

doi: 10.11835/j.issn.1000-582X.2025.11.008

引用格式:叶彬强,曹雪杰,李东,等.基于时空图卷积网络的水质多参数预测算法[J].重庆大学学报,2025,48(11): 92-105.



基于时空图卷积网络的水质多参数预测算法

叶彬强^{1a,2},曹雪杰²,李东^{1a},陈昶宏²,刘宏²,汤斌²,冯鹏^{1b}

(1. 重庆大学 a. 微电子与通信工程学院; b. 光电技术及系统教育部重点实验室,重庆 400044;

2. 重庆理工大学 两江人工智能学院,重庆 400054)

摘要:水质预测是水资源生态治理体系的核心环节,在时间演变、随机扰动及环境变迁等多重因素作用下,水质监测数据呈现出非平稳动态特征与多维非线性耦合关系。针对流域水质多参数预测中时空依赖复杂、传统模型难以有效融合动态拓扑与长周期特征的问题,文中提出了一种基于时空图卷积网络的水质预测算法。在时间维度上,设计掩码子序列 Transformer 模块,通过随机掩码自监督预训练任务使模型从长周期数据中提取趋势特征,结合扩张因果卷积捕获水质特征的累积效应,解决传统模型对突变事件的响应滞后问题。在空间维度上,构建动态图结构学习模块,融合基于站点物理距离的预定义邻接矩阵与动态残差图生成动态图结构。实验结果表明,相较于其他水质时空预测模型,该模型具有更好的预测精度,对所有水质指标的预测 R^2 均能达到 93% 以上。

关键词:时空图卷积网络;水质预测;动态图卷积

中图分类号:TP399; X832

文献标志码:A

文章编号:1000-582X(2025)11-092-14

A multi-parameter prediction algorithm for water quality based on spatial-temporal graph convolutional networks

YE Binqiang^{1a,2}, CAO Xuejie², LI Dong^{1a}, CHEN Changhong²,

LIU Hong², TANG Bin², FENG Peng^{1b}

(1a. School of Microelectronics and Communication Engineering; 1b. Key Laboratory of Optoelectronic Technology and Systems of Ministry of Education, Chongqing University, Chongqing 400044, P. R. China;
2. School of Artificial Intelligence, Chongqing University of Technology, Chongqing 400054, P. R. China)

Abstract: Accurate water quality prediction is essential for effective ecological water management. However, water quality exhibits complex non-stationary dynamics and multi-dimensional nonlinear relationships driven by temporal evolution and environmental variability. In multi-parameter river water quality prediction, intricate spatial-temporal dependencies make it difficult for traditional models to effectively integrate dynamic topology

收稿日期:2025-05-27

基金项目:中央高校基本科研业务费资助项目(2023CDJKYJH085);重庆市九龙坡区科技计划项目(2023TJ2001);重庆市青少年创新人才培养计划(CY240903)。

Supported by the Fundamental Research Funds for the Central Universities (2023CDJKYJH085), the Science and Technology Planning Project of Jiulongpo District, Chongqing (2023TJ2001), and the Chongqing Innovative Talent Training Program for Adolescents (CY240903).

作者简介:叶彬强(1981—),男,博士,主要从事人工智能、环境监测方向研究,(E-mail) ybq@cqu.edu.cn。

通信作者:李东,男,教授,博士生导师,(E-mail) lidongcuit@126.com。

and long-period features. To address this challenge, we propose a spatial-temporal graph convolutional network (STGCN) model. In the temporal dimension, a masked sub-series transformer module is employed to extract long-term trends through self-supervised pretraining. Combined with dilated causal convolution, it captures cumulative water quality effects and alleviates the response lag common in traditional models when facing abrupt changes. In the spatial dimension, a dynamic graph learning module integrates a predefined station-distance adjacency matrix with a dynamic residual map to generate adaptive graph structures. Experimental results demonstrate that the proposed model outperforms existing methods in water quality prediction, achieving an R^2 greater than 0.93 across all water quality indicators.

Keywords: spatial-temporal graph convolutional networks; water quality prediction; dynamic graph convolution

水质监测对于生态保护、水资源管理及公共健康至关重要^[1]。随着水污染问题的日益严重,传统的水质监测方法常常面临反应速度慢、成本高、空间覆盖不广等问题。因此,多参数水质预测成为了一项重要的研究课题,尤其是水质的实时监测和动态预测具有重要的应用前景。水质的评估通常依赖于一系列物化参数,如 pH、溶解氧(dissolved oxygen, DO)、化学需氧量(chemical oxygen demand, COD)等^[2]。随着数据科学和机器学习的进步,传统的水质预测方法逐渐被更为高效、准确的算法替代。在水质时间序列数据处理和多变量预测方面,水质预测技术已取得显著进展^[3]。

当前,国内外主要采用的预测方法可以分为 3 大类:统计分析方法、机器学习方法和深度学习方法^[4]。其中,基于统计分析的方法是最早应用于水质预测的手段。研究者们通过时间序列分析、回归分析及参数统计等手段提取水质数据的基本趋势并发掘水质数据的周期性和季节性变化。Agaj 等^[5]使用自回归差分移动平均模型(autoregressive integrated moving average model, ARIMA)对科索沃两处水站的历史水位月度数据进行预测,验证了该统计分析模型在预测水位方面的有效性。Du 等^[6]通过马尔可夫状态转移矩阵修正 ARIMA 残差的累积误差,首次将动态误差校正引入用水预测,解决周期性数据随机波动难题。Wang 等^[7]引入了 Holt-Winters 季节模型对 ARIMA 进行优化,并建立了总磷和总氮预测模型,通过模型的自校正大幅提高了水质指标的预测准确率。

相较于统计方法,机器学习在处理噪声和复杂环境问题上表现更为出色。研究者主要围绕数据建模和模型超参数等方面进行优化。马创等^[8]提出在支持向量机(support vector machine, SVM)模型的基础上使用遗传算法进行特征选择能够有效提升水质分类的预测性能。巫莉莉等^[9]利用基于随机森林(random forest, RF)的集成式机器学习方法对养殖水体中的氨氮含量进行预测,实现了循环水养殖水体中氨氮含量精准预测。Nasir 等^[10]评估了多种机器学习分类器以及集成模型对水质指标的分类效果,验证了堆叠集成模型的有效性。

深度学习方法通过构建多层神经网络捕捉长时依赖性和多变量交互作用,实现从大量数据中自动提取复杂的时空特征。其中,以长短期记忆网络(long short-term memory, LSTM)为代表的时序模型能够有效捕捉局部时间依赖性^[11]。Wu 等^[12]提出多任务深度学习模型 MTL-CNN-LSTM,通过共享多断面水质时空关联特征使水质 COD 预测误差大幅降低。Wang 等^[13]提出基于贝叶斯优化驱动的 LSTM-GRU 混合模型,通过特征选择与数据时频优化,使水质预测 RMSE 较单一 LSTM 降低 18.1%,验证了时序特征增强在动态水质建模中的协同优势。Qiao 等^[14]提出 FSGCN(spatial-temporal graph convolution neural networks, FSGCN)模型将时间注意力机制、图卷积网络与时间卷积残差结构相结合,实现传感器间信息的有效传递与融合。

尽管深度学习方法在水质预测中取得了显著的进展,但仍存在以下几个问题。1)现有模型主要集中在单一水质指标的预测上,对于多个水质参数的联动预测效果较差。2)水质指标通常存在空间和时间上的相关性,这些相关性来自于水体的不同位置、污染源的不同影响等因素。现有的模型在捕捉这些复杂的依赖关系方面存在不足,导致预测精度难以提升。3)水质数据通常呈现长期的趋势性变化,虽然循环神经网络(recurrent neural network, RNN)和 LSTM 在处理时间序列数据方面表现优异,但这些模型仍然在处理长期依赖

