J. Lake Sci.(湖泊科学),2026,38(6):000-000
DOI 10.18307/2026.0645
©2026 by *Journal of Lake Sciences*

JMD-Gformer 模型及其在大辽河多尺度径流预测中的应用*

许星晗[1]，苗星一[2]，胡磊[2]，刘建卫[1]**
（1：大连理工大学建设工程学部，大连 116000）
（2：大连理工大学控制科学与工程学院，大连 116000）

摘　要：在全球气候变化加剧的背景下，水文循环过程呈现显著的非平稳性、多尺度性和强非线性特征，传统机理模型难以深入揭示气象–径流响应机理，而深度学习方法在长期依赖刻画、空间关联建模和序列解耦方面仍存在不足。为此，本文提出一种融合跳跃–调幅调频模态分解（JMD）、稀疏有向图网络与 Transformer 的混合预测模型（JMD-Gformer）。该模型借助 JMD 将非平稳径流序列解耦为跳跃分量（如突发径流事件）和周期分量（如季节性和年际尺度波动），以抑制噪声干扰；进一步构建基于气象—水文节点的稀疏有向图结构，刻画上下游驱动关系；并引入多头自注意力机制，实现长程时序依赖建模。以大辽河流域为研究区，开展多时间尺度径流预测试验。结果表明，相比次优对比模型，JMD-Gformer 在主要评价指标（如 MAPE 和 RMSE）上分别下降 36.4%和 41.2%，在复杂水文条件下表现出良好的鲁棒性和预测精度，可为流域智慧水务管理提供新的技术途径。
关键词：跳跃–调幅调频模态分解（JMD）；稀疏有向图网络；Transformer；季节性特征；径流预测

JMD-Gformer Model and Its Application to Multi-scale Runoff Prediction in the Daliao River Basin

Xu Xinghan[1], Miao Xingyi[2], Hu Lei[2], Liu Jianwei[1]**
(1: *Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116000, P. R. China*)
(2: *College of Control Science and Engineering, Dalian University of Technology, Dalian 116000, P. R. China*)

Abstract: Global climate change is intensifying, and hydrological cycles present obvious non-stationary, multi-scale, and strong nonlinear characteristics. Traditional mechanistic models cannot fully characterize the response mechanisms between meteorological factors and runoff. Current deep learning methods also have limitations in capturing long-term dependencies, spatial correlations, and sequence decoupling features in hydrological data. To solve the above problems, this study develops a hybrid forecasting model named JMD-Gformer. The model integrates Jump plus AM-FM Mode Decomposition (JMD), sparse directed graph networks, and Transformer-based structures. In the proposed framework, JMD decomposes non-stationary runoff time series into jump components and periodic components. The jump components reflect abrupt runoff events, while the periodic components represent seasonal and inter-annual runoff fluctuations. This decomposition method effectively reduces data noise and mode aliasing. A sparse directed graph with meteorological and hydrological nodes is constructed to describe the upstream and downstream hydrological interactions. Furthermore, a multi-head self-attention mechanism is adopted to capture long-range temporal dependencies of the runoff series. Multi-time-scale runoff prediction experiments are carried out in the Daliao River Basin. Compared with the second-best benchmark model, the JMD-Gformer model decreases MAPE by 36.4% and RMSE by 41.2%. The model exhibits good robustness and high prediction accuracy under complex hydrological conditions. This study provides a feasible new method for intelligent watershed management.
Keywords: Jump-plus AM-FM mode decomposition (JMD); Sparse directed graph network; Transformer; Seasonal characteristics; Runoff prediction

河流径流量预测对于流域水资源管理、防洪安全和生态调控具有重要意义。随着全球气候变化加剧，水文循环过程呈现显著的非平稳、多尺度与强非线性特征，降水、气温、风速等气象因素对径流的影响日益突出[1]。我国多数流域降水季节性明显，旱雨季交替期间气象要素的复杂交互作用进一步增加了预测难度[2-3]。因此，在构建径流预测模型时，需综合考虑季节性变化、多源气象条件与地理环境因素，通过数据融合与多维建模提升预测的准确性与可靠性。

传统径流预测多依赖机理模型或统计模型，难以从数据驱动视角准确刻画复杂的气象与径流的响应关系[4]。近年来，深度学习模型如长短期记忆网络（long-short term memory, LSTM）、时序卷积网络（Temporal Convolutional Networks, TCN）[5]以及基于自注意力机制的 Transformer 及其改进模型（Informer、Autoformer）等[6-9]，在非线性特征提取和长序列建模方面取得显著进展。同时，物理信息神经网络（Physics Informed Neural Networks, PINN）通过将物理模型嵌入损失函数，实现了在少量数据下的有效学习，提供了“白箱”

* 2025-12-12 收稿;2026-04-10 收修改稿.
国家自然科学基金项目(52309079)资助.
** 通讯作者；E-mail: jwliu@dlut.edu.cn

结合“黑箱”混合建模的新方法[10]。例如，Nazari 等[11]将 PINN 应用于河道径流预测，通过贝叶斯推断增强参数估计的物理合理性。但由于流域径流序列的非平稳性和多尺度波动，这类模型在长期依赖和不规则水文空间关系建模上仍存在局限。针对空间异质性问题，基于图神经网络（Graph Convolutional Networks, GCN）增强时空建模能力的水文预测方法受到广泛关注。Li 等[12]提出双流 GCN 结构，通过显式建模河流上下游流动力学，结合掩码策略，实现了较高精度的径流预测；Bi 等[13]提出基于 Transformer 的时空图融合模型，利用时空注意力编码器与自适应动态邻接矩阵生成器，动态挖掘河流的时空特征。

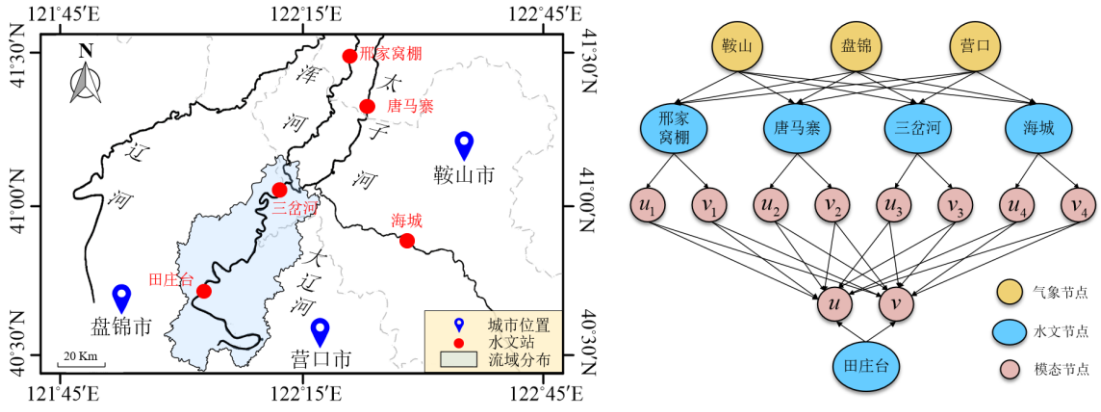
此外，天然水体环境复杂多变，水质与径流时序常受气象波动、人为干扰等噪声影响，直接对非平稳序列进行端到端学习易导致模态混叠[14-15]。为此，学者常使用时序分解策略，如变分模态分解（Variational mode decomposition,VMD），特征模态分解（Feature mode decomposition, FMD）等方法，以降低噪声干扰，从而提升预测精度与可解释性[16-18]。例如，Zhang 等[18]将 VMD-TCN 模型应用于渭河流域，缓解了径流非平稳性对预测精度的影响；Tang 等[19]基于自适应小波包分解和门控循环单元在黑河流域实现了优于 LSTM 的月径流预测；Chen 等[20]结合局部加权回归、VMD 与 Informer 提升了永定河流域的中长期径流预报能力。尽管分解-预测框架在缓解非平稳性方面取得一定进展，但现有方法对突发事件和跨尺度耦合过程的处理仍显不足，且与图网络和 Transformer 等先进时序模型的融合研究相对较少。值得注意的是，跳跃加调幅-调频模态分解（Jump plus AM-FM Mode Decomposition, JMD）近年来在光伏功率预测、机械故障诊断等领域显示了强大的非平稳信号处理能力[21-23]。JMD 通过在统一框架内实现对跳变（如阶跃、脉冲）与 AM-FM 振荡的协同分解，提供了处理复杂非平稳信号的新思路。然而，JMD 在水文领域的应用尚处于空白，特别是在处理水流径流序列中的突变事件与周期分量方面，亟待深入探究。

综合以上分析，目前径流预测研究仍存在以下不足：1）对非平稳、多尺度径流序列中跳跃和周期分量的解耦不足；2）气象—水文要素间的空间拓扑关系未有效利用；3）分解方法与图网络、Transformer 等深度学习框架的协同建模研究相对薄弱。为此，本文提出一种融合 JMD、稀疏有向图网络与 Transformer 的混合径流预测模型 JMD-Gformer，并以大辽河流域为研究区开展多时间尺度预测实验。本文的主要贡献包括：1）利用 JMD 将径流序列显式分解，解耦年际趋势、季节波动与降雨脉冲，显著降低建模复杂度与噪声干扰。2）构建上游驱动、下游信息增强的有向图结构，引入气象—水文节点耦合机制，精准刻画河流时空依赖关系。3）嵌入 Transformer 的自注意力机制，自适应建模长程依赖，实现高精度、高效率的径流预测。4）通过多时间尺度实验评估模型稳定性与适应性，全面考察其在不同时间跨度下的性能表现。

1 研究区域与数据

1.1 研究区域

本文选取辽河流域的大辽河段作为研究区。该河段位于我国东北温带季风气候区，年均气温约 8-10℃，年降水量 500-600mm，降水主要集中在 6-9 月（约占全年的 70%）。径流季节性显著，夏季受东南季风影响暴雨频发，冬季枯水期径流主要依赖地下水补给。上游洪水频发、径流变率大，中下游则面临水资源短缺。因此，分析并预测其径流演变对水资源调控、区域发展和生态保护至关重要。流域水系分布见图 1（a），大辽河由浑河与太子河汇合后注入渤海，干流全长 97.5 km，集水面积 1962 km²，河网拓扑结构如图 1（b）所示。



(a) 大辽河水系分布图　　(b) 河网拓扑结构图

图 1 大辽河段水系分布及拓扑结构图

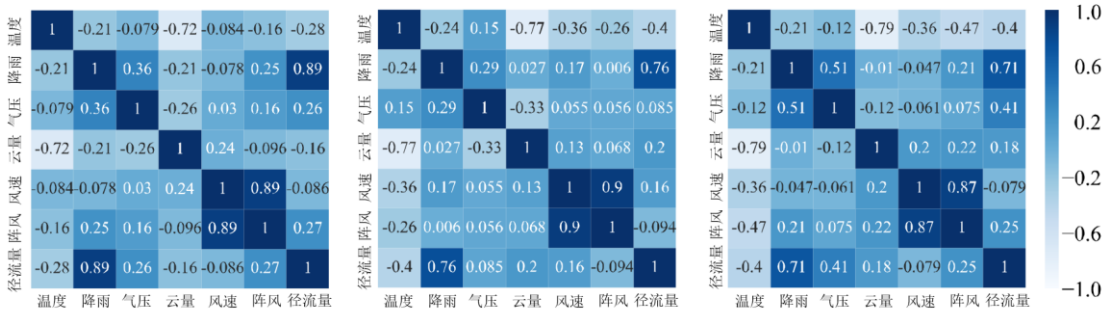
Fig. 1 Distribution of water systems in the Daliao River Section

