管网末梢的余氯浓度是保障饮用水水质安全的关键性指标,但其动态衰减规律受到水龄、管道材质、管壁生物膜附着状态等多种复杂因素的综合影响,导致基于物理化学反应机理的传统预测模型在实际应用中存在精度不足的局限性。为有效提升对管网系统末梢余氯浓度变化的预测准确性,从而为水厂优化加氯调控策略提供科学依据,进一步强化供水安全保障能力,文章提出基于 Transformer 智能算法的余氯衰减预测新方法。【方法】针对余氯衰减序列数据所具有的高度复杂性与时空耦合特性,文章引入 Transformer 深度学习算法,该算法凭借其特有的自注意力机制与强大的长序列特征捕捉能力,能够深入挖掘余氯监测数据中蕴含的复杂非线性依赖关系。文章以重庆市巴南区某典型供水管网片区为研究对象,系统收集了相关监测数据。为进一步提升基础 Transformer 模型的预测性能,文章采用 Patch 模块对模型结构进行优化,同步开展参数调优,最终成功建立了 Patch-Transformer 余氯衰减智能预测模型。【结果】最终所构建的 Patch-Transformer 余氯衰减智能模型,能够实现对管网末梢余氯浓度提前 12 h 的有效预测。模型水质预测误差仅为 5.34%,具有较高的精度。【结论】本文成功构建了基于 Patch-Transformer 的供水管网余氯衰减智能预测模型,可辅助水厂优化加氯策略,实现精准投加,有助于提升水厂的智慧化管理水平和安全保障能力。

关键词 供水管网 余氯衰减 Transformer 智能算法 深度学习 水质预测

中图分类号: TU991 文献标志码: A 文章编号: 1009-0177(2025)10-0064-10

DOI: 10.15890/j.cnki.jsjs.2025.10.007

Application of Water Supply Network in Residual Chlorine Decay Intelligent Prediction Based on Patch-Transformer Model

LI Xiangrong*, ZHONG Lixia, SUN Jiao, LI Kunpeng

(Chongqing Yunan Tap Water Co., Ltd., Chongqing 401320, China)

Abstract [Objective] Residual chlorine concentration at the end of water supply network is a critical indicator for ensuring drinking water safety. However, its dynamic decay behavior is influenced by multiple complex factors including water age, pipe material, and biofilm attachment on pipe walls, resulting in limited accuracy of traditional prediction models based on physicochemical reaction mechanisms. To enhance prediction accuracy for terminal residual chlorine concentration variations, thereby providing a scientific basis for optimizing chlorination strategies and strengthening water supply security, this paper proposes a Transformer-based intelligent approach for residual chlorine decay prediction. [Methods] To address the high complexity and spatiotemporal coupling characteristics of residual chlorine decay sequences, this paper introduced the Transformer deep learning algorithm. Based on its inherent self-attention mechanism and robust long-sequence feature extraction capabilities, this architecture effectively captures complex nonlinear dependencies within monitoring data. This paper systematically collected monitoring data from a representative water distribution network in Banan District, Chongqing. To further enhance the predictive performance of the base Transformer framework, a Patch module was incorporated for structural optimization alongside comprehensive parameter tuning, ultimately establishing the Patch-Transformer intelligent prediction model. [Results] The developed Patch-Transformer model achieved effective 12-hour-ahead

[收稿日期] 2024-07-25

[通信作者] 黎相荣(1972—),男,研究方向为供水管网余氯衰减机制, E-mail: 652520891@qq.com。

predictions of residual chlorine concentration at pipeline termini, demonstrating high precision with a prediction error of only 5.34%.

[**Conclusion**] This paper successfully establishes an intelligent Patch-Transformer-based prediction model for residual chlorine decay in water supply networks. The model supports water treatment plants in optimizing chlorination strategies and enabling precise dosing, thereby advancing smart management capabilities and safety assurance in water treatment operations.

Keywords water supply network chlorine residual decay transformer intelligent algorithm deep learning water quality prediction

供水管网中的余氯是非常重要的水质指标,若供水管网中余氯不足,会导致致病菌在饮用水中大量繁殖,而余氯过多则会导致产生过量的消毒副产物,消毒副产物的毒性已有许多研究^[1],包括细胞毒性、遗传毒性、致癌性、致畸性等。对于供水管网中的余氯,目前我国大部分的供水厂依赖日检、月检、年检等现场采样或使用在线监测仪器进行监测,但若在监测到余氯异常后再调整加氯量,从清水池到管网末梢输送调整加氯量后的水,仍存在滞后性,因此有必要对管网余氯进行预测。

目前,对供水管网的余氯衰减预测的研究主要可以分为机理模型^[2]及机器学习模型^[3]。其中,机理模型在管网末梢易受水力条件变动的影响而导致精度下降^[4],而目前余氯衰减预测的研究中所使用的机器学习模型,包括唐宇等^[5]使用改进的径向基函数神经网络(RBFNN)模型优化水厂加氯量,Bowden等^[6]使用广义回归神经网络(GRNN)进行余氯衰减预测,以及Wakdar等^[7]使用前馈神经网络(FFNN)模型对余氯在内的27个指标进行预测等,采用的均是比较经典的机器学习模型,尚未应用计算机领域内新的深度学习技术。

近年来,随着计算机算力的提升,深度学习模型在水处理的许多领域均有应用,由于生产实际中有许多数据均随时间的变化而受到季节、周期性和不规则变动等因素的影响,呈现出稳定的变化规律,许多研究中都应用了深度学习时间序列预测模型。如Shim等^[8]与Chen等^[9]基于长短期记忆(LSTM)模型进行污水处理厂进出水水质的预测,Jiang等^[10]使用门控循环单元(GRU)模型进行工业废水中重金属的预测。