关系时存在困难,在捕捉长期依赖时容易出现梯度消失或爆炸的问题,影响模型的稳定性和准确性。

文中提出了一种水质多参数预测模型,为了更好地获取水质指标之间的长期依赖,设计了基于多头自注意力的特征注意机制模型,通过掩码重建任务进行预训练获得更好的压缩和上下文子序列时间表示。考虑到相邻站点水质指标之间隐藏的关系,设计了动态图结构学习模块,通过可学习参数平衡先验知识与训练过程中发现的隐藏图结构,生成站点间最优拓扑。最后,将不同周期的水质特征进行融合,实现水质多参数指标的准确预测。在真实数据集上对所提出的框架进行了评估,实验结果表明,文中模型具有更好的预测精度,能够实现水质参数的高精度预测,为流域污染监控和水资源保护提供科学依据。

1 方法理论

1.1 时序卷积网络

时间卷积网络(temporal convolutional network, TCN)是一种基于一维卷积的时序建模方法,适用于捕捉时间序列中的动态模式,相比传统的LSTM或RNN等循环神经网络,TCN具有计算效率高、不存在梯度消失问题等优势,非常适合捕捉时间序列中的短期波动和长期趋势。在水质预测中,TCN可用于分析各站点水质指标随时间变化的规律。其卷积操作定义为

$$y = (x * w)(t) = \sum_{k=0}^{K-1} x(t-k) w(k), \quad (1)$$

式中: x 为输入时间序列; w 为卷积核;长度为 K ; $y(t)$ 为时间步 t 的输出。TCN通过堆叠多层一维卷积并结合扩张因果卷积,能够捕捉水质数据中的长期趋势。扩张卷积通过在卷积核中插入间隔扩展网络的感受野,其结构如图1所示。同时,为了防止信息的泄漏,网络还采用了因果卷积,即采用了按照时间序列顺序的方法进行卷积运算。时刻 t 的卷积输出只与时刻 t 和时刻 $t-1$ 或更早的数据相关,并通过补零操作使得隐藏层与输入层之间的长度总是保持恒定。

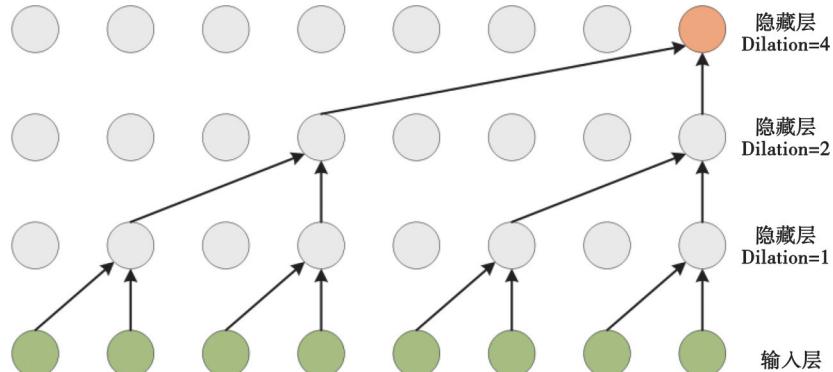


图1 扩张因果卷积示意图

Fig. 1 Schematic diagram of dilated causal convolution

1.2 自注意力机制

自注意力机制是深度学习领域的重要突破,核心思想是通过动态计算序列内部元素之间的相关性权重,从而捕捉长距离依赖关系并建立全局上下文关联。主要过程是对于输入序列中的每个元素,通过计算其与序列中所有元素的相似度得分,生成注意力权重分布,进而对输入特征进行加权聚合。这种机制突破了传统循环神经网络RNN和CNN在建模长程依赖时的局限性,同时具备并行化计算能力。

对于给定输入序列 $X = [x_1, x_2, \dots, x_n] \in R^{n \times d}$,其中, n 为序列长度, d 为特征维度。自注意力的计算主要可以分为3步。通过线性变换将每个输入元素 x_i 映射为3个不同的向量,查询向量(Query)键向量(Key)和值向量(Value):

$$q_i = \mathbf{W}^Q x_i, \quad (2)$$

$$k_i = \mathbf{W}^K x_i, \quad (3)$$

$$v_i = \mathbf{W}^V x_i, \quad (4)$$

其中, $\mathbf{W}^Q, \mathbf{W}^K, \mathbf{W}^V \in R^{d \times d_k}$ 为可学习参数矩阵。通过计算 Query 与所有 Key 的点积相似度, 并对其进行行方向的 Softmax 归一化处理, 得到注意力权重矩阵:

$$\alpha_{ij} = \text{Softmax} \left(\frac{q_i^T k_j}{\sqrt{d_k}} \right). \quad (5)$$

将所有值向量计算加权得到输出向量 z_i :

$$z_i = \sum_{j=1}^n \alpha_{ij} v_j. \quad (6)$$

这种全局交互机制使得每个位置都能直接访问序列中所有其他位置的信息, 特别适合处理需要全局语义理解的场景。为进一步增强模型的表达能力, 还可以采用多头注意力机制, 将查询、键、值向量分别投影到 h 个不同的子空间进行并行计算, 每个子空间对应一个注意力头, 再将各头的结果拼接后通过线性变换得到最终输出。多头机制允许模型在不同子空间中学习多样化的注意力模式, 提升模型的表征能力。

2 水质多参数预测网络

在水质多参数预测问题中, 水质监测网络采集的多元时序数据可视为时空图结构数据。设水质监测网络中包含 V 个空间节点, 每个节点在 t 时刻采集 C 维水质参数, 构成特征向量 $\mathbf{x}_t^{(v)} \in R^C$ 。时空图结构通过邻接矩阵 $\mathbf{A} \in R^{V \times V}$ 描述节点间的空间拓扑关系。水质多参数预测任务旨在建立 1 个映射函数 $f(\cdot)$, 在给定历史观测窗口条件下 $\mathcal{X}_t = \{X_{t-T+1}, \dots, X_t\} \in R^{T \times V \times C}$ (其中, T 为时间窗口长度, $\mathbf{X}_t = [\mathbf{x}_t^{(1)}, \dots, \mathbf{x}_t^{(V)}]^T$ 表示 t 时刻所有节点构成的水质参数矩阵), 输出未来 τ 个时间步的全节点多参数预测序列 $\mathcal{Y}_t = \{Y_{t+1}, \dots, Y_{t+\tau}\} \in R^{\tau \times V \times C'}$, 其中 C' 为待预测参数的维度。

该问题的核心在于同时建模水质参数的时空依赖性, 主要包括: 1) 时间维度上参数值的动态演变规律; 2) 空间维度上监测节点间的扩散传播效应。因此, 文中基于时空图卷积网络设计了 1 个水质多参数预测模型, 主要通过时序特征学习、空间拓扑感知和特征融合 3 个部分实现水质时空模式的联合建模。

2.1 模型总体框架

水质数据具有明显的时空依赖性, 这是由于水质受自然节律周期和污染降解的影响。水体的自净作用、污染物扩散等过程都具有时间累积和空间扩散效应, 传统模型难以捕捉水质的复杂时空依赖。因此, 文中设计了基于掩码子序列 Transformer (masked sub-series transformer, MST) 的时序特征学习模块增强模型模拟水质数据时间依赖性的提取能力, 并引入自适应邻接矩阵结合预定义拓扑, 增强模型对扩散效应的建模能力。模型的整体框架如图 2 所示。