所用气象与水文资料均针对大辽河流域，时段统一为 2007 年 1 月 1 日至 2017 年 12 月 31 日。气象数据为基于站点的观测数据，来源于世界气象数据平台（https://www.worldweatheronline.com/），涵盖鞍山、盘锦、营口三市每日 12:00 的温度、降水、云量、气压、风速及阵风共 6 项指标。水文数据采用 VIC-CN05.1[24] 所提供的空间分辨率为 0.25°×0.25° 的网格化径流数据。首先，依据研究区水系矢量数据与河网资料提取拓扑关系，并结合水文站点人工核验，形成图 1（b）的河网拓扑结构；随后，以流域内所有陆地网格的面积加权平均径流量作为日径流序列，将其视为流域出口站点（水文节点）的天然径流估计值。后续将完整数据按时间划分为三个独立时段：2007–2008 年（共 2 年）用于参数敏感性分析，2009–2015 年（共 7 年）作为训练集用于模型训练，2016–2017 年（共 2 年）作为测试集用于模型验证。

1.2 径流数据诊断分析

为评估径流序列的预测挑战，本小节对其非平稳性、趋势及异常值进行了系统分析。径流序列的非平稳性统计检验结果（补充材料附图 1）表明目标径流序列是一种具有波动趋势，突变异常点，尖峰厚尾分布的复杂非平稳序列。趋势结果（补充材料附图 2）显示年际尺度上径流量波动剧烈，丰枯水期差异显著，且波动强度在时间轴上呈现非恒定特征。补充材料附图 2（a）中红色趋势线两侧分布着明显的离群点，表明流域径流过程受到异常值的影响。补充材料附图 2（b）进一步揭示，年径流总量存在显著的年际差异。综上分析，非平稳性、异方差性及异常波动特征共同增加了后续预测模块对径流序列建模的难度。

为量化大辽河气象要素与径流的关系，本小节通过 Pearson 相关系数分析鞍山、盘锦和营口三市 6 项气象指标与出口径流的相关性。结果显示，多数气象要素（如气温、降水）与径流呈正相关，且变量间交互复杂。这要求模型具备强特征提取能力，以揭示多要素耦合模式，从而辅助径流高精度估计。



(a) 鞍山气象径流要素相关系数矩阵　(b) 盘锦气象径流要素相关系数矩阵　(c) 营口气象径流要素相关系数矩阵

图 2 气象径流要素相关系数矩阵

Fig. 2 Correlation coefficient matrix between meteorological and runoff elements

2. 研究方法

2.1 JMD-Gformer 混合预测模型

为解决大辽河流域复杂水力连接下多源气象—水文变量间时空耦合关系的建模挑战，提升多尺度径流预测的精度与鲁棒性，本节提出一种融合 JMD、稀疏有向图网络与 Transformer 的混合径流预测模型—JMD-Gformer，图 3 展示了 JMD-Gformer 混合预测模型的架构，其中 JMD 分解模块负责将输入的水文数据分解为周期和跳跃变量，图网络通过捕捉空间依赖关系对数据进行建模，Transformer 模块则通过自注意力机制建模时序依赖。最终，三个模块协同工作，实现对水文数据的准确预测。

2.2 JMD 分解方法

径流序列通常同时包含多尺度周期波动、由暴雨等事件引发的突发跳变成分以及高频噪声，呈现出显著的非平稳特征[20]。针对上述特性，本节引入 JMD 分解方法，通过优化联合目标函数，在统一框架内同步提取周期分量和跳跃分量，从而更好地处理具有季节性和突发事件的水文数据。与主要关注振荡模态分解的 VMD、FMD 等方法不同，JMD 在分解过程中刻画周期分量和跳跃分量的协同提取，有效降低水文信号在多模态间分散导致的模态混叠问题。同时，JMD 通过自适应参数调节机制增强了对非平稳水文序列的适应能力，可将径流数据分解为若干本征模态，实现对周期变化、趋势项及突发响应过程的有效分离。在非平稳性较强的情形下，JMD 相较于 VMD、FMD 等分解方法表现出更好的鲁棒性与分解精度。

2.2.1 分解变量定义　为降低原始径流序列 $f(t)$ 非平稳性对后续模型预测的干扰，JMD 将其分解如下。

$$f(t)=\sum_{n=1}^{N}u_n(t)+v(t)+n(t) \tag{1}$$

其中，$u_n(t)$ 表示第 n 个周期分量，表示多尺度周期规律；$v(t)$ 表示跳跃分量，用于刻画径流突变、突涨等

不连续突变情况； $n(t)$ 表示噪声项。

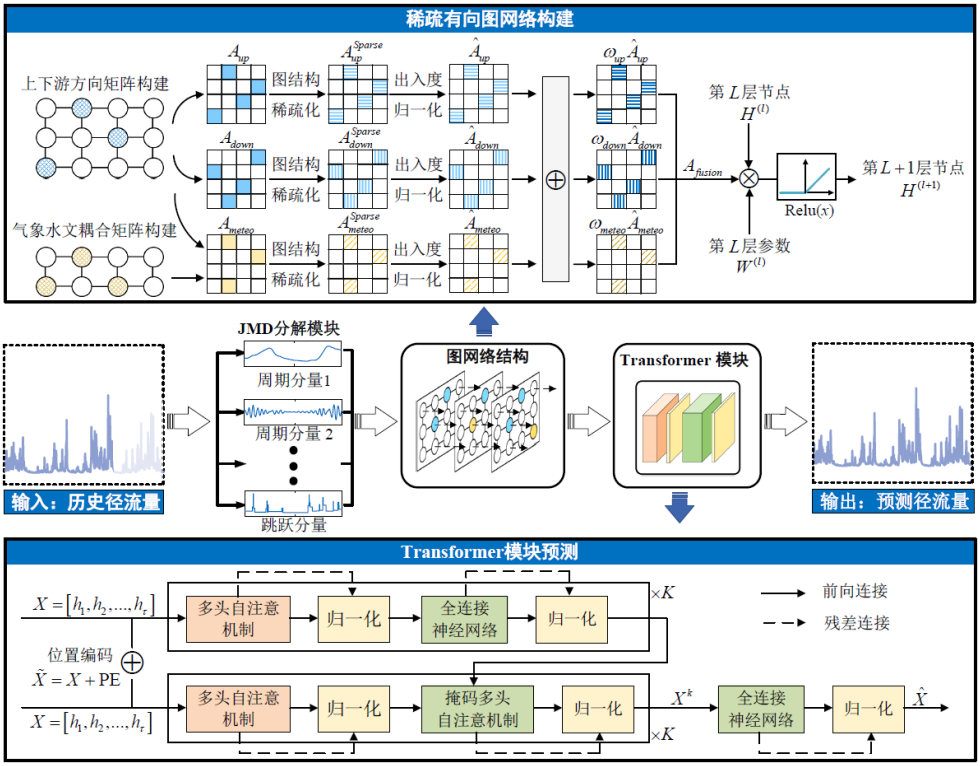


图 3 JMD-Gformer 混合预测模型整体架构示意图

Fig. 3 Diagram of the overall architecture of the hybrid prediction model of JMD-Gformer

2.2.2 联合优化目标　周期分量与跳跃分量的优化目标分别定义为 $J_1(\{u_n\})$ 与 $J_2(v)$，其具体表达式详见补充材料附件。综合上述分量目标，JMD 的联合优化问题可表示为如下形式：

$$J\left(\{u_n\},v\right)=\alpha J_1(\{u_n\})+\beta J_2(v)+\left\| f(t)-\left(\sum_{n=1}^{N}u_n(t)+v(t)\right)\right\|_2^2 \tag{2}$$

其中，α,β 是控制 J_1 与 J_2 贡献强度的正则化参数；$\|\cdot\|_2^2$ 表示重构误差项，约束分解后对原始序列的还原程度。N,α,β 的合理配置对分解效果至关重要，其敏感性分析详见第 4.2 节。综上，JMD 分解模块首先将输入的历史水文数据分解为多个模态成分，分别表征不同时间尺度下的周期变化及长期趋势。通过这一过程，模型能够有效地消除水文数据中的非平稳性，提取出更具规律性的特征，从而为后续的模型训练提供更清晰的输入。

2.3 稀疏有向图网络结构

径流过程具有显著的空间依赖性，在河网系统中主要体现为上下游之间的水力传导关系。传统图 GCN 通常基于无向图结构，默认节点间影响为对称关系，难以反映实际河道中“上游驱动下游”的单向水动力过程[12]。针对上述问题，本节构建了一种稀疏有向时空图网络，用于刻画河网结构下的空间依赖特征。该图网络属于时空图神经网络范畴，通过在图构建阶段引入水力方向性约束，并结合多源水文与气象节点的耦合建模，实现对流域空间依赖关系的更精细表达。在图卷积层中，模型以加权求和方式融合相邻节点特征，从而有效建模水文数据中的空间依赖特征，提升模型对复杂水文过程的建模能力。

2.3.1 节点与边的物理定义　图中包含三类节点与三类有向边：水文节点由 VIC 模型输出的流域径流序列构成，反映空间水文动态；气象节点对应鞍山、盘锦、营口等站点，提供气温、降水等驱动要素；模态节点源于 JMD 分解的跳跃与周期分量，表征多尺度物理过程。节点间通过空间边（基于河网拓扑的上游-下游连接，权重反映径流输移强度），模态边（基于相关系数的模态间统计关联）和耦合边（连接气象与水文节点，刻画气候-水文驱动关系）进行有向连接。

2.3.2 有向邻接矩阵构建　本小节定义两个方向性邻接矩阵，其中上游驱动矩阵 A_{up} 体现上游节点对下游节点的水动力影响，下游反馈矩阵 A_{down} 作为补充机制，捕获下游对上游的隐性反馈，计算式分别如下。

$$A_{up}=\left[A_{uu}\quad A_{uv}\quad A_{vv}\right],\quad A_{down}=A_{up}^{T} \tag{3}$$

其中，A_{uu},A_{uv},A_{uv} 分别表示由相关系数计算的周期分量间、周期-跳跃分量间、跳跃分量的连接矩阵，此外，权重系数 $\alpha_u,\alpha_{uv},\alpha_v$ 控制连接矩阵的贡献成分，分别设置为 $\{\alpha_u,\alpha_{uv},\alpha_v\}=\{0.85,0.2,0.9\}$。

同时考虑气象条件对径流变化的影响，引入气象节点与水文节点之间的耦合边。耦合权重由气象相似性和跨节点相关性共同决定，在计算这些指标前，所有原始气象数据（温度、降水、风速等）均进行了 Z-score 标准化处理，以消除量纲影响并突出各要素的相对波动。

$$A_{meteo}(i,j)=\mathrm{sig}\left(\frac{\mathrm{CosineSim}(M_i,M_j)+\mathrm{CrossCorr}(u_i,M_j)}{2}\right) \tag{4}$$