Transformer智能算法是由Vaswani等^[11]在2017年首先提出的一种深度学习算法,在自然语言处理领域已经取得了广泛的研究与应用,如语音辨识、机器翻译、语义提取等。相较于传统人工神经网络模型需按时间步顺序计算处理数据,Transformer通过矩阵并行运算一次性处理整个序列,显著地提

升训练和推理速度,因此近年来,也有学者将Transformer智能算法应用于时间序列预测领域,并取得了显著的成果。如Huang等^[12]首先将Transformer智能算法应用于污水厂出水水质的预测,在对COD、氨氮、TN 3项指标的预测中,最低的准确率也能达到96.8%,超过用于对比的多层感知机(MLP)、卷积神经网络(CNN)及LSTM模型;而刘雅亭等^[13]基于Transformer模型对某独立计量区域(DMA)分区进行供水量预测,结果表明,相较于LSTM模型和支持向量回归(SVR)模型,Transformer模型预测精准度更高稳定性也更好。Wang等^[14]则使用萨维茨基-戈莱(SG)滤波器与Transformer智能算法相结合的多步骤流域水质预测模型,并使用黄河的4个水质监测站点数据检验其预测精度,对比LSTM模型,SG-Transformer在平均绝对误差(MAE)、均方根误差(RMSE)及纳什-萨特克利夫效率系数(NSE)3项性能指标中均取得了更佳的表现。供水管网中的余氯也有着明显的时间序列特征。目前,大多数余氯衰减的机器学习模型也都使用一些经典的人工神经网络模型及其变体,未来随着管网余氯在线监测设备不断增加,可以尝试使用更复杂的模型,以取得更好的预测效果。因此,本研究将基于Transformer模型进行供水管网余氯衰减智能模型的构建与优化。

1 研究方法

1.1 Transformer 智能算法

要使用Transformer智能算法进行模型的构建,首先需要理解Transformer智能算法的结构,如图1所示。

Transformer智能算法主要由 n 层的编码器和 n 层的解码器2部分组成。Transformer智能算法的核心是自注意力机制,它能够学习序列中不同位置之间的关联性,并在预测时利用这些关联性来捕捉时间序列数据的长期依赖关系。具体来说,自注意力机制通过注意力权重矩阵 W_Q 、 W_K 及 W_V 来确定每个输入位置对于当前预测的重要性。这些权重用于

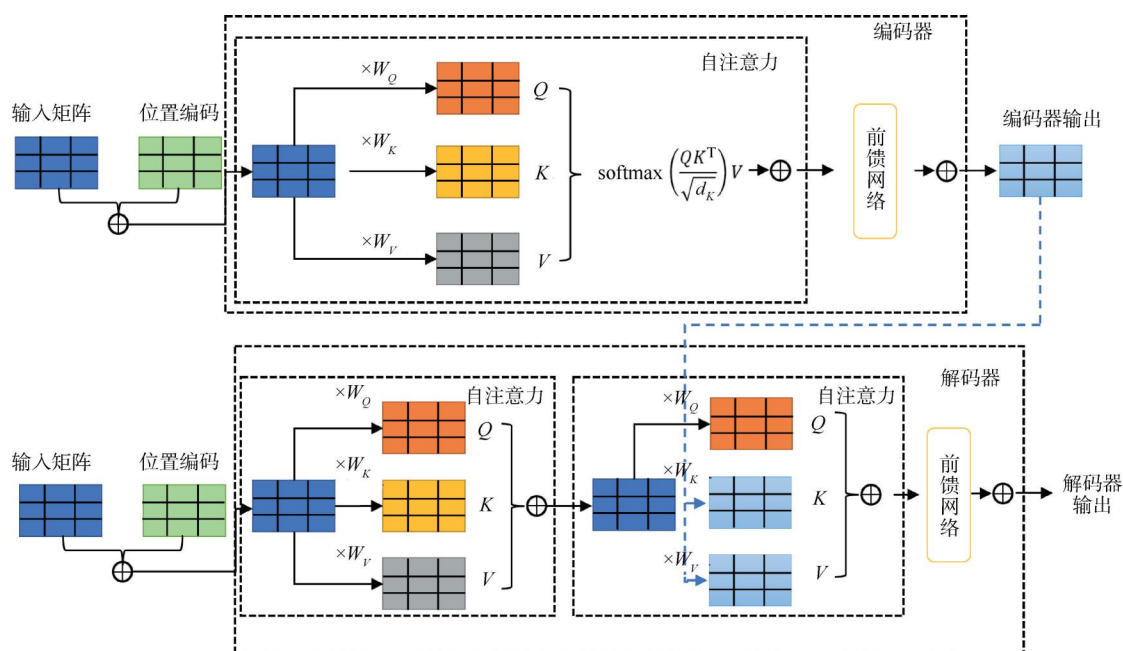


图 1 Transformer 智能算法结构

Fig. 1 Structure of Transformer Intelligent Algorithm

加权求和输入序列的编码表示,以产生模型的输出。查询向量 (Query, Q), 键向量 (Key, K) 和值向量 (Value, V) 3 个矩阵均来自同一输入, 并按照以下步骤计算。首先, 计算 Q 与 K 之间的点积, 为了防止其结果过大, 会除以 $\sqrt{d_k}$ (d_k 为 K 向量的维度), 这是为了使整个训练过程的梯度保持稳定, 避免出现梯度消失的问题, 然后利用 Softmax 操作将其归一化为概率分布, 再乘以矩阵 V 就得到权重求和的表示。整个计算过程如式(1)。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \quad (1)$$

在编码器中, 输入矩阵在自注意力机制计算后经过一层前馈网络进行输出, 解码器中有 2 次自注意力计算, 其中第 2 次计算时的矩阵 K 与 V 来自编码器的输出。最终经过一层前馈网络, 进行整个模型的输出。

总的来说, 使用 Transformer 智能算法进行时间序列预测的原理是通过自注意力机制捕捉序列中的长期依赖关系, 并在预测时考虑不同位置的相关性。通过合理设计模型结构、选择适当的损失函数和优化算法以及处理时间序列数据的特定问题, 可以实现准确的时间序列预测。