该模型主要包括 2 个部分, 时序特征提取模块和特征融合模块。其中, 时序特征提取模块为一个自监督掩码重构任务, 该模块进行时序特征的解耦, 对于模型输入的长周期数据 $\mathcal{X}_{\text{long}}$, 将其切分为多个不重叠的短周期数据 $\mathcal{X}_{\text{short}}$, 通过 Transformer 编码器提取时序的压缩表示 S_{unmasked} , 训练过程中设置较高的掩码比, 迫使模型学习更长的时间序列特征。另一个部分为特征融合模块, 将时序特征提取模块得到的特征表示 S_{unmasked} 通过堆叠的扩张因果卷积进行进一步的压缩, 得到水质的长期趋势特征, 再使用时空图卷积网络分别处理 $\mathcal{X}_{\text{short}}$ 和 S_{unmasked} 得到水质的短期波动特征, 最后将长期趋势特征与短期波动特征进行拼接, 得到融合特征, 最后通过全连接层输出模型的预测结果。

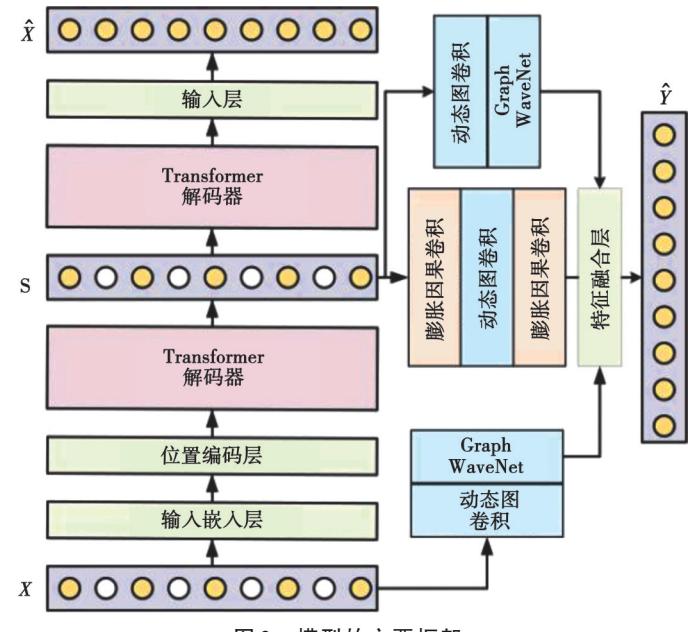


图 2 模型的主要框架

Fig. 2 Main framework of the model

2.2 时序特征提取模块

基于 MST 的时序特征提取模块结构如图 3 所示, 该过程通过掩码子序列重构任务, 迫使模型从长时序数据中学习压缩的子序列表示。将输入的长历史序列划分为子序列, 以高比例随机掩码后, 利用 Transformer 编码器从未掩码子序列及其上下文中提取潜在特征, 并通过重构掩码部分优化模型。这一设计旨在解决传统模型依赖短期数据的局限性, 通过自监督学习隐式水网的长期趋势和周期性特征, 同时增强模型对缺失数据的鲁棒性, 为下游预测任务提供更好的时空表示基础。

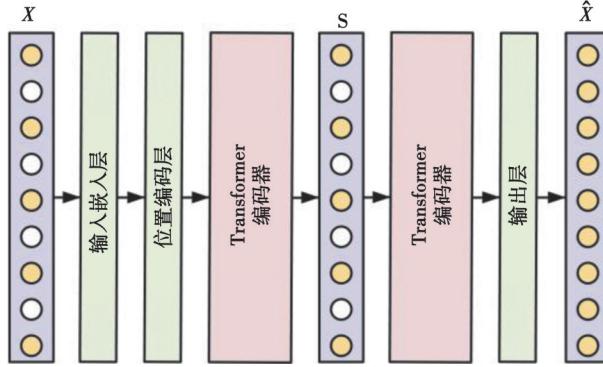


图 3 预训练自监督模块

Fig. 3 Pretrained self-supervised module

首先, 对给定的长时序输入 $X_{\text{long}} \in R^{N \times L}$ 进行划分, 其中, N 为水质指标的数量, L 为时间步长度。将其划分为 $I = \lfloor L/S \rfloor$ 个非重叠的子序列, 每个子序列包含 s 个时间步:

$$X_{\text{long}} = [X^{(1)}, X^{(2)}, \dots, X^{(I)}], \quad (7)$$

式中: $X^{(i)} \in R^{N \times s}$ 表示第 i 个子序列。得到子序列后, 对输入子序列进行随机掩码, 以一定的比例随机选择子序列进行掩码, 生成掩码集 M 和剩余未掩码子序列 U 。掩码后的输入表示为

$$X_{\text{masked}} = X^{(i)} \odot I(i \notin M), \quad (8)$$

式中: \odot 为逐元素乘积; $I(\cdot)$ 为指示函数。得到掩码输入后, 通过时序特征学习模块提取未掩码子序列的上下文表示为

$$S_{\text{unmasked}} = \text{STRL}(X_{\text{unmasked}}) \in R^{|U| \times d}, \quad (9)$$

式中: $X_{\text{unmasked}} = [X^{(i)}]_{i \in U}$ 为未掩码子序列的集合; d 为编码器隐藏层的维度。

然后, 进行重构任务, 将未掩码子序列的表示 S_{unmasked} 与可学习的掩码占位符 $S_{[\text{MASK}]} \in R^{|M| \times d}$ 拼接后, 输入到 Transformer 层和线性层进行掩码部分序列的重构, 该过程表示为

$$[\hat{X}_{\text{masked}}, \hat{X}_{\text{unmasked}}] = \text{TaskHead}([S_{\text{unmasked}}, S_{[\text{MASK}]}]). \quad (10)$$

目标是使重构的 \hat{X}_{masked} 尽可能接近真实的 X_{masked} , 在计算损失时只计算掩码部分的损失, 该模块的损失表示为

$$L_{\text{pretrain}} = \frac{1}{|M|} \sum_{i \in M} \|X_{\text{masked}} - \hat{X}_{\text{masked}}\|_2^2 \quad (11)$$

在该过程中, 可以通过设置高掩码比迫使模型从稀疏上下文中学习压缩的时间序列表示, 为下游预测任务提供稳健的表示基础。

2.3 动态图构建

模型结合预定义拓扑与自适应邻接矩阵进行水网动态图结构的构建, 通过可学习参数平衡先验知识与训练过程中发现的隐藏图结构。该模块基于实际水质检测站点的节点空间距离和拓扑关系生成预定义邻接矩阵 $A_{\text{pre}} \in R^{N \times N}$, 随机初始化一个可学习的嵌入矩阵 $E_1, E_2 \in R^{N \times d}$ 和自适应系数 $\lambda \in R$, 其中 d 为嵌入维度, N

为站点数量,通过反对称变换捕获图结构的潜在关联模式,自适应邻接矩阵的生成如式(12)所示,再将预定义邻接矩阵 A_{pre} 和自适应邻接矩阵 A_{res} 进行加权得到动态图,其计算式为

$$A_{\text{res}} = \text{Tanh} \left(k \left(E_1 E_2^T - E_2 E_1^T \right) \right), \quad (12)$$

$$A = A_{\text{pre}} + \lambda \cdot \text{ReLU}(A_{\text{res}}), \quad (13)$$

式中: k 为缩放因子; λ 为可训练参数,用于控制残差图的贡献度。 A 为融合后图结构,在训练过程中,嵌入矩阵通过反向传播优化,其计算式为

$$\frac{\partial \mathcal{L}}{\partial E_1} = \frac{\partial \mathcal{L}}{\partial A} \cdot \frac{\partial A}{\partial E_1}. \quad (14)$$

同理,可计算 E_2 的梯度。在训练过程中,每隔一段时间还需要对 A_{res} 进行稀疏化处理,首先计算 A_{res} 的绝对值均值 μ ,再对 A_{res} 进行阈值过滤,当 $A_{\text{res}}^{ij} < 0.5\mu$ 时,将其置为0。通过该方法动态生成水网的图结构不仅可以充分利用现有的水质站点之间的关系,还通过学习残差关联图优化了图结构,使模型能够学习由于流速、落差等因素造成的隐藏图结构。