其中，CosineSim(·,·) 表示余弦相似度，用来衡量节点 i 和 j 的气象条件相似度，CrossCorr(·,·) 表示交叉相关系数，用以衡量分解后的水文数据与不同节点气象的相关性。 sig(·) 表示 sigmoid 函数。

2.3.3 稀疏化与归一化　为消除冗余连接，对有向邻接矩阵实施基于分位数的稀疏化，为提高训练稳定性，避免高入度造成的特征过度放大，对所有稀疏化后的矩阵进行出入度归一化，公式分别如下：

$$A_{ij}^{Sparse}=A_{ij}\mathbf{1}\quad(A_{ij}>q_\theta) \tag{5}$$

$$\hat{A}=D_{\mathrm{out}}^{-1/2}(A^{Sparse}+I)D_{\mathrm{in}}^{-1/2} \tag{6}$$

其中，A_{ij} 表示所有邻接矩阵，θ 为预设的稀疏化比例，q_θ 为对应的分位数阈值，4.2 节通过敏感性分析评估其最优取值。D_{in} 和 D_{out} 是对角矩阵，分别表示入度和出度。这种对称归一化处理可以缓解高连通节点对聚合过程的主导作用，保证信息在网络中的传播更加均匀。

2.3.4 多图融合与特征更新　为整合上游驱动、下游反馈与气象耦合三类空间关系，采用加权融合策略生成统一邻接矩阵，融合后通过 L 层图卷积更新节点特征，融合策略与特征更新如下。

$$A_{fusion}=\omega_{up}\hat{A}_{up}+\omega_{down}\hat{A}_{down}+\omega_{meteo}\hat{A}_{meteo} \tag{7}$$

$$H^{(l+1)}=\mathrm{Relu}(A_{fusion}H^{(l)}W^{(l)}) \tag{8}$$

其中，ω 为可学习权重，且满足 $\sum\omega=1$，A_{fusion} 为融合后的矩阵，$H^{(l+1)}$ 和 $H^{(l)}$ 为第 $l+1$ 层和第 l 层的隐藏状态，$W^{(l)}$ 是图网络第 l 层的可学习参数，Relu(·) 表示非线性激活函数。

2.4 Transformer 结构预测

为精确建模径流序列的长期依赖特性，本节选用 Transformer 作为时序预测模块。相较于 LSTM（存在梯度消失，难以捕捉长程关联）[4]和 TCN（感受野受卷积核尺寸限制，需堆叠多层以扩展覆盖范围）[5]，，Transformer 模块通过多头自注意力机制捕捉长时间序列数据中的时序依赖[25]，特别是在水文数据中，季节性变化和突发事件往往跨越较长的时间段。通过自注意力机制，模型能够自动关注时间序列中关键的时间步，从而提高预测的准确性和鲁棒性。给定图网络增强后的节点特征序列 $X=h_i,i=1,...,\tau$，Transformer 编码器的核心计算流程如下：

2.4.1 位置编码　通过正弦余弦函数为输入序列 X 注入顺序信息[25]，通过各时间步位置编码矩阵得到最终输入矩阵 $\tilde{X}\leftarrow X+\mathrm{PE}(1:\tau)$

2.4.2 多头自注意力机制　通过多个注意力头并行捕捉不同子空间的依赖关系，其计算式如下：

$$\mathrm{MultiHead}(x)=\mathrm{Concat}(\mathrm{Head}_1,...,\mathrm{Head}_n)W^o$$

$$\mathrm{Head}_i=\mathrm{Softmax}\left(\frac{\tilde{X}W_i^Q\cdot(\tilde{X}W_i^K)^T}{\sqrt{d_k}}\right)\tilde{X}W_i^V \tag{9}$$

其中，MultiHead(·) 表示多头注意力机制，Concat(·) 表示向量连接操作，Softmax(·) 是归一化激活函数，W_i^Q，W_i^K，W_i^Q 分别是第 i 个注意力头的查询，键，值映射矩阵，W^o 是权重系数，d_k 是键向量维度。

2.4.3 前馈神经网络与输出预测　本小节采用编码器加全连接输出层的设计，而非完整编码器-解码器结构。该简化适应径流预测的序列回归任务特性，避免自回归解码带来的误差累积，并与 JMD 分解所提取的平稳分量相匹配，在保证精度的同时提升计算效率。全连接网络的输出 $\hat{X}$ 可通过下式计算。

$$X^k=\mathrm{LN}\left(\mathrm{MultiHead}(X^{k-1})+X^{k-1}\right) \tag{10}$$

$$\hat{X} = \mathrm{LN}\left(X^k + \mathrm{FFN}(X^k)\right) \tag{11}$$

其中， $k = 1,...,K$ 表示编码器的堆叠层数， LN(·) 表示归一化层， FFN(·) 为全连接神经网络。

3 结果及分析

3.1 实验细节及模型参数设置

为贴合径流预报业务实际需求，设置涵盖短（3d、7d）、中（30d）、长期（90d、180d）的多预见期预测任务，其中预见期指模型向前预报的天数。特征因子与标签变量的时延关系通过滑动时间窗口机制定义，模型输入为历史时段 $t-T$ 至 t 的多源特征序列（含径流、气象要素），输出为未来时段 t 至 $t+S+1$ 的径流量预测值，其中 S 为预见期长度， T 为回溯窗口。不同预见期预测的关注点与实际意义各异，3 天预测捕捉强降雨等突发事件的瞬时响应；7 天预测反映连续天气过程的累积效应；30 天及以上预测则依赖季节性气候背景场，刻画年际水文规律。在应用层面，短期预测支撑防洪应急与实时调度，中期预报服务于农业灌溉与水资源调度，长期预测为流域长期规划提供依据。

本节将 JMD-Gformer 与七种经典预测模型进行比较以评估其预测性能，包括基于 Transformer 的 Autoformer 和 FEDformer，以及基于图网络的 DGCN 和 STGFT 等。所有基准模型的简介、结构参数与输入/输出维度详见补充材料附表 1 与附表 2。在数据预处理阶段，针对水文监测中常见的“审查数据”（即仪器在低于检出限或接近饱和限时输出的固定值），首先剔除非物理可信的审查数据并视其为缺失值，再采用 T-LSTM-VAE 方法[26]进行统一插补。该方法基于变分自编码器框架，能同时捕捉时间依赖与变量间相关性，具体流程见补充材料。插补后，对所有输入特征进行 Z-score 标准化，并统一应用于所有对比模型，以保障数据尺度一致。为确保对比实验的公平性与可复现性，所有模型均遵循相同的训练流程：数据集按 7:1:2 的比例划分训练集、验证集与测试集，并统一使用 Adam 优化器进行训练，批量大小设置为 32。训练过程中，若验证集损失连续 10 轮未改善，则触发早停机制；最大训练轮数设为 100 轮。对于已开源模型（如 Autoformer、FEDformer、Dlinear），严格遵循其原始论文的预测步骤与超参数设置；对于未公开代码的模型，则依据论文描述进行复现，并在同一超参数空间内进行网格搜索以确定最优配置。所提 JMD-Gformer 模型的关键超参数汇总于表 1。

为准确评估各模型的径流预测性能，选用三个指标用于评估误差：均方根误差（root mean squared error, RMSE），平均绝对百分比误差（mean absolute percentage error, MAPE）和决定系数（coefficient of determination, R^2）。评价指标的具体计算公式见补充材料。

表 1 JMD-Gformer 关键超参数值设置

Tab. 1 Key hyperparameter value Settings for JMD-Gformer

超参数设置	注意力头数	丢弃率	编码器层数	解码器层数	学习率	正则化参数
参数值	8	0.1	2	2	10^{-3}	10^{-4}

3.2 年度径流量趋势预测结果

为评估 JMD-Gformer 年度尺度径流量预报能力，以 3 天预见期对比所提模型与 7 种基线模型。采用 RMSE、MAPE、R^2 及平均运行时间作为评价指标，结果汇总于表 2；图 4 通过散点图展示预测值与真实值的散点分布，密度分布直观反映预测精度，点云贴近对角线表明模型精度越高。

表 2 不同模型在年度性径流量预测的综合性能对比

Tab. 2 Comparison of comprehensive performance of different models in annual runoff prediction

对比模型	MAPE	RMSE	R^2	运行时间（秒）
DCRNN	0.2165	0.4348	0.8830	113.85
Autoformer	0.2074	0.4074	0.8112	152.21
FEDformer	0.1813	0.3419	0.8433	246.32
PatchTST	0.1555	0.3293	0.8775	67.66
Dlinear	0.1422	0.3037	0.9082	**22.80**
DGCN	0.1118	0.2085	0.9546	61.04
STGFT	0.1246	0.2392	0.9382	102.33
JMD-Gformer	**0.0626**	**0.1241**	**0.9854**	53.62

如表 2 所示，所提模型在各项精度指标上均表现优异。具体而言，相较于次优模型 DGCN，MAPE 降低 44%，RMSE 降低 40.48%。径流过程受多源气象要素与空间拓扑共同影响，传统图模型难以有效建模气象-径流的交互机制，而 Transformer 类模型直接处理原始非平稳序列，训练难度大且易受噪声干扰。为

此，本研究引入 JMD 作为前置处理器，从而在结构上分离年际趋势、季节波动与突发降雨等物理过程。模型随后针对这些规律性更强、物理意义更清晰的模态分量，结合图网络学习其空间传播机制。该设计降低了学习难度，提升了可解释性与泛化能力，使模型在 MAPE、RMSE 与 R^2 上取得领先。在计算效率方面，模型运行时间为 53.62 秒，低于 Autoformer、FEDformer 等方法，实现了精度与效率的协同优化。图 4 结果显示，所提模型的预测偏差相对更低，散点沿对角线集中分布。相比之下，Dlinear 与 DGCN 在高值区间出现明显离群点，表明其对非线性强波动时序的适应能力有限；DCRNN 与 Autoformer 的预测点云较为分散，整体偏差较大；STGFT 的散点集中度亦不及本模型。综上，JMD-Gformer 能够有效提取气象-径流耦合特征，并对序列波动趋势保持稳健跟踪能力，从而实现了更精准、稳定的径流预测。