1.2 Transformer 模型的训练

将完成预处理的数据集划分为训练集、验证集与测试集, 进行模型的训练与测试。以 Transformer 模型为例, 模型训练的模式如图 2 所示, 需要使用损失函数计算该轮训练的损失, 简单来说, 损失函数就是用于评估每一轮训练后, 该模型预测的值与实际值之间的差距, 即模型在该轮训练后的性能, 并将损失输入到优化器中, 对模型的权重进行调整, 常用的损失函数包括均方误差 (MSE)、MAE 等。一般而言, 其余深度学习模型也具有类似的训练方式。即在训练集上进行模型权重的学习, 有验证集时同步在验证集上进行验证, 当模型训练结果达到最优后, 停止训练并在测试集上进行测试。

1.3 模型评价指标

本研究中, 使用以下 3 个指标作为模型性能的评价指标。

(1) MSE

MSE 值计算如式(2)。

$$M_{\text{MSE}} = (y_i - \hat{y}_i)^2 \quad (2)$$

其中: M_{MSE} —— MSE 值, $(\text{mg/L})^2$;

y_i —— 真实值, mg/L ;

$\hat{y}_i$ —— 预测值, mg/L 。

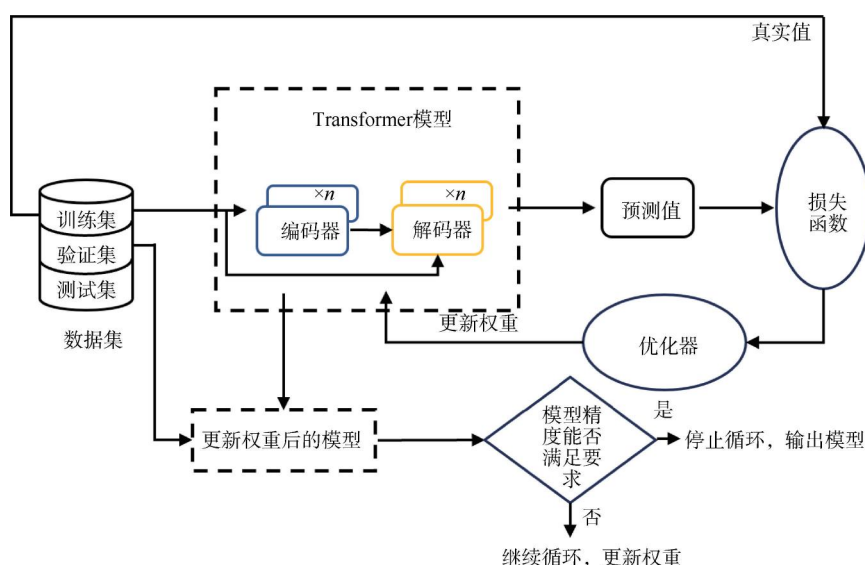


图2 Transformer模型的训练

Fig. 2 Training of Transformer Model

MSE 的优点在于计算简单,同时对于大误差会给予更大的惩罚,便于模型减少大的预测误差。缺点是模型会更加关注异常值而导致过拟合的情况出现。

(2) MAE

MAE 值计算如式(3)。

$$M_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

其中: M_{MAE} ——MAE 值,mg/L。

MAE 的优点在于不会过度关注异常值而导致模型过拟合,但 MAE 与 MSE、均方根误差不同,其在零点处不可导,可能导致优化器反向传播导数时在零点附近收敛困难。

(3) 水质预测误差(δ)^[15]

δ 值计算如式(4)。

$$\delta = \frac{|\hat{x} - x|}{\lambda} \times 100\% \quad (4)$$

其中: δ ——水质预测误差值;

$\hat{x}$ ——游离氯预测值,mg/L;

x ——游离氯实测值,mg/L;

λ ——出厂水游离氯均值,mg/L。

2 实例研究

2.1 Transformer 模型的构建

本研究的目标供水片区管网,位于重庆市巴南

区。该片区由单一供水厂进行供水,设计供水量为 40 万 m^3/d ,供水范围如图 3 所示,供水面积约为 57 km^2 。供区西方属于江河畔,地势较低,东方属山脉地形,地势较高,整体地面标高为 188~310 m,地势高差较大,供区共有 3 座加压泵站。

本研究所收集的数据集为目标供水片区某供水厂的在线监测设备所记录的数据,设备型号为瑞明 RM-SG 型 10 水质多参数在线监测仪,收集数据的时间为 2023 年 9 月 27 日 13:00:00—2024 年 3 月 1



图3 供水范围

Fig. 3 Range of Water Supply

日 00:00:00, 为确保在线监测设备精度, 收集数据期间每 2 周对各监测点的仪器进行校正, 收集的数据包括出厂水余氯 (C_0)、出厂水 pH (pH 值)、出厂水浑浊度 (T_{turb}) 以及位于管网末梢的管网控制点余氯 (C_3), 每 10 min 记录 1 次, 总共 22 388 行, 部分数据集如表 1 所示。

检查该数据集的完整性, 发现部分时刻的数据存在缺失, 主要是由于仪器异常或离线, 使用插值法将缺失值补齐。补齐后的数据集, 如图 4 所示。

表 1 数据集的收集
Tab. 1 Collection of Datasets

时间	$C_3/$ ($\text{mg} \cdot \text{L}^{-1}$)	$C_0/$ ($\text{mg} \cdot \text{L}^{-1}$)	pH 值	$T_{\text{turb}}/$ ($\text{mg} \cdot \text{L}^{-1}$)
2023-09-27 13:00:00	0.28	0.748	7.92	0.063
2023-09-27 13:10:00	0.28	0.745	7.93	0.067
...	...	...	...	...
2024-02-29 23:50:00	0.16	0.736	8.10	0.093
2024-03-01 00:00:00	0.16	0.730	8.09	0.092