3 数值实验

3.1 数据集构建

研究数据来源于国家地表水水质自动监测实时数据发布系统,覆盖了2021年6月—2025年1月全时段连续监测记录,涉及重庆市内的42个自动化监测站点。各点位采集9项关键参数,包括水温、pH、溶解氧、电导率、浊度、高锰酸盐指数(COD_{mn})、氨氮(NH3-N)、总磷(TP)、总氮(TN)。每站点数据采集时间间隔为4 h,得到可用于分析处理的数据共计707周。

为了筛选出其中关联性较强的特征,研究采用相关系数分析法,计算特征序列之间的皮尔逊相关系数,绘制热力图如图4所示。以0.4为分界点,剔除与其他水质指标的相关系数绝对值均 ≤ 0.4 的指标,保留6个水质特征,分别为温度、pH、溶解氧、电导率、高锰酸盐指数、总氮。

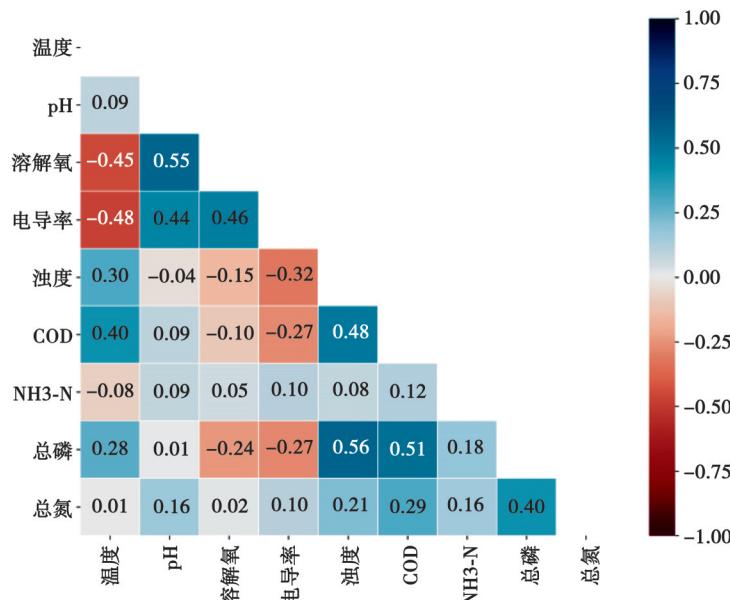


图4 不同水质指标间的相关系数图

Fig. 4 Correlation matrix of different water quality indicators

3.2 数据预处理

3.2.1 缺失数据处理

水质多参数监测数据缺失现象是水质检测系统运行中不可避免的问题。传感器电极老化、藻类生物附着、监测站点周期性维护以及突发性灾害都可能导致水质参数的间歇性或连续性丢失。对于水质数据的缺失需要考虑水质时空数据中的复杂依赖关系,而不是仅仅依赖于局部时间或空间邻点,因此文中使用SAITS^[15]、TimesNet^[16]、CSDI^[17]、USGAN^[18]和MRNN^[19]等针对长连续缺失和跨时空混合缺失等复杂场景设计的模型对数据中的连续缺失进行处理。这些模型利用神经网络捕捉数据中的复杂非线性关系,对数据的时间和空间维度的依赖关系建模,避免了邻近点插值忽略远距离相关性或拓扑结构的问题。实验结果如表1所示,由表可知TimesNet模型的损失最低,可采用TimesNet模型对该数据集的缺失值进行填充。

表1 不同模型的缺失插值实验结果

Table 1 Experimental results of missing data interpolation for different models

参数	SAITS	MRNN	CSDI	TimesNet	USGAN
MAE	0.145 1	0.726 8	0.136 7	0.109 5	0.289 5

3.2.2 空间关系构建

空间关联网络的具体构建过程为:先基于修正后的河网拓扑关系,新建有向线要素图层,将相邻站点按河道流向连接,对存在河道弯曲或江心洲分割的复杂河段,采用手动方式连接,再运用ArcMap的几何测量工具计算各连接线的真实河道长度,该参数将作为图结构的权重系数参与后续模型的训练。同时,模型的训练还需要构建各站点的空间拓扑图。根据站点间空间距离和站点的拓扑结构计算站点邻接矩阵的空间权重。站点邻接关系的定义和空间权重计算为

$$A_{ij} = \begin{cases} 1, & d_{ij} \leq d_{\max} \\ 0, & \text{others} \end{cases}, \quad (15)$$

$$\tilde{W} = \frac{\frac{A_{ij}}{1 + \alpha d_{ij}}}{\sum_{k=1}^n \frac{A_{ik}}{1 + \alpha d_{ik}}}, \quad (16)$$

式中: \tilde{W} 为归一化后的权重矩阵; d_{ij} 为站点间实际河道距离; d_{\max} 为最大有效作用距离,用于排除远距离伪关联; α 为衰减系数,用于量化距离衰减效应。监测站点间的空间关联强度随距离增大呈现非线性衰减特性。当两站点间距超过预设距离阈值 d_{\max} 时,其关联路径将被关闭,能够有效消除远距离伪相关对模型的干扰。同时,通过预定义站点空间权重,将原始全连接网络的边数量级从 $O(N^2)$ 压缩至 $O(N)$,从而降低模型的计算复杂度。

3.2.3 数据集划分

为适配时序预测任务,确保模型训练时仅使用历史数据预测未来状态,避免数据泄露风险,采用滑动窗口的方式生成样本数据,设定训练集验证集和测试集的比例为0.6:0.2:0.2,使用大小为28的滑动窗口沿时间轴截取数据段。原始4 246个检测周期的数据经窗口滑动后得到的训练集、测试集和验证集数量分别为2 548、849、849组。

3.3 实验结果

3.3.1 预测效果

模型对温度、总氮、电导率、溶解氧、pH、高锰酸盐指数6项水质指标的均方根误差(RMSE)分别为0.573 9、0.374 0、15.220 7、0.690 9、0.104 3、0.683 3,预测效果如图5所示,可以看出,本模型对水质指标的预测较为准确,对于水质指标的突变反应灵敏。