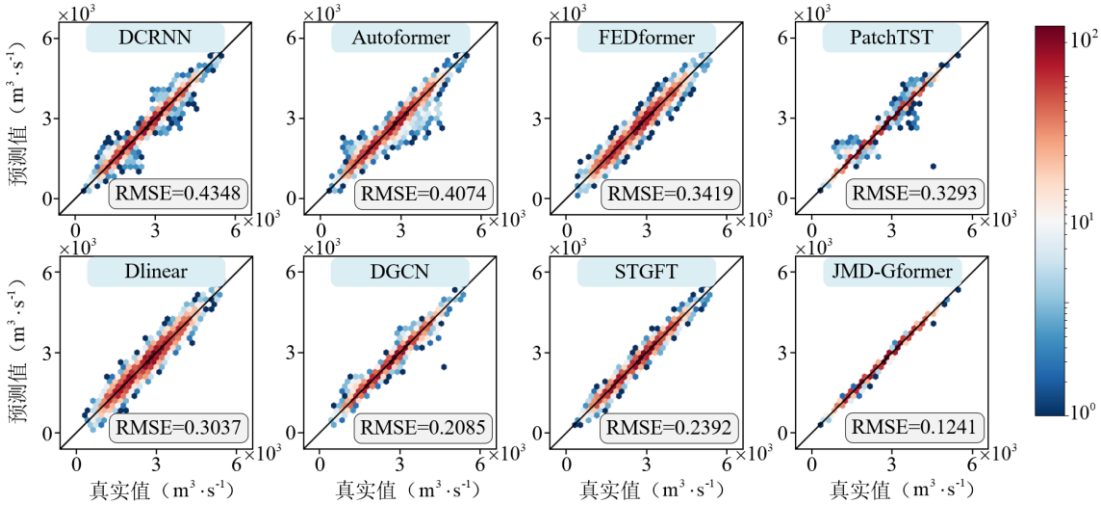


图 4 不同模型预测值与真实值的散点分布与拟合效果

Fig. 4 Scatter distribution and fitting effect of predicted and true values of different models

3.3 不同预见期径流量趋势预测结果

为系统评估不同模型在多预见期下的预测性能，构建涵盖 7 天至 180 天的多时间尺度预报框架。表 3 对比各模型在不同预见期下的性能，图 5 直观展示所提模型的预测结果及误差分布特征。

表 3 不同预见期下径流量预测性能综合对比

Tab.3 Comprehensive comparison of runoff prediction performance under different prediction step sizes

对比模型	评价指标	不同预见期长度				平均运行时间（秒）
		7	30	90	180	
DCRNN	MAPE	0.2275	0.2256	0.2588	0.2661	112.44
	RMSE	0.4437	0.5093	0.5365	0.5938	
	R^2	0.7626	0.7568	0.7016	0.6054	
Autoformer	MAPE	0.2002	0.2117	0.2141	0.2379	154.16
	RMSE	0.4093	0.4386	0.4633	0.5176	
	R^2	0.8074	0.7845	0.7632	0.6816	
FEDformer	MAPE	0.1819	0.1942	0.2017	0.2642	259.60
	RMSE	0.3920	0.4257	0.4385	0.5317	
	R^2	0.8423	0.8195	0.8002	0.7024	
PatchTST	MAPE	0.1675	0.1820	0.2045	0.2393	68.25
	RMSE	0.3587	0.3869	0.4175	0.4623	
	R^2	0.8643	0.8255	0.7936	0.7491	
Dlinear	MAPE	0.1437	0.1493	0.1625	0.1960	23.82
	RMSE	0.2991	0.3235	0.3394	0.4405	
	R^2	0.9125	0.9044	0.8643	0.8004	
DGCN	MAPE	0.1164	0.1490	0.1522	0.1748	66.93
	RMSE	0.2677	0.3093	0.3356	0.3932	
	R^2	0.9435	0.9058	0.8847	0.8366	
STGFT	MAPE	0.1352	0.1575	0.1780	0.2009	127.32
	RMSE	0.2304	0.3136	0.3611	0.4481	
	R^2	0.9251	0.8984	0.8453	0.7802	
JMD-Gformer	MAPE	**0.0681**	**0.0772**	**0.0943**	**0.1084**	55.29
	RMSE	**0.1256**	**0.1362**	**0.1683**	**0.2023**	
	R^2	**0.9817**	**0.9746**	**0.9521**	**0.9495**	

在多预见期预测任务中（7 天至 180 天），所有模型的预测误差（MAPE 与 RMSE）均随预见期增加呈逐步上升趋势，体现出时序预测中普遍存在的时间衰减效应。其中，JMD-Gformer 在全年尺度预测（180 天）中表现稳健，其 MAPE 仅为 0.1084，低于其它对比模型，表明该模型在捕捉长期依赖关系方面具备明显优势。值得注意的是，JMD-Gformer 的误差随时间步长增加的幅度也最为平缓，表明该模型在不同时间尺度上均能保持较好的预测一致性。基于 Transformer 的模型（如 Autoformer、FEDformer、PatchTST）在短期与中期预测中表现尚可，但在长期预测任务中误差上升较为明显，这可能与注意力机制对远距离依赖关系的捕捉能力受限有关。相比之下，DGCN 通过显式建模流域空间拓扑结构，增强了对水文传导过程的表征能力，因此在长期预测中表现出优于传统 Transformer 模型的稳定性。而 DCRNN 作为时序图网络，其整体性能落后于 DGCN，说明简单的图卷积与时序网络组合在建模复杂非平稳径流过程方面仍有局限。综合多步长预测结果（表 3，图 5），JMD-Gformer 在周、月尺度均表现稳定，仅在 180 天预见期下误差略有波动，表明所提模型具有更强的适应能力和更低的预测偏差。

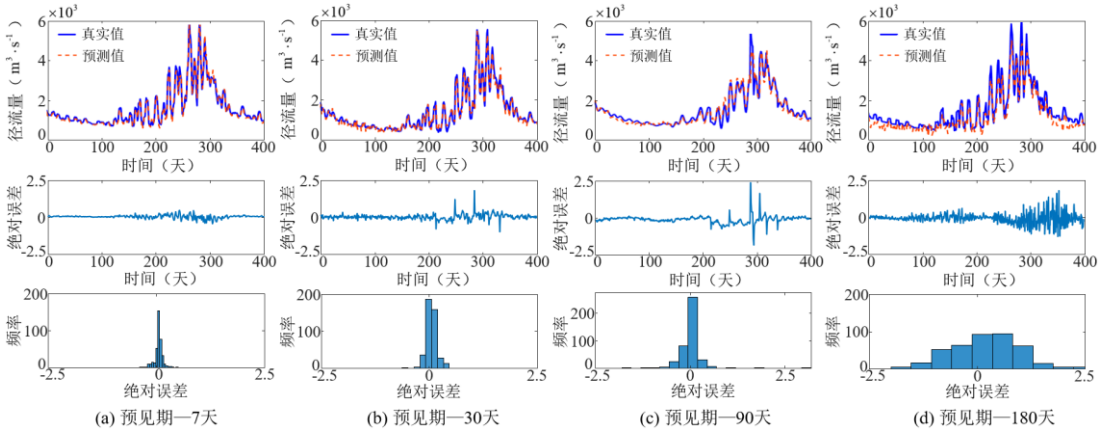


图 5 不同预见期下所提模型的预测结果和误差分布

Fig. 5 Prediction results and error distribution of the proposed model with different prediction step sizes

3.4 季节性径流量趋势预测结果

为分析所提模型的季节敏感性，本小节设置 3 天预见期，采用完整训练集进行模型训练，依据气象季节特征将测试集划分为四个典型气候季度[30]：冬季（12 月-2 月）、春季（3 月-5 月）、夏季（6 月-8 月）和秋季（9 月-11 月），分别评估各季度内的预测结果与误差分布特征。表 4 列出各季度的起止时间及径流量统计特征（径流量范围、均值、标准差）；表 5 汇总模型在各季度的评价指标；图 6 展示四个季度下的径流量分布及预测误差分布情况。

表 4 四个季度的径流量统计特征（$10^3 m^3 \cdot s^{-1}$）

Tab. 4 Statistical characteristics of runoff volume of the four quarters ($10^3 m^3 \cdot s^{-1}$)

季度划分	开始日期	终止日期	径流量范围	径流量均值	径流量标准差
第一季度（冬季）	2016-12-01	2017-02-28	0.876-1.367	1.044	0.183
第二季度（春季）	2017-03-01	2017-05-31	0.645-1.907	1.102	0.840
第三季度（夏季）	2017-06-01	2017-08-31	0.673-6.422	2.465	1.454
第四季度（秋季）	2017-09-01	2017-11-30	0.989-5.696	1.993	0.426

根据表 4、表 5 与图 6 的综合分析，所提模型总体上表现出良好的泛化能力与季节适应性，各季度测试集的 R^2 均保持在 0.9 以上，说明其具备稳定的预测性能，但精度仍存在季节性差异。模型在第四季度（秋季）的表现最佳（MAPE=0.0548, RMSE=0.0637），预测曲线与实际观测值高度吻合。该优势可能源于秋季径流过程较为平稳，水文条件相对稳定，加之模型对时序特征捕捉能力较强。第一季度（冬季）次之（MAPE=0.0951, RMSE=0.1369）。结合统计特征可知，秋、冬两季径流标准差较低（分别为 0.426 和 0.183），数据波动较小，有利于模型学习规律性变化。模型在第三季度（夏季）表现中等（MAPE=0.0530, RMSE=0.1508）。该季节以暴雨径流为主，流量峰值显著，均值和标准差分别为 2.465 和 1.454，数据波动剧烈，导致误差有所上升。相比之下，模型在第二季度（春季）的预测波动最为明显，出现跟踪延迟现象，且部分时段误差显著增大。这可能由于春季受降雨与融雪径流共同影响，流量过程复杂多变，导致训练数

据分布偏移，影响了预测一致性。

尽管存在季节性波动，各季度预测结果整体仍处于可控范围内—平均 MAPE 为 0.0779，RMSE 为 0.1509，R^2 为 0.9607—表明模型具备良好的综合预测能力，在不同季节条件下均具有可靠的实用价值。

表 5 所提模型在四个季度中的评价指标结果

Tab. 5 Evaluation index results of the proposed model in four quarters

季度划分	MAPE	RMSE	R^2
第一季度（冬季）	0.0951	0.1369	0.9768
第二季度（春季）	0.1085	0.2521	0.9116
第三季度（夏季）	0.0530	0.1508	0.9691
第四季度（秋季）	0.0548	0.0637	0.9853
平均值	0.0779	0.1509	0.9607

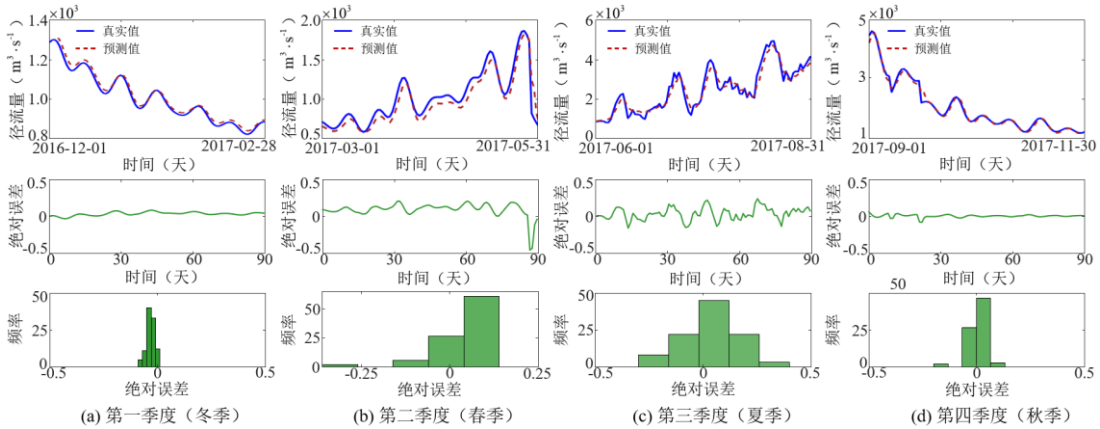


图 6 不同季度的径流量预测值及误差分布

Fig. 6 Runoff prediction values and error distribution in different quarters

4. 讨论

4.1 消融实验

为定量评估各组件对模型性能的贡献，本小节在 3 天预见期下系统开展消融实验。实验分为两部分：第一部分通过删减关键模块，定量分析对比完整模型、删减 JMD 分解模块及删减图网络结构三种配置的预测精度与计算效率差异；第二部分为明确 JMD 分解的量化重要性，保持其余架构不变，分别前置 VMD、FMD 两种分解方法，系统评估其对预测结果的影响。评价从预测精度（RMSE、MAPE、R^2）和计算效率（平均运行时间）两个维度展开，结果如表 6 和图 7 所示。

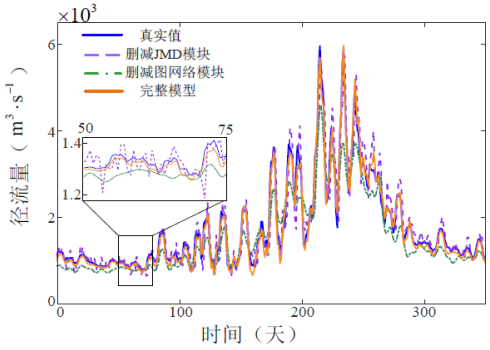
表 6 删减模块以及不同前置分解方法与完整模型的预测性能对比

Tab. 6 Comparison of prediction performance of the deleted modules and different predecomposition methods with the completed model.