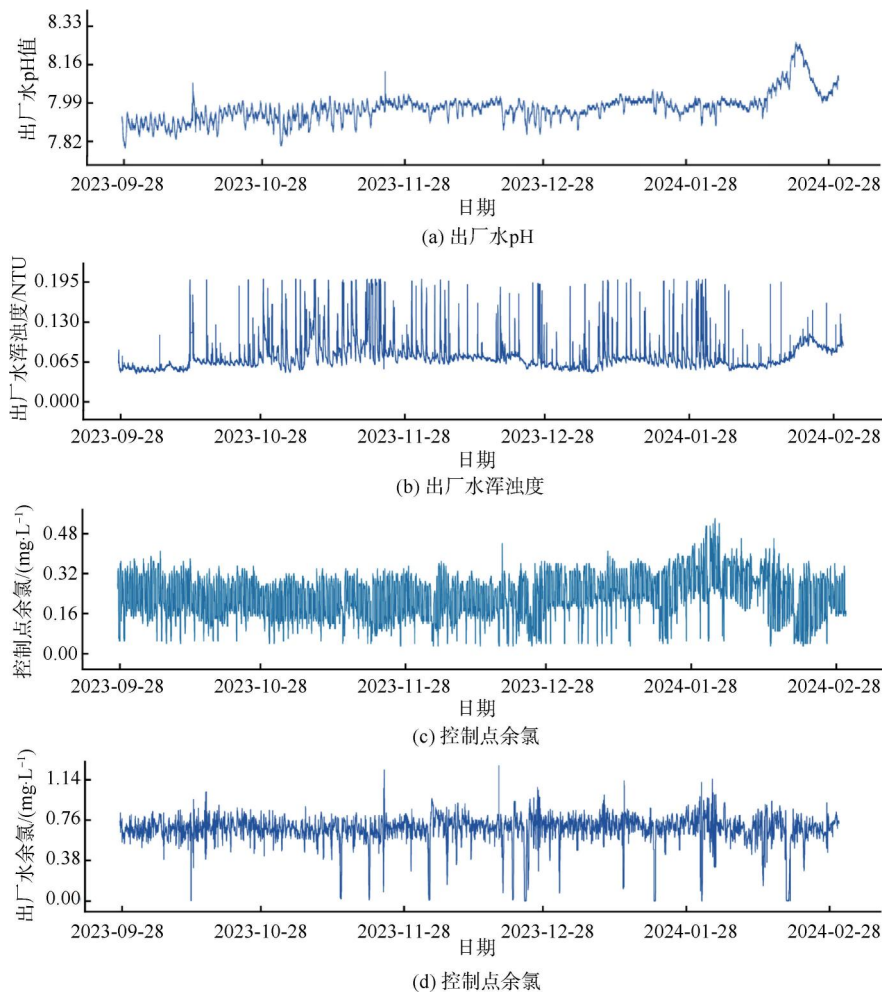


图 4 预处理后的数据集

Fig. 4 Preprocessed Dataset

将补齐后的 C_3 序列进行时间序列检验。使用 Python 进行时间序列的增强迪基-富勒 (ADF) 检验。ADF 检验是一种常用的单位根检验方法, 用于检验一个时间序列是否具有平稳性。平稳性是时间序列分析中的一个重要概念, 它指的是时间序列的统计特性不随时间变化。

ADF 检验结果包括检验统计量、临界值和 p 值等。如果检验统计量小于临界值, 且 p 值小于显著性水平 (通常可定为 0.05), 则可以拒绝原假设, 即数据不存在单位根, 认为该数据是平稳的。如果检验统计量大于临界值, 或 p 值大于显著性水平, 则无法拒绝原假设, 即数据存在单位根, 认为该数据为非

平稳的。

使用 Python 进行时间序列的 ADF 检验结果如表 2 所示。

表 2 ADF 检验结果
Tab. 2 Results of ADF Test

滞后阶数	10
p 值	0
ADF 检验统计量	-25.529
1%临界值	-3.431
5%临界值	-2.862
10%临界值	-2.567

由表 2 可知, p 值为 0, 且 ADF 检验统计量为 -25.529, 远小于 1% 临界值 -3.431, 即在 99% 的置信水平下可以认为该序列是平稳的, 具有不随时间变化的时间序列统计特性。

为构建余氯衰减智能模型, 本研究基于 Python3.8 的 PyTorch 库进行 Transformer 模型的构建。具体使用的 PyTorch 版本为 2.2.0, 所有模型训练次数均设置为 10 次, 以排除模型随机种子带来的偶然性。

首先, 按照 1.2 节所述的 Transformer 智能算法结构进行模型构建。Transformer 模型所使用的参数如表 3 所示。为了减少季节变化、周期波动和随机事件对数据的影响, 研究采用时间序列分割法对数据集进行统一划分。时间序列分割法将数据按照时间顺序分成训练集、验证集和测试集, 可以确保训练、验证和测试集之间的时间依赖性不会被破坏, 避免未来信息泄露, 同时将最新的数据保留为测试集, 可以评估模型应对最新趋势和不规则变动的能力。本研究统一将数据集划分为 60% 的训练集, 20% 的验证集与 20% 的测试集, 在数据输入模型与输出预测结果时, 执行归一化与反归一化操作, 相关模型参数设置如表 3 所示。

使用初步构建的 Transformer 模型进行特征消融试验, 首先将输入序列长度 (seq_len) 及预测序列长度 (pred_len) 均设置为 144, 即模型训练、验证、测试时, 使用前 24 h 的数据预测后 24 h 的数据, 验证不同特征组合下, 该模型的预测能力。其中, C_0 作为与预测目标管网 C_3 有直接时间相关性的变量, 不进行特征消融, 对出厂水 pH (pH 值) 与出厂水浑浊度 (T_{turb}) 2 个特征进行特征消融试验, 试验结果如表 4 及图 5 所示。