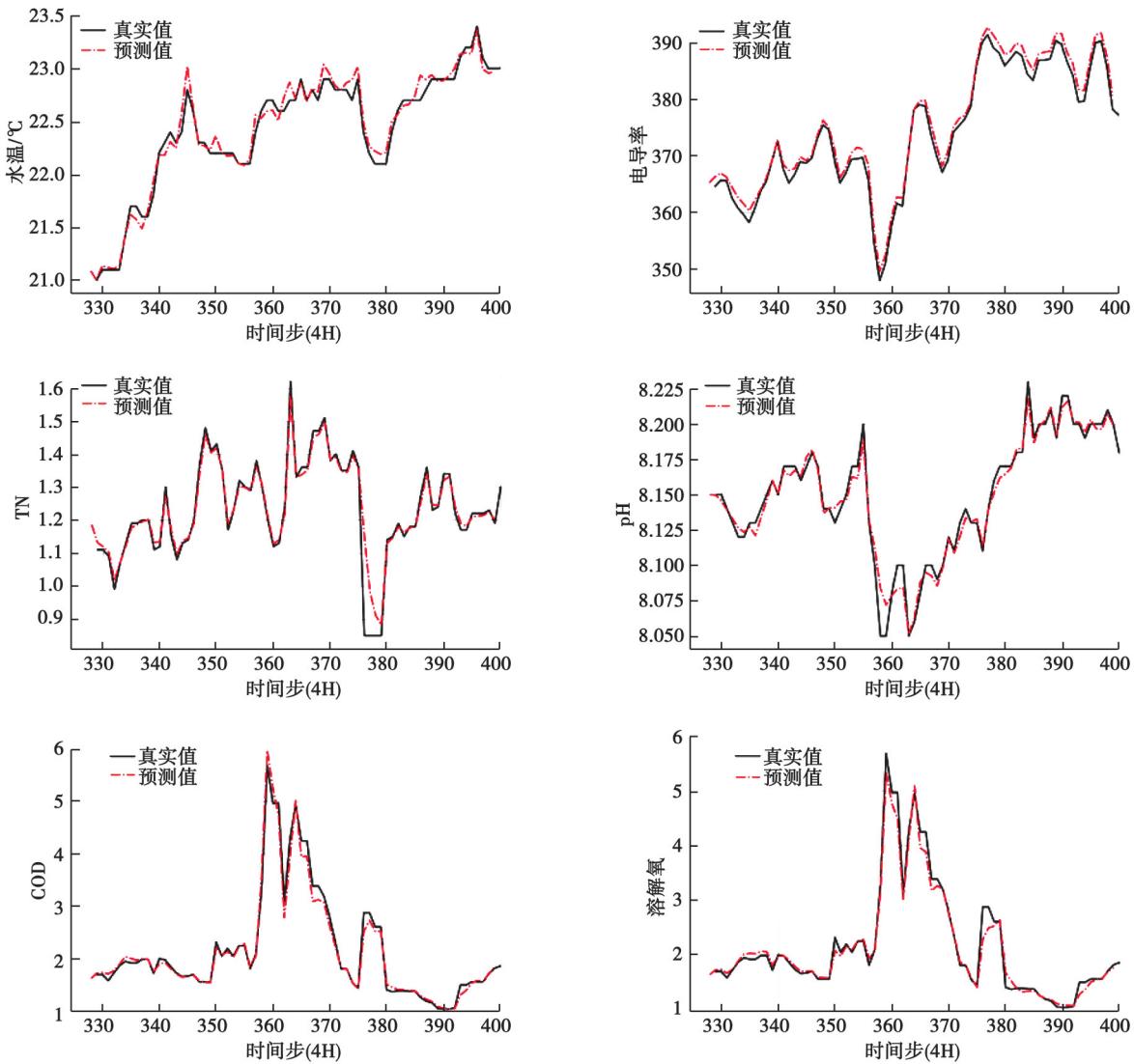


图 5 文中方法水质各指标的预测效果

Fig. 5 Prediction performance of various water quality indicators using the proposed method

3.3.2 对比实验

为全面评估研究的效果, 实验选取 DCRNN^[20]、STGNN^[21]、MTGNN^[22]、StemGNN^[23]、STAEformer^[24]这 5 类具有代表性的多元时空序列预测模型作为基准, 不同方法对于水质多参数的检测效果如表 2 所示, 从模型整体预测结果来看, 基于 Transformer 架构的 STAEformer 与文中提出模型在多数参数上表现最佳, 而 MTGNN 因自适应图学习机制缺乏物理约束, 整体性能显著落后。文中模型在温度、电导率、溶解氧、pH 和高锰酸盐指数等 5 个参数的 MAE、RMSE 指标上均达到最低值, R^2 值最高, 总氮的 MAE、RMSE 达到了次优, R^2 值最高。验证了其时空特征融合架构的有效性。STAEformer 在总氮预测中表现最优 ($R^2=0.9408$), 但在其他水质参数上略逊于文中模型, 表明其注意力机制对多参数协同建模的泛化能力存在局限。而 MTGNN 因依赖纯数据驱动的图结构生成, 在温度预测中 MAE(1.5399) 和 RMSE(2.3894) 均异常偏高, $R^2(0.8467)$ 显著低于其他模型, 验证了水质预测中物理先验嵌入的必要性。

为直观评估模型对水质参数动态变化的捕捉能力, 研究随机选取 1 个站点, 选择 COD 和 TN 这 2 个具有典型波动特征的参数进行预测结果可视化, 结果如图 6 和图 7 所示。其中, COD 反映水体有机物污染程度, 其浓度受工业排放与微生物降解作用影响, 常呈现短时脉冲波动; TN 表征营养盐负荷, 突变多由农业面源径流或沉积物释放引发。实验结果表明, 文中模型在复杂动态预测中均展现出显著优势。

表2 不同方法的水质多参数预测效果

Table 2 Multi-parameter water quality prediction performance of different methods

参数	指标	DCRNN	STGNN	MTGNN	StemGNN	STAEformer	OURS
温度	MAE	0.399 4	0.500 4	1.539 9	0.527 5	0.341 3	0.309 7
	RMSE	0.690 8	0.752 8	2.389 4	0.845 2	0.602 9	0.573 9
	$R^2/\%$	98.87	98.64	84.67	98.19	99.10	99.20
总氮	MAE	0.112 8	0.121 2	0.118 7	0.146 2	0.099 5	0.101 9
	RMSE	0.393 1	0.388 4	0.437 8	0.412 0	0.378 8	0.374 0
	$R^2/\%$	92.04	93.20	86.55	90.97	94.08	93.47
电导率	MAE	6.287 7	8.998 2	7.371 6	6.740 7	5.873 8	5.858 2
	RMSE	17.800 4	19.929 7	19.152 1	17.403 7	15.184 3	15.220 7
	$R^2/\%$	98.38	97.81	97.96	98.37	98.78	98.75
溶解氧	MAE	0.229 4	0.244 0	0.752 0	0.296 8	0.216 4	0.213 1
	RMSE	0.713 9	0.732 9	1.160 9	0.799 6	0.697 0	0.690 9
	$R^2/\%$	92.83	92.26	87.14	90.15	93.03	93.27
pH	MAE	0.080 9	0.094 8	0.091 5	0.100 2	0.056 1	0.052 2
	RMSE	0.135 7	0.141 6	0.158 9	0.176 9	0.108 6	0.104 3
	$R^2/\%$	98.80	98.64	88.11	97.56	99.23	99.28
高锰酸盐指数	MAE	0.229 5	0.245 7	0.252 6	0.298 8	0.219 3	0.210 6
	RMSE	0.713 2	0.728 9	0.722 7	0.802 5	0.705 5	0.683 3
	$R^2/\%$	92.77	91.88	89.92	90.04	92.86	93.50

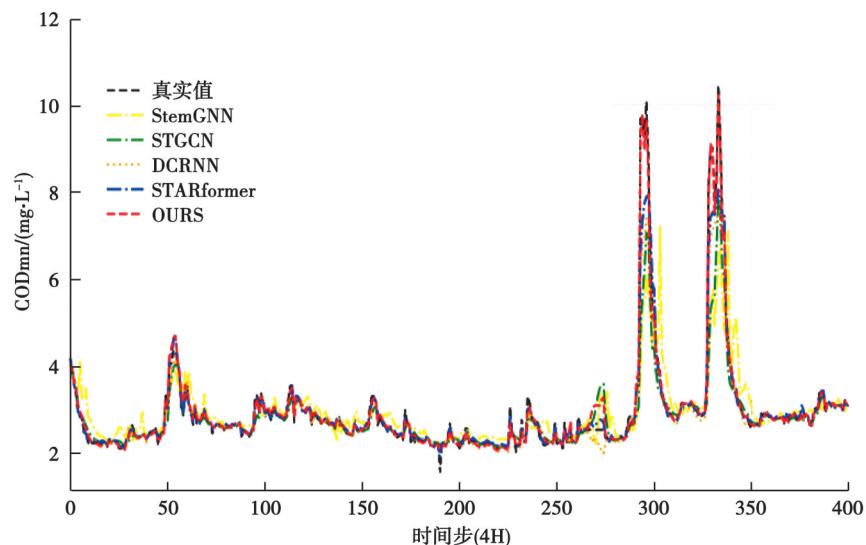


图6 各方法 COD 预测的效果

Fig. 6 COD prediction performance of different methods

由图可知,针对COD预测的突变响应结果,COD真实值在时间步300和340出现毛刺状突变。各模型预测表现差异较大,文中模型预测曲线紧密贴合真实值轨迹,滞后时间小于1个时间步;StemGNN模型受频域分解的平滑效应影响,预测曲线对突变响应迟缓,峰值增幅仅为真实值的70%左右,且滞后时间相较其他模型最大;DCRNN在275个时间步出现异常,可能是扩散卷积假设导致模型对上游污染源的过度敏感。图7显