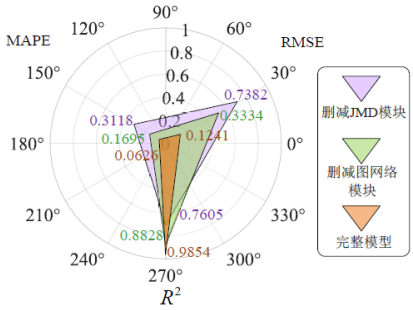
模型变体	MAPE	RMSE	R^2	运行时间（秒）
删减 JMD 模块	0.2118	0.6482	0.7605	**46.31**
删减图网络模块	0.1692	0.3334	0.8828	52.29
完整模型	**0.0626**	**0.1241**	**0.9854**	53.62
前置 VMD 模块	0.2278	0.5760	0.8012	**49.85**
前置 FMD 模块	0.1837	0.4315	0.8370	84.73
前置 JMD 模块	**0.0626**	**0.1241**	**0.9854**	53.62

如表 6 所示，引入 JMD 分解的完整模型在预测精度方面表现较好，MAPE 为 0.0626，RMSE 为 0.1241，R^2 也明显提升（达到 0.9854）。与删减 JMD 分解模块和删减图网络模块的模型相比，完整模型在 MAPE、RMSE 等指标上均有所降低，反映出 JMD 分解所提供的时序特征与图网络所提取的空间依赖关系对模型性能具有积极作用。在三种前置分解方法中，JMD 在各项精度指标上均优于 VMD 和 FMD，且其运行时间较低，在精度与效率之间取得了较好的平衡。相比之下，FMD 虽然精度高于 VMD，但计算耗时增加 58.1%。从可视化结果来看，如图 7（a）所示，在径流量波动阶段（50 至 75 时段），完整模型的预测曲线

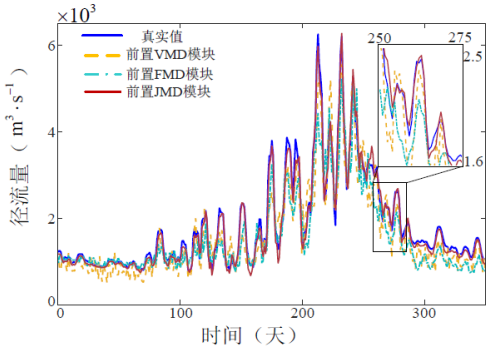
与实测值更为接近，而删减 JMD 模块后，预测轨迹出现在峰值和谷值等关键特征点出现显著偏差，删减图网络结构虽能大致跟踪趋势，但预测误差范围有所扩大。如图 7（c）所示，面对剧烈突变的径流变化（250 至 275 时段），前置 JMD 模块的预测结果与实测序列贴合更紧密，能较好地捕捉快速波动和趋势变化，而前置 VMD、FMD 模块则出现一定的相位滞后与幅值低估现象。图 7（b）、（d）综合显示：完整模型在所有评价维度上均优于删减 JMD 分解模块或图网络模块后的模型，且前置 JMD 分解的模型整体性能领先于前置 VMD、FMD 等分解模块的模型。



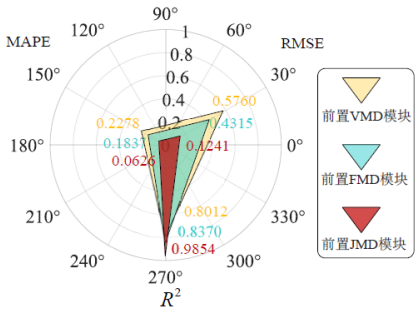
(a) 各模块对完整模型预测性能的贡献度分析



(b) 模块消融实验下的模型性能指标对比



(c) 不同前置分解方法与JMD方法的预测效果对比



(d) 前置分解方法对模型性能影响的雷达图评估

图 7 模型性能评估与消融实验对比分析

Fig. 7 Comparative analysis of model performance evaluation and ablation experiments

4.2 参数敏感性分析

JMD 分解方法的关键超参数中，周期分量数量 N 控制分解强度：N 过大易导致过分解，N 过小则引发欠分解与模态混叠；正则化参数 α 过大会使周期分量的频带过度收缩而丢失有效信息，过小则导致频带重叠与模态混合；正则化参数 β 设置不当将影响跳跃分量提取的准确性，为系统评估 N 、α 、β 和网络稀疏化程度 θ 对预测精度与计算效率的影响，进行如下参数敏感性分析。参数敏感性分析实验在每次测试中仅调整一个参数，其余参数保持默认值。各参数的具体范围设定如下：模态数 $N \in \{2, 3, 4, 5, 6, 7\}$，周期分量正则化参数 $\alpha \in \{1000, 3000, 5000, 7000, 9000, 11000\}$，跳跃分量数 $\beta \in \{0.005, 0.010, 0.015, 0.020, 0.025, 0.030\}$，图网络稀疏化程度 $\theta \in \{10\%, 20\%, 30\%, 40\%, 50\%, 60\%\}$。通过对比不同参数组合下的预测精度与效率，最终确定一组在效率与精度之间取得最优平衡的参数配置为 $\{N, \alpha, \beta, \theta\} = \{3, 5000, 0.020, 30\%\}$。参数敏感性分析结果见图 8 和补充材料附表 3。

4.3 JMD 分解的水文机理贡献分析

为验证 JMD 方法在径流预测中的水文物理合理性，本小节从水文过程角度分析了 JMD 所提取跳跃分量与实际暴雨洪水事件之间的时间一致性。以检验其对极端降水引发快速径流响应的识别能力。

选取研究期内典型的暴雨洪水过程，对比分析原始径流序列与 JMD 分解结果（降雨事件与径流突增时间见补充材料表 4）。图 9（b）显示，在多个时段（如第 50、110、325 天附近）阶段，跳跃分量（紫色曲线）与降雨事件（浅蓝色峰值）同步出现显著尖峰，且其发生时刻与径流响应（黑色峰值）高度吻合，而周期分量在这些阶段未呈现明显突变。这一同步性表明跳跃分量有效捕捉了极端气象事件驱动的快速径流响应。此外，无降雨或小降雨量时期（125-150 天时段）跳跃分量值接近零，径流序列仍保持平滑基流波动，说明 JMD 未将缓变基流误判为跳变。

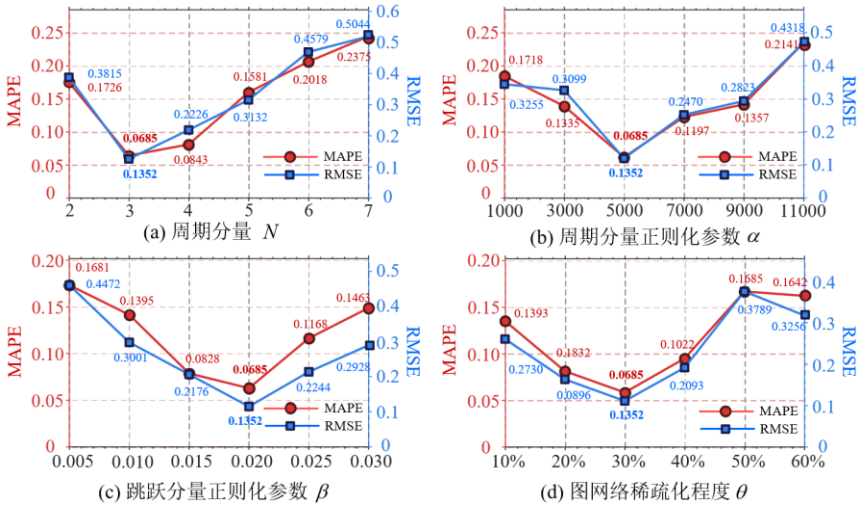


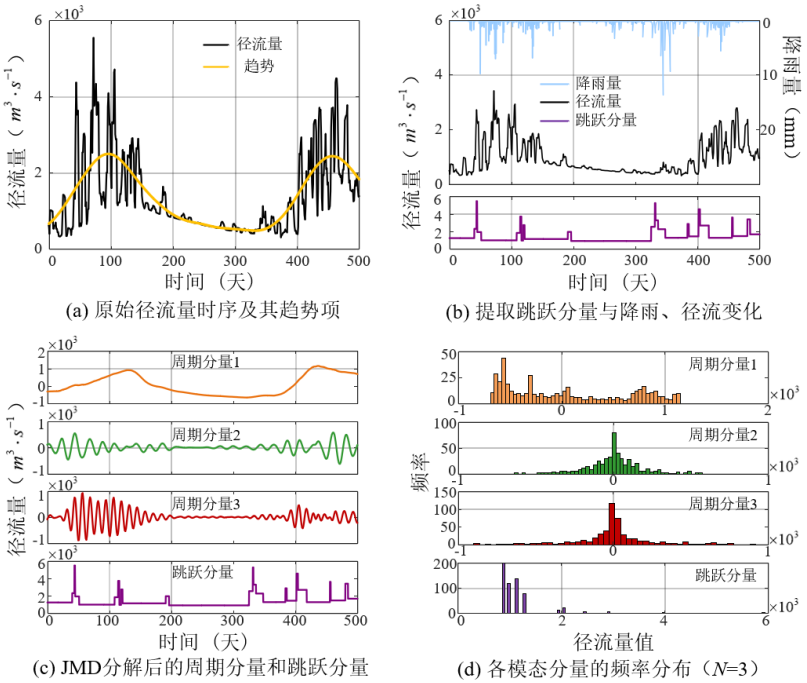
图 8 JMD-Gformer 参数敏感性分析

Fig. 8 Sensitivity analysis of the parameters of the JMD-Gformer

从水文机理看，JMD 能区分径流序列中的不同物理过程：跳跃分量捕捉如暴雨等突发事件引发的快速响应，而周期分量则刻画季节或年际尺度上的缓慢演变。这种物理意义明确的分解，使预测模型能分别学习各类水文过程的规律，在保持整体稳定性的同时，增强对极端径流事件的捕捉能力。该方法适用于具有显著突变特征和周期性波动的径流序列。此外，其迁移潜力在于，JMD 提取的分量对应普适的水文物理过程，通过针对目标流域特性调整分解参数（如 N,α,β ），可适配不同地理区位的河网系统。