表 3 Transformer 模型参数
Tab. 3 Parameters of Transformer Model

模型参数	设置/选择	参数值	备注
Enc-layers	/	3	编码器层数
Dec-layers	/	3	解码器层数
Learning-rate	/	0.000 1	学习率
Dropout	/	0.1	训练时不更新权重的比例
Optimizer	Adam	/	优化器
Activation	gelu	/	激活函数
Batch-size	/	32	批量大小

表 4 不同特征组合的预测性能
Tab. 4 Prediction Performance with Different Feature Combinations

特征组合	MSE 均值	MAE 均值	δ 值
C_3-C_0	0.006 19	0.051 0	7.53%
$C_3-C_0\text{-pH}$	0.006 32	0.057 5	8.47%
$C_3-C_0-T_{\text{turb}}$	0.004 45	0.040 6	5.98%
$C_3-C_0\text{-pH-}T_{\text{turb}}$	0.004 82	0.045 8	6.75%

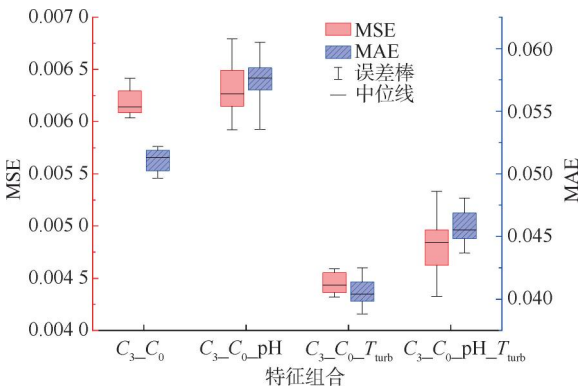


图 5 不同特征组合的模型预测性能
Fig. 5 Model Prediction Performance with Different Feature Combinations

如表 4 及图 5 所示, 当使用全部 4 个特征进行输入时, 初步构建的 Transformer 模型 δ 值为 6.75%。特征消融试验结果显示, pH 在本次模型构建中属于负面特征, 从全部特征的组合中将 pH 去除, δ 值从 6.75% 降至 5.98%, 预测性能比去除前提升 11.4%; 从 $C_3-C_0\text{-pH}$ 三特征组合中将 pH 去除, δ 值从 8.47% 降至 7.53%, 预测性能比去除前提升 11.1%。而 T_{turb} 则属于正面特征, 从全部特征的组合中将 T_{turb} 去除, δ 值从 6.75% 升至 8.47%, 预测性能比去除前下降 25.5%; 从 $C_3-C_0-T_{\text{turb}}$ 三特征组合中将 T_{turb} 去除, δ 值从 5.98% 升至 7.53%, 预测性能

比去除前下降 25.9%。

根据消融试验的结果,使用 C_0 、 C_3 及 T_{turb} 3 个特征进行管网控制点余氯的预测时,模型预测性能达到最佳水准。MSE 均值为 0.004 45, MAE 均值为 0.045 8, 同时水质预测误差达到 5.98%。

然后,使用 C_0 、 C_3 及 T_{turb} 三特征组合,探究模型最佳的序列长度及预测长度,试验结果如表 5 及图 6 所示。

表 5 不同序列长度、预测长度的预测性能
Tab. 5 Prediction Performance of Different Sequence Lengths and Prediction Lengths

序列长度	预测长度	MSE 均值	MAE 均值	δ 值
144	144	0.004 47	0.041 0	6.04%
144	72	0.004 31	0.040 3	5.95%
108	108	0.004 49	0.041 5	6.12%
108	72	0.004 45	0.041 5	6.12%
72	72	0.004 43	0.042 5	6.28%

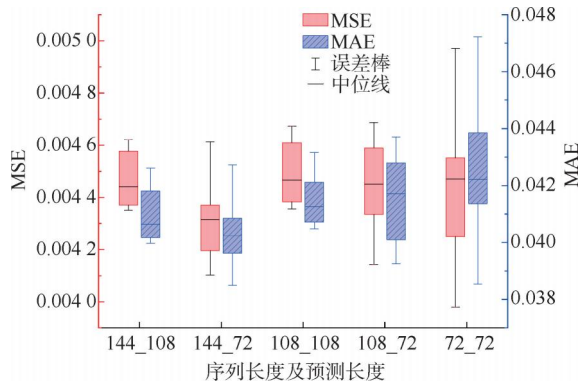


图 6 不同序列长度、预测长度的预测性能
Fig. 6 Prediction Performance of Different Sequence Lengths and Prediction Lengths

如表 5 及图 6 所示,模型测试结果中,当序列长度为 144 (即 24 h),预测长度为 72 (即 12 h) 时,初

步构建的 Transformer 余氯衰减智能模型精度达到最佳, MSE 均值为 0.004 31, MAE 均值为 0.040 3, δ 达到 5.95%。考虑管网实际的水龄及预测需求, 12 h 的预测长度已经能够满足出厂水余氯调整的需要。而当序列长度为 144 时,与更短的序列长度相比,模型的预测效果更加稳定。这是因为序列长度更长时,模型能够更好地捕捉序列中的特征,从而得到更稳定的预测结果,因此,后续试验中均确定以 C_3 、 C_0 、 T_{turb} 的三特征组合作为模型输入,同时将序列长度确定为 144,预测长度确定为 72。

2.2 Transformer 模型的优化

进一步探究 Transformer 余氯衰减智能模型的优化方式,使用 Nie 等^[16]提出的 PatchTST 模型进行尝试,该模型中包含的 Patch 模块是对 Transformer 模型的矩阵输入进行调整,如图 7 所示。