示了各个模型对 TN 尖峰的捕捉能力, TN 在 290 和 330 个时间步时出现了 2 次陡峭尖峰, 峰值分别达到 2.9 mg/L 和 2.6 mg/L。文中模型成功捕捉 2 次突变的上升沿斜率, 这是由于其自适应图卷积模块通过融合流速数据, 准确识别了支流输入与沉积物再悬浮两类污染源的空间传导路径; 而 STGNN 使用静态图结构导致空间响应僵化, 在时间步 150 处出现突变后, 预测值产生明显的漂移。StemGNN 因频域能量分散表现最差, 其峰值预测结果最低; DCRNN 的扩散过程假设与 TN 的平流输送特性冲突, 预测曲线有时会呈现异常; STAEformer 在大部分情况下表现比较好, 但对于尖峰突变响应不如文中模型。

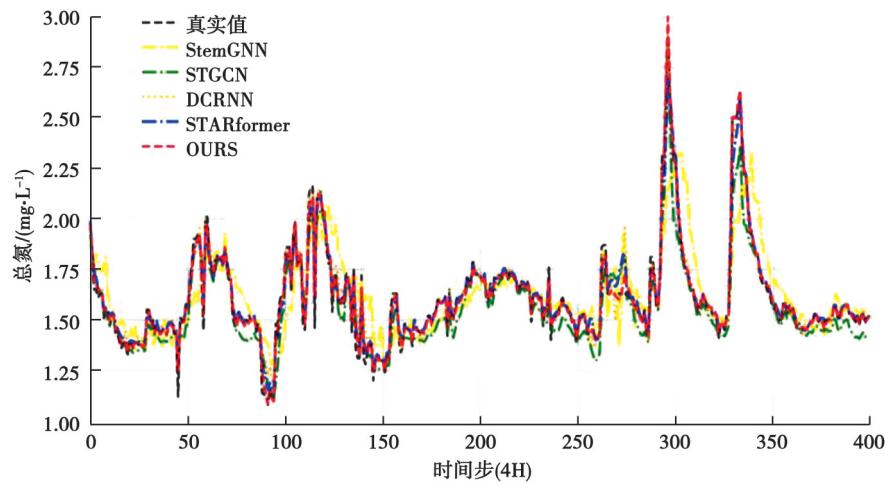
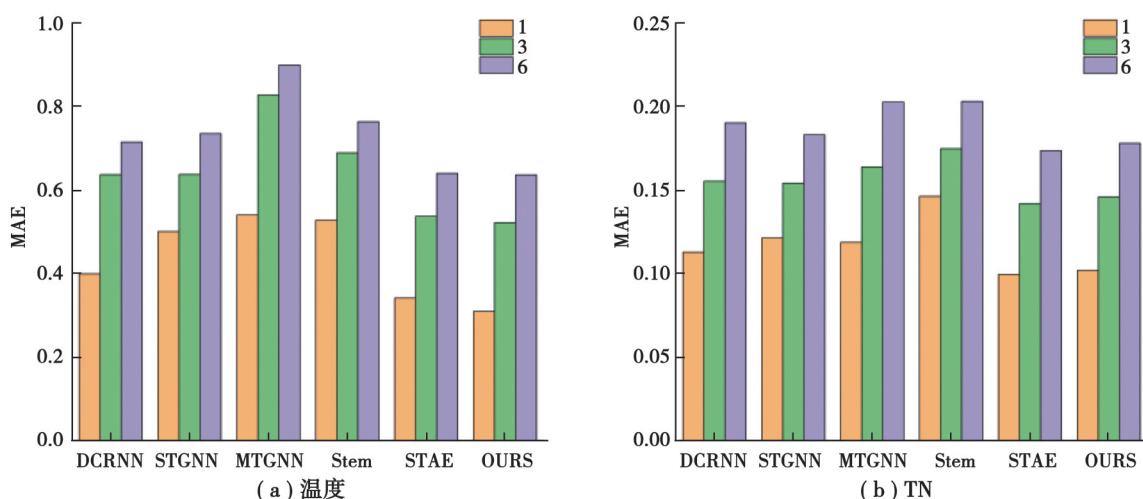


图 7 各方法对水质 TN 预测的效果

Fig. 7 TN prediction performance of different methods

为评估模型在不同预测步长下的预测性能以及模型对单个指标的偏好, 进行了模型的多步长预测实验, 结果如图 8 所示, 观察文中模型和其他模型在预测步长分别等于 1、3、6 条件下的 MAE 损失情况可以看出, 预测步长为 1 时, 模型对温度、电导率、溶解氧、pH 和高锰酸盐指数的预测性能最佳, STAEformer 在总氮预测中表现最优, 但随着预测步长增加到 3, STAEformer 的误差开始增加, 当预测步长增加到 6 时, 模型对温度、电导率、溶解氧、pH 和高锰酸盐指数的预测性能提高到最佳。

总之, 随着预测步长的增加, 所有模型的 6 项水质指标的 MAE 损失都在增加, 但模型对不同指标有所偏好。MTGNN 对温度的预测较其他模型更弱, 但对电导率的预测比 DCRNN 和 STGNN 强。STAEformer 步长短时效果较好, 说明具有较好的局部波动捕捉能力, 但随着预测步长的增加, 其误差增加幅度增大。模型性能随预测步长增加而出现的分化反映了模型对时空特性的适应差异。STAEformer 的损失变化说明了其局部波动捕获能力较强, 而模型的损失幅度随着步长增加而放缓, 说明模型具有较好的时空耦合建模效能。