4.4 时间复杂度分析

为评估模型的计算复杂度，根据不同预见期的平均 RMSE、平均推理时间（秒）和峰值内存占用（GB）三个方面，将所提模型与 7 种基线方法进行对比。时间复杂度结果（补充材料图 3）表明，扩散卷积模型 DCRNN 内存占用较低，但推理速度较慢且误差较高；基于 Transformer 的改进模型精度有所提升，但整体计算成本较高。相比而言，JMD-Gformer 在精度与效率之间取得了较好的平衡。这主要得益于 JMD 分解对非平稳信号的预处理，降低了 Transformer 的建模复杂度。与同样依赖图网络的 DGCN 和 STGFT 相比，JMD-Gformer 通过对跳跃-周期分量的显式解耦，提高了模型的预测精度。



(a) 原始径流量时序及其趋势项　(b) 提取跳跃分量与降雨、径流变化

(c) JMD分解后的周期分量和跳跃分量　(d) 各模态分量的频率分布（N=3）

图 9 径流序列分解结果与跳跃分量分析

Fig. 9 Decomposition results and jump component analysis of runoff series

5 结论

本文针对气候变化背景下径流序列的非平稳性、多尺度特征与复杂空间依赖等挑战，提出了一种融合 JMD、稀疏有向图网络与 Transformer 的混合预测框架 JMD-Gformer。该模型通过“分解—空间—时序”协同建模策略，实现了对径流演变过程中不同时间尺度与空间驱动机制的联合刻画。

研究结果表明，JMD 能够将径流序列中的突发响应过程与周期性变化过程进行显式分离，为后续模型提供物理意义更加清晰的输入特征；基于上游驱动关系构建的稀疏有向图结构有效刻画了河网水力传导及气象—水文耦合特征；在此基础上，Transformer 对解耦模态的长程依赖建模进一步提升了模型对复杂时序特征的表征能力。在大辽河流域的多尺度预测实验中，JMD-Gformer 在年度及多步长预测任务中均表现出较为稳定的误差控制能力，体现了良好的预测精度与鲁棒性。

值得指出的是，所提 JMD-Gformer 框架在方法机制上具备与实时径流预报场景相兼容的潜力。JMD 分解仅依赖当前时刻之前的径流观测序列，可在滑动时间窗内对实时更新的数据进行动态分解，并将提取的多尺度特征输入预测模型，从而形成“实时分解—滚动预测”的预报流程。尽管实验主要基于离线方法验证模型有效性，该框架在理论上支持向实时预报应用的拓展。同时需要说明的是，本文模型的验证主要基于单一流域数据，其空间记忆结构及水文响应模式在不同气候条件、下垫面特征或河网形态下的外推能力仍有待进一步检验。此外，部分特殊水文过程（如积雪消融，下游潮汐影响等）尚未被显式纳入建模框架。上述问题为后续在多流域泛化验证及复杂水文过程耦合建模提供了进一步研究空间。

6 附件

补充材料附件、附图及附表见电子版（DOI 10.18307/2026.0645）。

7 参考文献

[1]. Wang L, Zhai WL, Zhang JH *et al*. Impact of Future Climate Change on Runoff in the Upper Hanjiang River Based on RBF-SWAT. *Journal of Yangtze River Scientific Research Institute*, 2023, **40**(04): 31-36. [王立, 翟文亮, 张爵宏等. 基于 RBF-SWAT 的气候变化下汉江上游流域径流预测及特征分析. 长江科学院院报, 2023, **40**(04): 31-36.]

[2]. Xu XH, Hu L, Liu JW *et al*. Data-Driven Solution for Water Resource Management of Intelligent Water Conservancy. *IEEE Sensors Journal*. DOI: 10.1109/JSEN.2025.3625228.

[3]. Wang BW, Li JZ, Feng P. Impacts of land use change on runoff in Panjiakou Reservoir watershed. *Journal of Hydraulic Engineering*, 2018, **49**(03): 379-386. DOI: 10.13243/j.cnki.slxb.20170962. [王博威, 李建柱, 冯平. 土地利用变化对潘家口水库控制流域径流影响. 水利学报, 2018, **49**(03): 379-386. DOI: 10.13243/j.cnki.slxb.20170962.]

[4]. Ye KJ, Liang ZM, Chen HY *et al*. Research on regional long-short term memory neural network (LSTM) flood forecasting model.

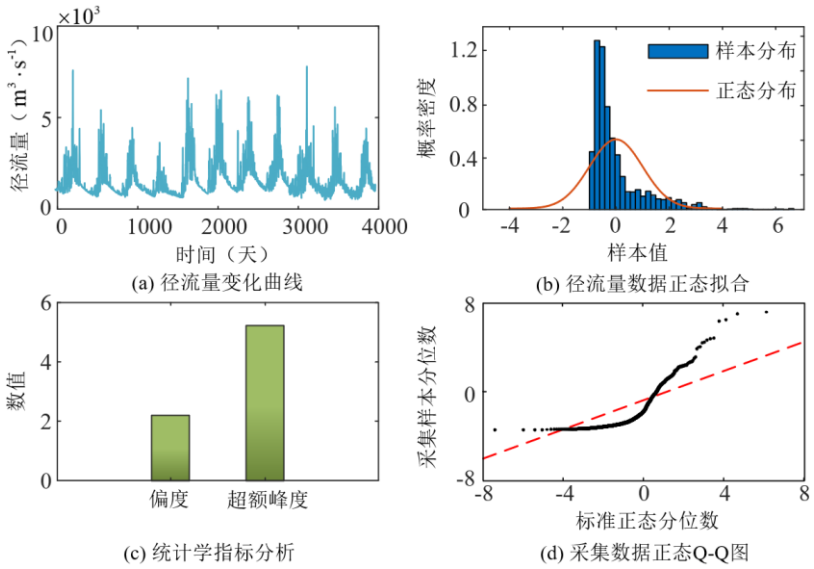
Journal of Lake Sciences, 2025, **37**(02): 651-659. [叶可佳, 梁忠民, 陈红雨等. 区域化长短期记忆神经网络(LSTM)洪水预报模型研究. 湖泊科学, 2025, **37**(02): 651-659.]
[5]. Wang J, Gao ZX, Shan CY. Multivariate Yellow River Runoff Forecast Based on TCN-Attention Model. *Yellow River*, 2022, **44**(11): 20-25. [王军, 高梓勋, 单春意. 基于 TCN-Attention 模型的多变量黄河径流量预测. 人民黄河, 2022 , **44**(11): 20-25.]
[6]. Zhou HY, Zhang SH, Peng JQ *et al*. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI conference on artificial intelligence*, 2021, **35**(12): 11106-11115. DOI: 10.1609/aaai.v35i12.17325.
[7]. Wu HX, Xu JH, Wang JM *et al*. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in neural information processing systems*, 2021, **34**: 22419-22430.
[8]. Zhou T, Ma ZQ, Wen QS *et al*. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. *International conference on machine learning. PMLR*, 2022: 27268-27286.
[9]. Xu XH, Hu L, Miao XY *et al*. Imputation of continuous missing values in water quality data using a temporal embedding-based self-attention variational autoencoder. *Journal of Hydrology*, 2026, **668**: 134937. DOI:10.1016/j.jhydrol.2026.134937.
[10]. Nazari, Luis Fernando, Eduardo Camponogara *et al*. Physics-informed neural networks for modeling water flows in a river channel. *IEEE Transactions on Artificial Intelligence*, 2022, **5**(3): 1001-1015.
[11]. Ye, JW, Zeng W, Nhu Cuong Do et al. Reconstructing transient pressures in pipe networks from local observations by using physics-informed neural networks. *Water Research*, 2024: 121648. DOI: 10.1016/j.watres.2024.121648.
[12]. Li ZL, Wang YN, Liu HX *et al*. Dual-stream graph convolutional networks enable accurate real-time chlorine prediction in sparsely monitored water distribution networks. *Water Research*, 2025: 124824. DOI: 10.1016/j.watres.2025.124824.
[13]. Bi J, Wang ZQ, Yuan HT *et al*. Long-term Water Quality Prediction with Transformer-based Spatial-Temporal Graph Fusion. *IEEE Transactions on Automation Science and Engineering*, 2025, **22**: 11392-11404. DOI: 10.1109/TASE.2025.3535415.
[14]. Qin R, Zhang XQ. Comparison on the application performance of LSTM flood forecasting model under different input method. *Journal of Lake Sciences*, 2025, **37**(04): 1470-1481. [覃睿, 张小琴. 不同输入设置对 LSTM 洪水预报模型应用效果的影响. 湖泊科学, 2025, **37**(04): 1470-1481.]
[15]. Zhu GW, Qin BQ, Zhang YL *et al*. Fluctuation of phosphorus concentration in Lake Taihu in the past 70 years and future control strategy. Journal of Lake Sciences, 2021, **33**(04): 957-973. [朱广伟, 秦伯强, 张运林等. 近 70 年来太湖水体磷浓度变化特征及未来控制策略. 湖泊科学, 2021, **33**(04): 957-973.]
[16]. Dragomiretskiy K, Zosso D. Variational mode decomposition. *IEEE transactions on signal processing*, 2013, **62**(3): 531-544.
[17]. Miao YH, Zhang BY, Li CH *et al*. Feature mode decomposition: New decomposition theory for rotating machinery fault diagnosis. *IEEE Transactions on Industrial Electronics*, 2022, **70**(2): 1949-1960.
[18]. Zhang SY, Luo JG, Shi GD *et al*. Monthly Runoff Prediction in Weihe River Basin Based on VMD-TCN Model. *Yellow River*, 2023, **45**(10): 25-29. [张上要, 罗军刚, 石国栋等. 基于 VMD-TCN 模型的渭河流域月径流量预测研究. 人民黄河, 2023, **45**(10): 25-29.]
[19]. Tang MZ, Yang YK, Zhang JW. Monthly runoff prediction model based on ASWPD-BO-GRU. *Journal of Water Resources and Water Engineering*, 2023, **34**(04): 84-91. [唐铭泽, 杨银科, 张菁雯. 基于 ASWPD-BO-GRU 的月径流量预测模型. 水资源与水工程学报, 2023, **34**(04): 84-91.]
[20]. Chen XZ, Wang ZJ, Liu D *et al*. A two-stage dynamic decomposition-prediction-reconstruction model for medium-long term runoff forecasting. *Journal of Hydraulic Engineering*, 2025, **56**(07): 933-944. [陈小泽, 王忠静, 刘丹等. 径流序列两段动态分解-预测-重构中长期预报模型. 水利学报, 2025, **56(**07): 933-944.]
[21]. Nazari M, Korshøj A R, ur Rehman N. Jump plus AM-FM mode decomposition. *IEEE Transactions on Signal Processing*, 2025. DOI: 10.1109/TSP.2025.3535822.
[22]. Cui XW, Yu XY, Gao YB *et al*. Ultra-short-term Photovoltaic Power Point-Interval Prediction Based on JMD-ITD Decomposition and Mamba2-MTransformer Parallel Architecture. *Smart Power*, 2025,**53**(12):68-77. [崔曦文, 于霄宇, 高一波等. 基于 JMD-ITD 分解与 Mamba2-MTransformer 并行的超短期光伏功率点-区间预测. 智慧电力, 2025, **53**(12): 68-77.]
[23]. Xie FY, Song CJ, Song MY *et al*, Rolling bearing life prediction based on adaptive JMD decomposition and improved TimesNet network. J*ournal of Railway Science and Engineering*, 2025, 1-14. [谢锋云, 宋成杰, 宋明桦等. 基于自适应 JMD 分解和改进 TimesNet 网络的滚动轴承寿命预测. 铁道科学与工程学报, 2025, 1-14. DOI: 10.19713/j.cnki.43-1423/u.T20250769.]
[24]. Miao Y, Wang AH. Evaluation of routed-runoff from land surface models and reanalyses using observed streamflow in Chinese river basins. *Journal of Meteorological Research*, 2020, **34**(1): 73-87. DOI: 10.1007/s13351-020-9120-z.
[25]. Han K, Wang YH, Chen HT *et al*. A survey on vision transformer. *IEEE transactions on pattern analysis and machine intelligence*, 2022, **45**(1): 87-110. DOI: 10.1109/TPAMI.2022.3152247.
[26]. Xu XH, Hu L, Liu JW *et al*. T-LSTM-VAE: A random missing data imputation model for multivariate time series. *Journal of the Franklin Institute*, 2025, **363**: 108231. DOI: 10.1016/j.jfranklin.2025.108231.