如图 7 所示, Patch 模块将原始的输入矩阵分割为 n 个 $1 \times P$ (时间窗宽度,反映水质变化的延时周期) 的矩阵,并分别进行输入,相比原先的 Transformer 模型直接将整个 M (特征维度数量,对应监测的水质参数种类数) $\times L$ (时间序列长度,代表历史监测时长) 的矩阵作为输入,减少了计算复杂度的同时,能够使模型更好地学习每个 Patch 内延时的时间依赖,这与本研究中数据集的特性相符,即出厂水余氯、浑浊度的变化需经过一段时间后才能够反应在管网控制点余氯的变化上。使用加入 Patch 模块的 Patch-Transformer 模型进行训练,能够使模型更好地捕捉出厂水余氯、浑浊度变化对管网控制点余氯变化的延时影响。

依然保持此前构建 Transformer 模型的最优输入参数,使用 C_0 、 T_{turb} 及 C_3 三特征组合进行输入,并设置 144 的序列长度及 72 的预测长度,其余模型参

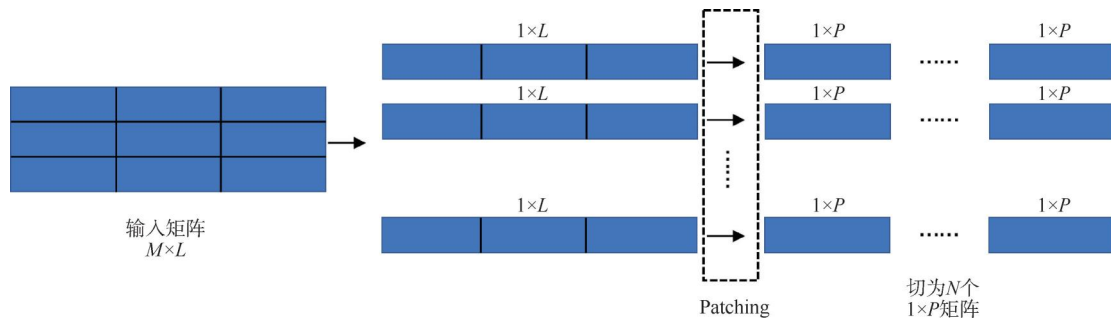


图 7 矩阵输入的调整
Fig. 7 Adjustment of Matrix Input

数保持与此前初步构建的 Transformer 模型一致,此外,对 Patch 模型的批量大小 (Batch_size) 进行调整,观察其性能变化。Batch_size 值决定了模型每次训练的样本数量,同时将影响 Patch-Transformer 模型对输入序列的分割情况。Batch_size 值一般可根据经验取 2 的幂,以便矩阵输入时在内存中对齐排列,将 Patch-Transformer 模型的 Batch_size 值分别取 32、64、128、256 进行试验,试验结果如表 6 及图 8 所示。

表 6 不同 Batch_Size 下 Patch-Transformer 模型的预测性能
Tab. 6 Prediction Performance of Patch-Transformer with Different Batch_Size

Batch_size	MSE 均值	MAE 均值	δ 值
32	0.004 26	0.036 7	5.41%
64	0.004 22	0.036 3	5.37%
128	0.004 18	0.036 2	5.34%
256	0.004 20	0.036 5	5.38%

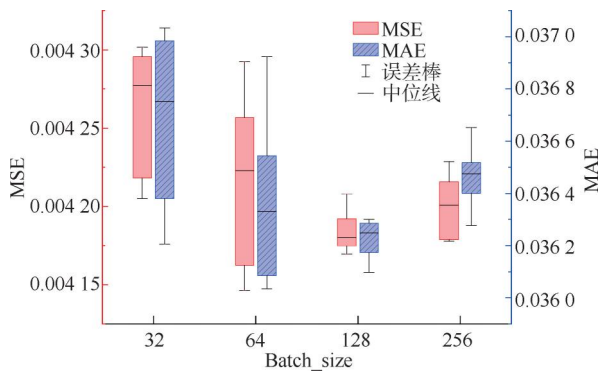


图 8 不同 Batch_Size 下 Patch-Transformer 模型的预测性能
Fig. 8 Prediction Performance of Patch-Transformer with Different Batch_Size

如表 6 及图 8 所示,当 Batch_size 取到 128 时, MSE、MAE 及 δ 值均达到最佳水平,其中水质预测误差 δ 能够达到 5.34%。另外,尝试使用增加 Dropout 值的方法缓解模型过拟合的情况, Batch_size 值使用预测性能最优的 128,其余参数均使用模型的最优参数,试验结果如表 7 及图 9 所示。

表 7 不同 Dropout 值下 Patch-Transformer 模型预测性能
Tab. 7 Prediction Performance of Patch-Transformer with Different Dropout Values

Dropout 值	MSE 均值	MAE 均值	δ 值
0.1	0.004 18	0.036 2	5.34%
0.3	0.004 13	0.036 5	5.38%
0.5	0.004 31	0.038 0	5.61%

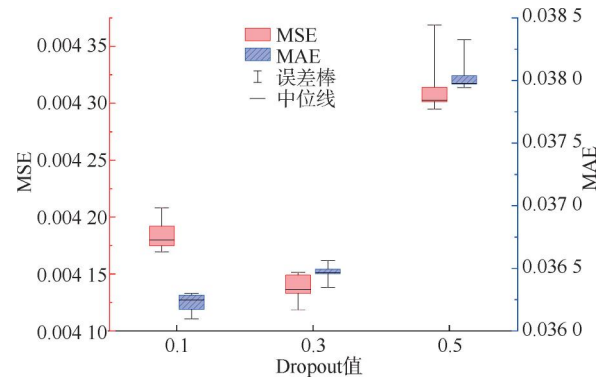


图 9 不同 Dropout 值下 Patch-Transformer 模型预测性能
Fig. 9 Prediction Performance of Patch-Transformer with Different Dropout Values