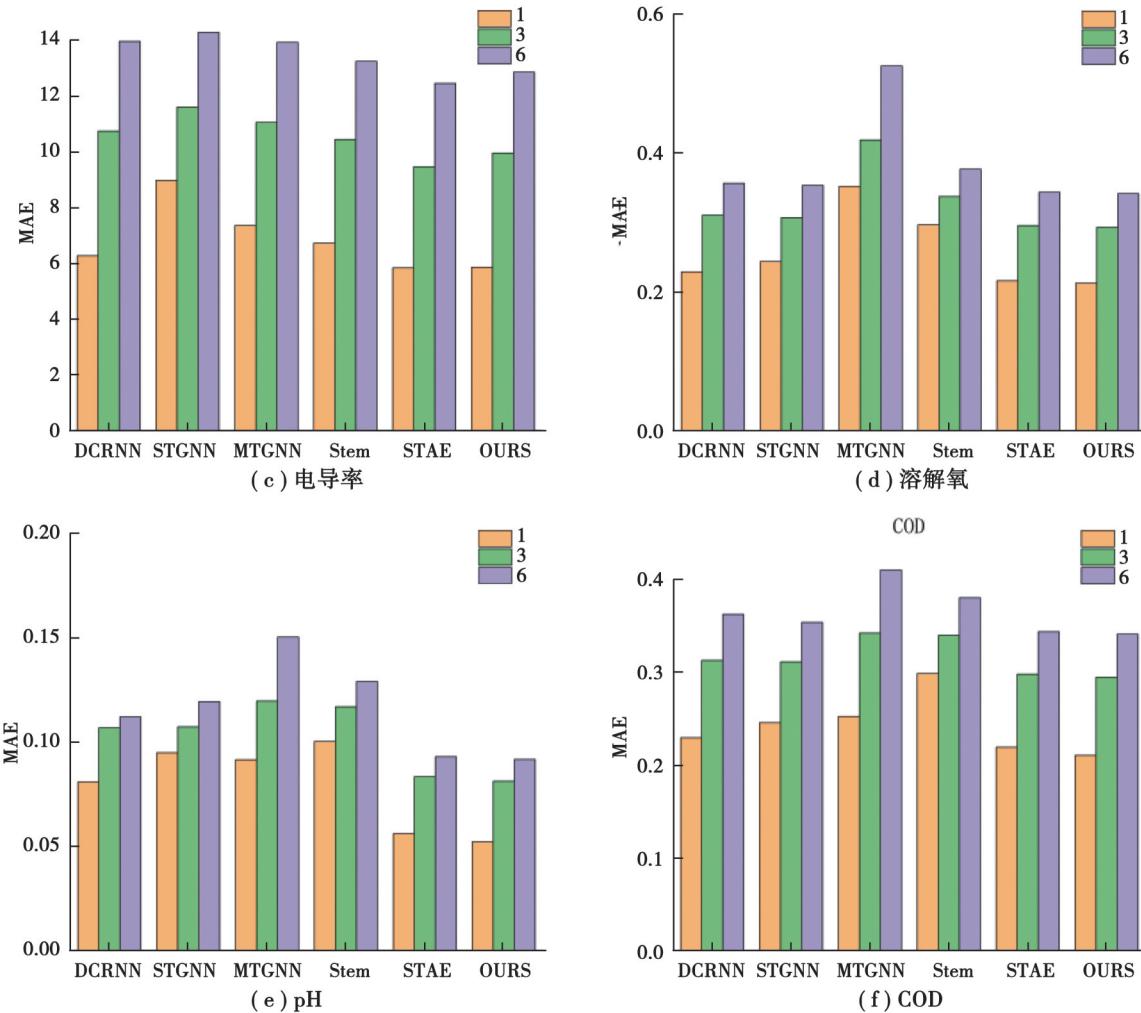


图8 不同预测步长下3种方法对各个指标的预测效果

Fig. 8 Prediction performance of three methods for each indicator at different forecast horizons

3.3.3 效率分析

为更全面地评估模型性能,在同一硬件环境(NVIDIA RTX 3090, 24 GB 显存; Intel i9-12900K)下对比了不同模型的模型损失、模型参数数量、训练及推理速度。文中模型与基线模型的效率指标对比如表3所示,其中,训练时间和推理时间分别表示每个Epoch模型的训练和推理所消耗的时间,MAE表示模型在测试集上的平均绝对误差损失。

表3 效率对比实验结果

Table 3 Experimental results of efficiency comparison

参数	DCRNN	STGNN	MTGNN	StemGNN	STAЕformer	OURS
MAE	1.223	1.701	1.688	1.352	1.134	1.124
参数数量	223 k	109 k	170 k	1 065 k	1 093 k	534 k
训练时间/s	7.37	2.64	1.35	3.13	4.44	6.78
推理时间/s	0.37	0.08	0.11	0.15	0.14	0.25

由表可知,文中模型的MAE为1.124,具有最优的预测精度。模型的MAE相较次优模型STAЕformer提升0.9%,参数量减少51.1%,可以大幅减少显存的占用;相较STGNN和MTGNN等参数数量较少的模型,预测精度分别提升了33.9%和33.4%,具有明显优势;相较DCRNN模型,在预测精度、训练和推理速度方面均有优势。但由于图结构需实时更新邻接矩阵,因此训练时间和推理时间高于STAЕformer和StemGNN等模

型。进一步的优化方向为动态图的稀疏化处理和注意力机制优化。由于该数据集的水质监测周期为 4 h, 依旧满足流域级水质预警需求。

3.3.4 消融实验

为了验证模型的各个主要组成部分的有效性, 进行了消融实验。设计了 3 种模型的变体, 主要包括: W/O LT, 表示移除了 MST; W/O GW, 表示移除了用于短周期时空融合中的 Graph WaveNet; W/O F, 表示在最后特征融合部分移除了长周期的 Graph WaveNet。所有变体以及原模型的实验结果如表 4 所示。

表 4 消融实验对比
Table 4 Ablation study comparison

类型	MAE	RMSE	$R^2/\%$
W/O LT	3.965	6.230	72.26
W/O F	1.733	3.732	81.83
W/O GW	1.504	3.477	85.62
Original	1.124	2.941	96.25

从时间建模方面看, 移除了长周期特征模块后 MAE 和 RMSE 分别增长了 54.18% 和 26.90%, R^2 降低了 14.98%, 所有指标的下降体现了长周期时间依赖的有效性; 移除了 MST 模块后 MAE 和 RMSE 分别增长了 253% 和 112%, R^2 降低了 24.92%, 性能下降幅度最大, 反映了 Transformer 在多尺度长短期时间模块中占据着主导地位。从空间建模方面看, 移除了 Graph WaveNet 后 MAE 和 RMSE 分别增长了 33.81% 和 18.23%, R^2 降低了 11.04%, 意味着图卷积网络模块在水质参数预测任务起到了至关重要的作用, 也证明了模型捕获空间依赖性的必要性。

3.3.5 超参数分析

为了进一步研究模型的预测性能, 重点分析了模型编码器隐藏层维度和掩码比例这 2 个关键超参数对模型预测性能的影响。其中, 编码器隐藏层维度为预训练阶段 Transformer 输出的维数, 是模型的一个重要的超参数。长期的水质数据具有时间连续性和低信息密度的特点, 掩码比例控制了预训练任务难度, 特定的掩码比例会影响模型对长期特征的提取能力。为了考虑模型的泛化性, 同时避免测试集信息泄漏, 实验使用测试集损失作为评估指标, 从 [16, 32, 64, 128] 中选择编码器隐藏层的最佳维度, 确定最佳维度后, 再从 [0.25, 0.5, 0.75, 0.9] 中选择最佳的掩码比例。参数扰动对预测性能影响的实验结果如图 9 所示。

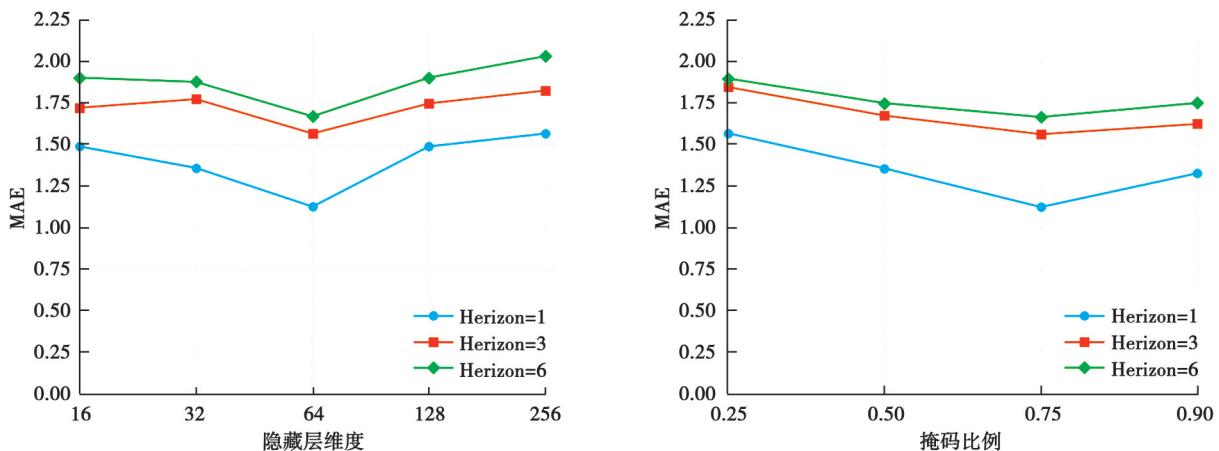


图 9 不同超参数下模型的预测损失
Fig. 9 Prediction loss of the model under different hyperparameters