JMD-Gformer 模型及其在大辽河多尺度径流预测中的应用

（DOI 10.18307/2026.0645）附件材料

1. 径流序列非平稳性分析

对原始径流序列进行系统诊断分析，表明其具有显著的非平稳。各项统计检验结果如下：1）单位根检验（ADF 检验：h=1, p=0.001）拒绝“存在单位根”的原假设，而平稳性检验（KPSS 检验：h=1, p=0.01）拒绝“序列平稳”的原假设。二者结合表明，该径流序列具有典型的趋势平稳特性，即序列本身非平稳，但可通过一阶差分转化为均值平稳序列。2）Mann-Kendall 趋势检验确认序列存在波动趋势（Z=-4.431, p <0.001），印证上述径流序列非平稳的判断。3）Pettitt 突变点检测识别出显著的全局突变点，表明序列均值曾发生跳跃式变化，可能与极端水文事件或降雨变化有关。4）分布特征（附图 1（c））显示，序列具有明显的右偏（偏度=2.168）与尖峰厚尾特性（超额峰度=5.209）。样本与标准正态分布的 Q-Q 图（附图 1（d））显示，其右端明显偏离参考线，说明极端高径流事件的发生概率高于正态分布预期，与附图 2（a）中观测到的剧烈波动和离群点相符。

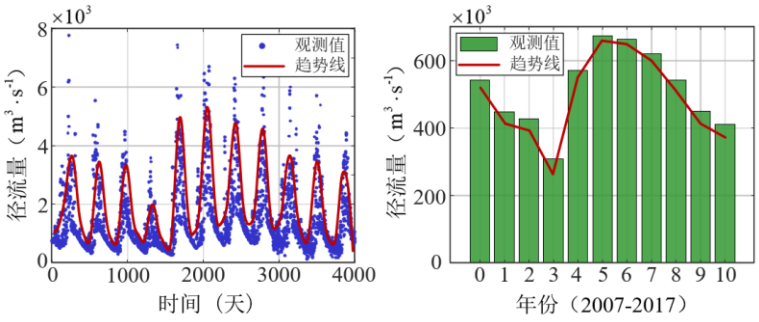


(a) 径流量变化曲线　(b) 径流量数据正态拟合

(c) 统计学指标分析　(d) 采集数据正态Q-Q图

附图 1　径流序列非平稳性诊断分析

App. Fig. 1 Diagnostic analysis of nonstationarity of runoff series



(a) 径流趋势分析　(b) 年径流总量变化

附图 2　径流序列趋势诊断分析

App. Fig. 2 Trend diagnostic analysis of runoff series

综上分析，本文面向的径流序列是一种具有波动趋势、突变异常点以及显著尖峰厚尾分布的复杂非平稳序列。尽管一阶差分可将其转化为平稳序列，但差分操作会同时消除序列中的确定性趋势和全局突变结构，而这些正是水文物理过程的关键表征。因此，本文选择采用 JMD 分解方法对其进行显式解耦。JMD 能够主动将序列分解为捕获长期趋势与周期波动的分量以及专门刻画突发事件的跳跃分量，从而在保留全部物理信息的前提下，有效降低原始非平稳序列的建模难度，为后续图网络结构结合 Transformer 的预测模块提供规律性更强、物理意义明确的输入序列。

2. 周期分量与跳跃分量优化目标

2.1 周期分量提取

为获取窄带 AM-FM 成分，JMD 对振荡模态的解析信号施加带宽约束，为确保得到的模态具有尽可能窄的瞬时频率带宽，构造如下周期分量带宽最小化目标：

$$J_1=\sum_N\left\|\partial_t\left(\tilde{u}_n(t)e^{-j\omega_k t}\right)\right\|_2^2 \tag{12}$$

其中，∂_t 表示对时间的偏导，令 $\tilde{u}_n(t)=u_n(t)+j\mathrm{H}u_n(t)$ 表示信号解析处理，H 表示希尔伯特变换，ω_k 为信号中心频率，$\|\cdot\|_2^2$ 表示 L_2 范数以惩罚带宽，使得 $u_n(t)$ 表现平滑瞬时频率变化的窄带 AM-FM 成分。

2.2 跳跃分量提取

径流突涨、突降通常表现为信号中的尖锐跳变。为保留这些水文过程的重要特征，JMD 采用重参数化和重缩放的凹惩罚函数对跳变项的导数施加稀疏约束。它旨在准确地保持分段常数信号分量的幅度，同时改善跳跃分量中尖锐不连续的表示，跳变项的优化目标和惩罚函数定义如下：

$$J_2=\int_0^\infty\phi\left(\left|\partial_t v(t)\right|;b\right)dt,\quad \phi(x;b)=\begin{cases}-\dfrac{b}{2}x^2+\sqrt{2b}x & x\in\left[0,\sqrt{2/b}\right)\\ 1 & x\in\left[\sqrt{2/b},\infty\right)\end{cases} \tag{13}$$

其中，参数 b 决定惩罚项的非凸程度，进而控制跳跃分量的稀疏性，当 $b\to 0$ 时，$\phi(x;b)\to|x|$，趋近于 L_1 正则，保留更多细节，$b\to\infty$ 时，$\phi(x;b)$ 逼近 L_0 正则，强化跳变稀疏性。该目标促进 $v(t)$ 形成少量突变点与较长的分段常数结构，从而匹配径流信号的物理特性。

3. 对比模型选取及参数设置

本节选取 7 个基准模型开展系统对比，涵盖 Transformer 系列与图结构系列方法，概要如下：DCRNN[1]使用扩散卷积来模拟信息（如污染物）在图结构上随时间的传播过程，再耦合门控循环单元（gated recurrent unit，GRU）捕捉时间动态。Autoformer[2]引入序列分解架构，将时间序列自动分解为趋势项和周期项，配合自相关机制聚合相似子周期。FEDformer[3]结合序列分解思想，通过在频域进行滤波和混合以更低的计算成本捕获全局依赖。PatchTST[4]通过划分时间补丁进一步捕捉局部语义模式，压缩序列长度，提升长序列建模效率。Dlinear[5]将序列分解为趋势项和残差项，分别用线性层建模再合并，兼顾可解释性与计算速度。DGCN[6]引入双流图卷积架构，并行刻画上下游水流动态，强化空间依赖捕捉。STGFT[7]集成时空注意力编码器与多图融合层，联合捕获河流网络的潜在空间关联与时序演化特征。

附表 2 列出了 JMD-Gformer 及 7 个基准模型的核心架构参数。针对多变量水文时间序列预测任务，这些参数主要基于各模型的原始论文及通用实现。对于已公开源代码的模型（如 Autoformer、FEDformer、Dlinear），严格遵循源论文中的模型结构和超参数设置；对于未公开代码的模型，依据论文公式和相关参数进行复现，并在必要时对关键参数进行网格搜索优化。所提模型与对比模型的结构设置及输入/输出维度详如下，其中特征因子与标签变量的时延关系通过滑动时间窗口机制定义，模型输入为历史时段 $t-T$ 至 t 的多源特征序列（含径流、气象要素），输出为未来时段 t 至 $t+S+1$ 的径流量预测值。B 表示批次大小，所有模型批次统一设置为 32，T 为回溯窗口长度，F 是特征数量，S 代表预测步长，D 指图网络的节点数，C 表示输出径流量维度。

附表 1 基准模型汇总

App. Tab. 1 Summary of baseline models

序号	对比模型	模型缩写	提出年份
1	Diffusion Convolutional Recurrent Neural Network	DCRNN	2018
2	Decomposition Transformer with Auto-Correlation	Autoformer	2021
3	Frequency Enhanced Decomposition Transformer	FEDformer	2022
4	Patch time series Transformer	PatchTST	2023
5	Decomposition-Linear	Dlinear	2023
6	Dual-stream Graph Convolutional Network	DGCN	2025
7	Spatial-Temporal Graph Fusion Transformer	STGFT	2025

附表 2 模型参数设置

App. Tab. 2 Model parameter settings

模型	模型定义	编码器/解码器结构	输入/输出维度
DCRNN	扩散卷积+ 门控循环单元 (GRU)	编码器：2 层扩散卷积 GRU 解码器：2 层扩散卷积 GRU GRU 神经元数：64	扩散步长：2，双向随机游走 输入:(*B*=32, *T*=128, *D*=8, *F*=6) 输出:(*B*=32, *S*, *C*=1)
Autoformer	序列分解+ 自相关机制+ Transformer 框架	编码器/解码器：各 2 层 自相关机制头数：8 前馈层神经元数：512 移动平均核：25	输入: (*B*=32, *T*=64, *F*=21) 输出: (*B*=32, *S*, *C*=1)
FEDformer	傅里叶 增强分解+ Transformer 框架	编码器/解码器：各 2 层 自注意力头数：8 前馈层神经元数：512 傅里叶变换随机选择：64 维	输入: (*B*=32, T=64, *F*=21) 输出: (*B*=32, *S*, *C*=1)
PatchTST	时序 Patch+ Transformer	编码器：3 层 自注意力头数：8 前馈层神经元数：512 Patch 长度：16	输入: (*B*=32, T=64, *F*=21) 输出: (*B*=32, *S*, *C*=1)
Dlinear	序列分解+ 线性层	线形层：1 个分解层+2 个线性层 前馈层神经元数：512	输入: (*B*=32, *T*=128, *F*=21) 输出: (*B*=32, *S*, *C*=1)
DGCN	双流图卷积网络	双流图卷积层：3 层 掩码层：随机掩蔽 50%图网络节点 前馈网络层神经元数：512	输入: (*B*=32, *T*=64, *D*=8, *F*=6) 输出: (*B*=32, *S*, *C*=1)
STGFT	时空图融合 Transformer	时空注意力编码器：1 层 注意力头数：2 自适应动态邻接矩阵生成器：1 层 图融合层：1 层 前馈网络层神经元数：512	输入: (*B*=32, *T*=64, *D*=8, *F*=6) 输出: (*B*=32, *S*, *C*=1)
JMD-Gformer	JMD 分解+ 图注意力+ Transformer 框架	JMD 分解层：分解周期数 *N*=3 图网络层：2 层，神经元数=256， 编码器/解码器：各 2 层 自注意力机制头数：8 前馈层神经元数：512	输入: (*B*=32, *T*=64, *D*=8, *F*=6) 输出: (*B*=32, *S*, *C*=1)