如表 7 及图 9 所示, Patch-Transformer 模型的预测性能随着 Dropout 值的增大而降低, δ 值由原先的 5.34% 升至 5.38% 及 5.61%, 对应预测性能分别降低了 0.75% 和 5.1%。结合前面在 DA-Transformer 模型中增加 Dropout 值, 模型预测性能也出现降低的现象, 说明模型没有因为过度学习训练集中的噪声而出现过拟合现象, 而是以目前的数据集长度及特征量, 模型已无法学习到更多的信息以达到更好的预测精度。

根据试验结果, 使用 Batch_size 设置为 128、Dropout 值设置为 0.1 的 Patch-Transformer 模型进行管网控制点余氯浓度预测, 效果最佳, δ 值为 5.34%, 使用 Python 将该模型在测试集上的预测结果进行可视化, 如图 10 所示, 模型对管网控制点余氯变化的整体趋势预测较为准确, 主要在一些数据不平滑的点预测效果较差, 需要在后续的数据收集关注这些不平滑点形成的原因。

为验证 Patch-Transformer 余氯衰减智能模型的优化效果, 使用供水片区原有的余氯衰减机理模型对管网控制点每日实际的余氯值进行预测, 预测时除未使用 Patch-Transformer 模型优化外, 其余条件均保持相同, 最终得到 δ 值为 9.13%, 优化后的 Patch-Transformer 模型预测性能相比余氯衰减机理模型提升了 41.51%, 能够不需人工校核调整参数的情况下, 更好地预测管网控制点的余氯浓度。

3 结论

为更好地预测管网末梢的余氯, 本文阐述了深度学习中 Transformer 模型的原理与训练过程, 并基于 Transformer 智能算法, 构建了余氯衰减智能模型

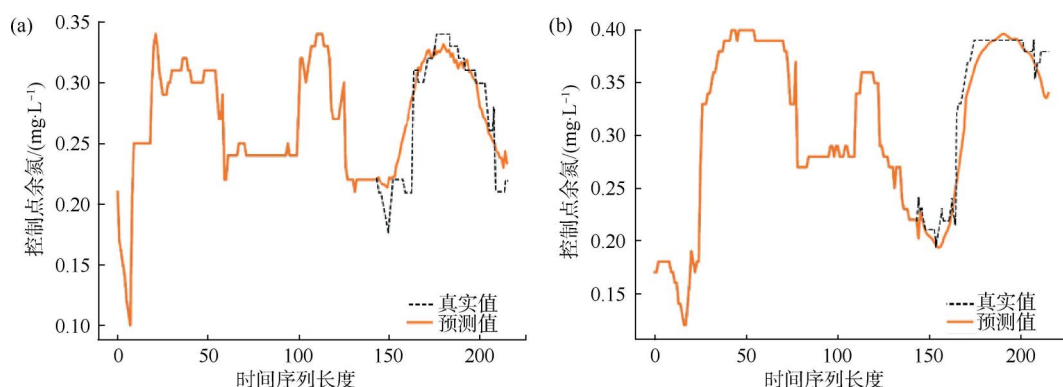


图 10 Patch-Transformer 模型测试集上的预测结果

Fig. 10 Prediction Results of Patch-Transformer on the Test Sets

用于预测管网控制点的余氯浓度预测,并对所构建的模型进行优化,得出结论如下。

(1)初步构建 Transformer 模型,通过特征消融试验确定最佳的特征组合,并探究不同序列长度和预测长度对模型预测性能的影响,使用 C_0 、 T_{turb} 及 C_3 的三特征组合进行模型构建,并将序列长度确定为 144,预测长度确定为 72。在该条件下,Transformer 模型的 δ 值为 5.95%。

(2)使用 Patch 模块对 Transformer 模型进行优化,并探究其最优参数,最终,Batch_size 为 128、Dropout 值为 0.1 的 Patch-Transformer 模型取得了最佳预测效果, δ 值为 5.34%。若将原有的机理模型用于每日的管网控制点余氯浓度预测, δ 值为 9.13%,而初始构建的 Transformer 模型 δ 值为 6.75%,相比之下,最终优化的 Patch-Transformer 余氯衰减智能模型的预测性能分别提升了 41.51%及 20.89%,效果提升明显。

(3)所构建的 Patch-Transformer 余氯衰减智能模型,能够有效地预测供水管网末梢的余氯,精度较高,后续应进一步探究其他深度学习模型在余氯衰减预测中的应用。

参考文献

- [1] 邓艳玲,曾强. 消毒副产物暴露对胎儿生长发育的影响:研究进展与展望[J]. 环境与职业医学, 2022, 39(6): 714-721.
- DENG Y L, ZENG Q. Effects of exposure to disinfection byproducts on fetal growth and development: Research progress and perspectives[J] Environment and Occupational Medicine, 2022, 39(6): 714-721.
- [2] 林晓丹,陈方亮,强志民,等. 供水管网水质微观机理模型

研究及应用进展[J]. 中国给水排水, 2021, 37(16): 41-47, 53.

LIN X D, CHEN F L, QIANG Z M, et al. Research and application progress on micro-mechanism model of water quality in drinking water distribution systems [J] China Water & Wastewater, 2021, 37(16): 41-47, 53.

- [3] 林晓丹,陈方亮,强志民,等. 供水管网水质模型的研究进展:经验统计模型[J]. 中国给水排水, 2021, 37(14): 1-7.

LIN X D, CHEN F L, QIANG Z M, et al. Research progress on water quality model of water distribution systems: Empirical statistic models[J] China Water & Wastewater, 2021, 37(14): 1-7.

- [4] ABOKIFA A A, YANG Y J, LO C S, et al. Water quality modeling in the dead end sections of drinking water distribution networks[J]. Water Research, 2016, 89: 107-117. DOI: 10.1016/j.watres.2015.11.025.

- [5] 唐宇,徐冰峰,山丕斌,等. Bagging 集成 RW-RBF 的水厂消毒预测模型[J]. 能源工程, 2023, 43(1): 66-71.