可以看出,模型的隐藏层维度和掩码比例存在性能拐点。隐藏层维度为 64 且掩码比例为 75% 时模型在所有预测步长上均取得最优 MAE。短步长预测中,64 维相较 32 维和 128 维的损失分别降低 17.2% 和 24.3%;长步长预测中,损失降幅分别为 11.0% 和 12.2%。这是由于隐藏层维度过低会导致长期特征提取能力不足,过高易导致模型过拟合且增加模型参数量。掩码比例为 75% 时模型的预测能力最佳,短期预测误差相较 50% 和 90% 掩码比例分别降低 17.2% 和 15.3%;长期预测误差分别降低 4.8% 和 4.9%。过低的掩码比例下,模型会退化为短期特征提取器;过高的掩码比例下,上下文的缺少导致模型对长期特征提取能力减弱。

4 结 论

针对流域水质多参数预测中时空依赖复杂、传统模型难以有效融合动态拓扑与长期特征的问题,文中基于时空图卷积网络构建了一种融合自适应图结构与多尺度时序特征的水质预测模型。通过设计时间特征提取模块,引入 Transform 掩码重建任务,利用掩码数据增强模型对长期特征的提取能力,设计自适应图结构学习模块构建水网动态图结构,通过时空融合的方式提高模型的预测效果。实验结果表明,文中模型对水质参数动态变化有较好的捕捉和跟踪能力。

参考文献

- [1] 马嘉怿,李志威,陈晨,等. 长江中下游江心洲陆生生境健康评价方法及应用[J]. 水资源与水工程学报, 2024, 35(6): 1-9, 22.
Ma J Y, Li Z W, Chen C, et al. Health evaluation method for terrestrial habitats and its application to mid-channel bars in the middle and Lower Yangtze River[J]. Journal of Water Resources and Water Engineering, 2024, 35(6): 1-9, 22. (in Chinese)
- [2] 涂倩倩,付国楷,罗羽豪,等. 人工构建沉水植物净化系统的净化效能[J]. 重庆大学学报, 2022(S1): 79-84.
Tu Q Q, Fu G K, Luo Y H, et al. Purification efficiency of artificial submerged plant purification system[J]. Journal of Chongqing University, 2022(S1): 79-84. (in Chinese)
- [3] 邱勇,毕怀斌,田宇心,等. 污水处理厂进水数据特征识别与案例分析[J]. 环境科学学报, 2022, 42(4): 44-52.
Qiu Y, Bi H B, Tian Y X, et al. Characterization model of influent data for wastewater treatment plants and its application in case studies[J]. Acta Scientiae Circumstantiae, 2022, 42(4): 44-52. (in Chinese)
- [4] Ni Q J, Cao X H, Tan C Q, et al. An improved graph convolutional network with feature and temporal attention for multivariate water quality prediction[J]. Environmental Science and Pollution Research, 2023, 30(5): 11516-11529.
- [5] Agaj T, Budka A, Janicka E, et al. Using ARIMA and ETS models for forecasting water level changes for sustainable environmental management[J]. Scientific Reports, 2024, 14: 22444.
- [6] Du H Y, Zhao Z H, Xue H F. ARIMA-M: a new model for daily water consumption prediction based on the autoregressive integrated moving average model and the Markov chain error correction[J]. Water, 2020, 12(3): 760.
- [7] Wang J, Zhang L Y, Zhang W, et al. Reliable model of reservoir water quality prediction based on improved ARIMA method[J]. Environmental Engineering Science, 2019, 36(9): 1041-1048.
- [8] 马创,王尧,李林峰. 基于遗传算法与支持向量机的水质预测模型[J]. 重庆大学学报, 2021, 44(7): 108-114.
Ma C, Wang Y, Li L F. A water quality prediction model based on genetic algorithm and SVM[J]. Journal of Chongqing University, 2021, 44(7): 108-114. (in Chinese)
- [9] 巫莉莉,黄志宏,何斌斌,等. 基于随机森林的循环水养殖氨氮预测模型研究[J]. 重庆理工大学学报(自然科学版), 2021, 35(4): 207-214.
Wu L L, Huang Z H, He B B, et al. Study on the ammonia nitrogen predictive model of recirculating aquaculture based on random forest[J]. Journal of Chongqing University of Technology (Natural Science), 2021, 35(4): 207-214. (in Chinese)
- [10] Nasir N, Kansal A, Alshaltone O, et al. Water quality classification using machine learning algorithms[J]. Journal of Water Process Engineering, 2022, 48: 102920.
- [11] Zhang Y T, Li C L, Jiang Y Q, et al. Accurate prediction of water quality in urban drainage network with integrated EMD-LSTM model[J]. Journal of Cleaner Production, 2022, 354: 131724.

- [12] Wu X J, Zhang Q, Wen F, et al. A water quality prediction model based on multi-task deep learning: a case study of the Yellow River, China[J]. Water, 2022, 14(21): 3408.
- [13] Wang K F, Liu L, Ben X C, et al. Hybrid deep learning based prediction for water quality of plain watershed[J]. Environmental Research, 2024, 262: 119911.
- [14] Qiao J F, Lin Y Z, Bi J, et al. Attention-based spatiotemporal graph fusion convolution networks for water quality prediction[J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 1-10.
- [15] Du W J, Côté D, Liu Y. SAITS: Self-attention-based imputation for time series[J]. Expert Systems with Applications, 2023, 219: 119619.
- [16] Wu H X, Hu T G, Liu Y, et al. TimesNet: temporal 2D-variation modeling for general time series analysis[EB/OL]. 2022: arXiv: 2210.02186. <https://arxiv.org/abs/2210.02186>.
- [17] Tashiro Y, Song J M, Song Y, et al. CSDI: conditional score-based diffusion models for probabilistic time series imputation[EB/OL]. 2021: arXiv: 2107.03502. <https://arxiv.org/abs/2107.03502>.
- [18] Miao X Y, Wu Y Y, Wang J, et al. Generative semi-supervised learning for multivariate time series imputation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(10): 8983-8991.
- [19] Yoon J, Zame W R, van der Schaar M. Estimating missing data in temporal data streams using multi-directional recurrent neural networks[J]. IEEE Transactions on Biomedical Engineering, 2019, 66(5): 1477-1490.
- [20] Li Y G, Yu R, Shahabi C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting[EB/OL]. 2017: arXiv: 1707.01926. <https://arxiv.org/abs/1707.01926>.
- [21] Jin G Y, Liang Y X, Fang Y C, et al. Spatio-temporal graph neural networks for predictive learning in urban computing: a survey[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(10): 5388-5408.
- [22] Wu Z H, Pan S R, Long G D, et al. Connecting the dots: multivariate time series forecasting with graph neural networks[C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. July 6-10, 2020, Virtual Event, CA, USA. ACM, 2020: 753-763.
- [23] Cao D F, Wang Y J, Duan J Y, et al. Spectral temporal graph neural network for multivariate time-series forecasting[C]// Proceedings of the 34th International Conference on Neural Information Processing Systems. December 6-12, 2020, Vancouver, BC, Canada. ACM, 2020: 17766-17778.
- [24] Liu H C, Dong Z, Jiang R H, et al. Spatio-temporal adaptive embedding makes vanilla transformer SOTA for traffic forecasting [C]//Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. October 21-25, 2023, Birmingham, United Kingdom. ACM, 2023: 4125-4129.

(编辑 陈移峰)