4. 模型评估指标

选用三个指标用于评估预测误差：均方根误差（root mean squared error, RMSE），平均绝对百分比误差（mean absolute percentage error, MAPE）和决定系数（coefficient of determination, R^2），公式定义分别如下。

$$\mathrm{MAPE}=\frac{1}{M}\sum_{i=1}^{M}\left|\frac{\hat{y}_i-y_i}{y_i}\right|\times 100\% \tag{14}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\hat{y}_i-y_i\right)^2} \tag{15}$$

$$R^2=1-\frac{\sum_{i=1}^{M}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{M}\left(y_i-\bar{y}_i\right)^2} \tag{16}$$

其中，M 为测试集样本总数，$\hat{y}_i$ 为预测径流量值，y_i 是真实径流量值，$\bar{y}_i$ 是真实平均径流量值。

5. T-LSTM-VAE 插补方法

T-LSTM-VAE 的整体工作流程如下，主要包含四个步骤[8]：

步骤 1：基模型预训练：利用完整源域数据（历史完整序列）训练 LSTM-VAE 基模型，通过编码器-解码器结构与变分推断学习通用时间模式，其目标为最大化证据下界（The evidence lower bound, ELBO）。

步骤 2：目标数据初始化与迁移：对含缺失值的目标数据，基于邻域加权均值进行缺失值初始化，保留局部时序平稳性；随后加载基模型，冻结浅层参数以保持通用特征提取能力，微调深层网络以适配目标域分布，并通过最大均值差异（MMD）对齐源域与目标域特征分布。其中整体损失函数为 $\mathrm{L_{total}}$，LSTM-VAE 基模型的损失函数为 $\mathrm{L_{LSTM\text{-}VAE}}$，最大均值差异为 MMD(·)。

$$\begin{aligned}\mathrm{L_{total}}&=\mathrm{L_{LSTM\text{-}VAE}}+\gamma\cdot\mathrm{MMD}^2\left(X_{Source},\bar{X}_{Target}\right)\\&=\left[\beta D_{KL}\left(q_\phi(z|\hat{x})\,\|\,p_\theta(z)\right)-\mathrm{E}_{q_\phi(z|\hat{x})}\log p_\theta(\hat{x}'|z)\right]+\gamma\left\|\frac{1}{T_1}\sum_{i=1}^{T_1}\phi(x_i)-\frac{1}{T_2}\sum_{j=1}^{T_2}\phi(\tilde{x}_j)\right\|^2\end{aligned} \tag{17}$$

步骤 3：超参数优化：采用贝叶斯优化自动搜索关键超参数（如学习率、正则化参数等）的最优组合。

步骤 4：数据插补：将预处理后的目标序列输入优化模型，通过解码器从隐分布采样重构缺失值，生成符合数据分布规律的插补结果。

6. 参数敏感性分析

为系统评估 N、α、β 和网络稀疏化程度 θ 对预测精度与计算效率的影响，本节进行如下参数敏感性分析。通过对比不同参数组合下的预测精度与效率，最终确定一组在效率与精度之间取得最优平衡的参数配置为$\{N,\alpha,\beta,\theta\}$={3, 5000, 0.020, 30%}。参数敏感性分析具体结果如附表 3 所示。

附表 3 不同参数值敏感性分析结果

App.Tab. 3 Results of sensitivity analysis for different parameter values

名称	周期分量数 N					
设定值	2	3	4	5	6	7
MAPE	0.1726	**0.0685**	0.0843	0.1581	0.2018	0.2375
RMSE	0.3815	**0.1352**	0.2226	0.3132	0.4579	0.5044
运行时间（秒）	**57.31**	58.03	60.81	64.78	63.64	67.86
名称	分解后周期分量的带宽 α					
设定值	1000	3000	5000	7000	9000	11000
MAPE	0.1718	0.1335	**0.0685**	0.1197	0.1357	0.2141
RMSE	0.3255	0.3099	**0.1352**	0.2470	0.2823	0.4318
运行时间（秒）	59.54	58.58	**58.03**	58.56	59.97	59.32

名称	提取跳跃分量的正则化参数 β					
设定值	0.005	0.010	0.015	0.020	0.025	0.030
MAPE	0.1681	0.1395	0.0828	**0.0685**	0.1168	0.1463
RMSE	0.4472	0.3001	0.2176	**0.1352**	0.2244	0.2928
运行时间（秒）	59.65	59.83	60.21	**58.03**	63.29	65.96
名称	图网络稀疏化程度 θ					
设定值	10%	20%	30%	40%	50%	60%
MAPE	0.1393	0.0896	**0.0685**	0.1022	0.1685	0.1642
RMSE	0.2730	0.1832	**0.1352**	0.2093	0.3789	0.3256
运行时间（秒）	65.92	60.54	58.03	56.25	57.46	**50.92**

7. JMD 分解的水文机理贡献分析

为验证 JMD 分解在水文预测中的物理合理性，本节选取了 2007 年至 2017 年间 8 次典型的降雨径流突变过程，定量分析了 JMD 所提取的跳跃分量与实际降雨、径流响应之间的时序关系。各事件的降雨、径流及跳跃分量的峰值时间、归一化后的强度及时序滞后详见附表 4。其中，时序滞后定义为跳跃分量峰值时间与降雨峰值时间之差，正值表明跳跃分量的出现滞后于降雨，符合“降雨-汇流-响应”的水文过程物理规律。

附表 4 显示，所有事件的时序滞后均为正值，平均滞后约 1.4d，与该流域的汇流时间尺度一致，表明跳跃分量并非随机噪声，而是滞后于降雨且与径流突增紧密耦合的物理信号。例如在强降雨事件 3（降雨强度 0.8547）中，跳跃分量强度为 0.7580，滞后为 1.5d，显示对极端降水的快速响应能力。在无雨或小雨时段，跳跃分量强度趋近于零，周期分量则保持平稳波动，说明 JMD 有效区分了事件性突变成分与缓变基流，未将基流误判为跳跃事件。综上，从数据驱动角度分析，JMD 是一种具有明确水文物理意义的信号解耦工具，能为后续时空预测模型提供可靠且可解释的特征输入。

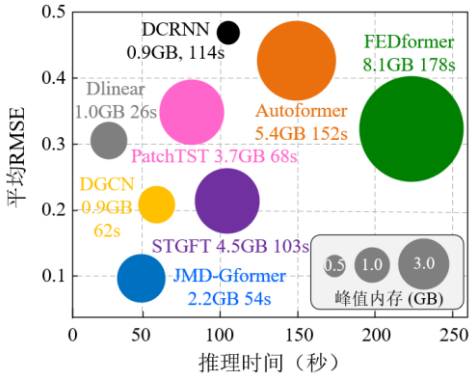
附表 4 暴雨事件时序特征与归一化强度分析

App. Tab. 4 Temporal characteristics and normalized intensity analysis of rainstorm events

事件编号	开始时间	结束时间	时序特征	降雨量	径流量	跳跃分量	时序滞后
1	2007-07-06	2007-07-06	峰值时间	186.5	187.1	187.4	0.9
			归一化强度	0.4091	0.3216	0.3562	/
2	2010-06-15	2010-06-17	峰值时间	1262.6	1263.8	1264.2	1.6
			归一化强度	0.3521	0.4385	0.3810	/
3	2012-07-19	2012-07-20	峰值时间	2027.3	2028.1	2028.8	1.5
			归一化强度	0.8547	0.8122	0.7580	/
4	2013-07-16	2013-07-16	峰值时间	2388.6	2389.0	2389.2	0.6
			归一化强度	0.5720	0.4988	0.5245	/
5	2015-03-21	2015-03-21	峰值时间	3001.4	3002.8	3003.5	2.1
			归一化强度	0.6205	0.6128	0.7050	/
6	2015-09-26	2015-09-27	峰值时间	3191.3	3191.8	3192.6	1.3
			归一化强度	0.7833	0.7321	0.8391	/
7	2016-05-03	2016-05-03	峰值时间	3410.5	3411.2	3411.3	0.8
			归一化强度	0.6182	0.6355	0.5724	/
8	2017-09-13	2017-09-14	峰值时间	3908.8	3909.5	3911.3	2.5
			归一化强度	0.5050	0.5320	0.4615	/

8. 时间复杂度分析

本节根据平均 RMSE、平均推理时间（秒）和峰值内存占用（GB），将所提模型与 7 种基线方法进行对比。时间复杂度分析结果如附图 3 所示。



附图 3 不同模型的时间复杂度分析

App. Fig. 3 Time complexity analysis of different models

9. 附件参考文献

[27]. Li YG, Yu RS, Shahabi C *et al*. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint* arXiv:1707.01926, 2017. DOI: 10.48550/arXiv.1707.01926.

[28]. Wu HX, Xu JH, Wang JM et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. Advances in neural information processing systems, 2021, **34**: 22419-22430.

[29]. Zhou T, Ma ZQ, Wen QS et al. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. International conference on machine learning. PMLR, 2022: 27268-27286.

[30]. Nie YQ, Nguyen N H, Sinthong P et al. A time series is worth 64 words: Long-term forecasting with transformers. arXiv preprint arXiv:2211.14730, 2022.

[31]. Zeng AL, Chen MX, Zhang L et al. Are transformers effective for time series forecasting? Proceedings of the AAAI conference on artificial intelligence. 2023, **37**(9): 11121-11128. DOI : 10.1609/aaai.v37i9.26317.

[32]. Li ZL, Wang YN, Liu HX et al. Dual-stream graph convolutional networks enable accurate real-time chlorine prediction in sparsely monitored water distribution networks. Water Research, 2025: 124824. DOI: 10.1016/j.watres.2025.124824.

[33]. Bi J, Wang ZQ, Yuan HT et al. Long-term Water Quality Prediction with Transformer-based Spatial-Temporal Graph Fusion. IEEE Transactions on Automation Science and Engineering, 2025, **22**: 11392-11404. DOI: 10.1109/TASE.2025.3535415.

[34]. Xu XH, Hu L, Liu JW *et al*. T-LSTM-VAE: A random missing data imputation model for multivariate time series. *Journal of the Franklin Institute*, 2025, **363**: 108231. DOI: 10.1016/j.jfranklin.2025.108231.