TANG Y, XU B F, SHAN P B, et al. Water plant disinfection prediction model based on bagging integrated RW-RBF[J] Energy Engineering, 2023, 43(1): 66-71.

- [6] 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 and Computer Modelling, 2006, 44(5/6): 469-484. DOI: 10.1016/j.mcm.2006.01.006.

- [7] WADKAR D V, NANGARE P, WAGH M P. Evaluation of water treatment plant using Artificial Neural Network (ANN) case study of pimpri chinchwad municipal corporation (PCMC) [J]. Sustainable Water Resources Management, 2021, 7(4): 52.

- [8] SHIM J, PARK S, CHO K H. Deep learning model for simulating influence of natural organic matter in nanofiltration [J]. Water Research, 2021, 197: 117070. DOI: 10.1016/j.watres.2021.117070.

(下转第 143 页)

- aerobic granular sludge at various organic loading rates [J]. *Process Biochemistry*, 2009, 44(2): 242–245.
- [19] WU J, JIANG B, FENG B, et al. Pre-acidification greatly improved granules physicochemical properties and operational stability of upflow anaerobic sludge blanket (UASB) reactor treating low-strength starch wastewater [J]. *Bioresource Technology*, 2020, 302: 122810. DOI: 10.1016/j.biortech.2020.122810.
- [20] BHUNIA P, GHANGREKAR M M. Required minimum granule size in UASB reactor and characteristics variation with size[J]. *Bioresource Technology*, 2007, 98(5): 994–999.
- [21] 周滢月, 张智, 李蕾, 等. 底物对厌氧消化系统 EPS 产生及起泡的影响[J]. *中国环境科学*, 2023, 43(8): 4046–4056.
ZHOU Y Y, ZHANG Z, LI L, et al. Influence of substrate composition on the production characteristics of extracellular polymers and foaming performance in anaerobic digestion systems [J]. *China Environmental Science*, 2023, 43(8): 4046–4056.
- [22] 季军远, 林久淑, 朱晓桐, 等. ANAMMOX-UASB 反应器启动过程中的生物特性[J]. *环境工程学报*, 2021, 15(10): 3358–3367.
JI J Y, LIN J S, ZHU X T, et al. Biological characteristics of ANAMMOX-UASB reactor during startup process [J]. *Chinese Journal of Environmental Engineering*, 2021, 15(10): 3358–3367.
- [23] DONG J, ZHANG Z, YU Z, et al. Evolution and functional analysis of extracellular polymeric substances during the granulation of aerobic sludge used to treat *p*-chloroaniline wastewater [J]. *Chemical Engineering Journal*, 2017, 330: 596–604. DOI:10.1016/j.cej.2017.07.174.
- [24] 李亚静, 王少坡, 刘璐, 等. 有机负荷对污泥胞外聚合物 (EPS)分泌特性及信号分子释放差异的影响[J]. *环境工程*, 2022, 40(2): 47–52.
LI Y J, WANG S P, LIU L, et al. Secretory characteristics of EPS and the signal molecules release under different organic loading [J]. *Environmental Engineering*, 2022, 40(2): 47–52.
- [25] SHENG G P, YU H Q, LI X Y. Extracellular polymeric substances (EPS) of microbial aggregates in biological wastewater treatment systems: A review [J]. *Biotechnology Advances*, 2010, 28(6): 882–894.
- [26] CHEN H, HU H, CHEN Q, et al. Successful start-up of the anammox process: Influence of the seeding strategy on performance and granule properties [J]. *Bioresource Technology*, 2016, 211: 594–602. DOI: 10.1016/j.biortech.2016.03.139.
- [27] 蔡春光, 刘军深, 蔡伟民. 胞外多聚物在好氧颗粒化中的作用机理[J]. *中国环境科学*, 2004, 24(5): 623–626.
CAI C G, LIU J S, CAI W M. Action mechanism of extracellular polymers on the aerobic granulation [J]. *China Environmental Science*, 2004, 24(5): 623–626.

(上接第 72 页)

- [9] CHEN S Y, HUANG J L, WANG P, et al. A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation [J]. *Water Research*, 2024, 248: 120895. DOI: 10.1016/j.watres.2023.120895.
- [10] JIANG Y Q, LI C L, SONG H X, et al. Deep learning model based on urban multi-source data for predicting heavy metals (Cu, Zn, Ni, Cr) in industrial sewer networks [J]. *Journal of Hazardous Materials*, 2022, 432: 128732. DOI: 10.1016/j.jhazmat.2022.128732.
- [11] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]. Long Beach: Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS), 2017.
- [12] HUANG Z L, ZHANG L Y, ZHANG Y, et al. Transformer based multi-output regression learning for wastewater treatment [C]. Washington D. C.: Proceedings of the IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), 2021.
- [13] 刘雅亭, 黄显怀, 杨伟伟, 等. 基于 Transformer 深度学习的供水量预测研究 [J]. *给水排水*, 2024, 50(3): 148–153, 162.
LIU Y T, HUANG X H, YANG W W, et al. Research of water volume prediction method based on Transformer deep learning [J]. *Water & Wastewater Engineering*, 2024, 50(3): 148–153, 162.
- [14] WANG R Q, QI Y, ZHANG Q, et al. A multi-step water quality prediction model based on the Savitzky-Golay filter and Transformer optimized network [J]. *Environmental Science and Pollution Research*, 2023, 30(50): 109299–109314.
- [15] 舒诗湖, 赵明, 何文杰, 等. 供水管网水力、水质模型校核标准探讨 [J]. *中国给水排水*, 2008, 24(18): 104–106.
SHU S H, ZHAO M, HE W J, et al. Discussion on calibration standard for hydraulic and water quality model of water distribution network [J]. *China Water & Wastewater*, 2008, 24(18): 104–106.
- [16] NIE Y, NGUYEN N H, SINTHONG P, et al. A time series is worth 64 words: Long-term forecasting with transformers [C]. Online International Conference on Learning Representations, 2022.