

刘康乐, 林涛, 孙军益, 等. 基于 Transformer 及其耦合模型的分钟级用水量预测新方法[J]. 净水技术, 2025, 44(9): 175–185.

LIU K L, LIN T, SUN J Y, et al. New method for minute-level water consumption prediction based on Transformer and the coupling models[J]. Water Purification Technology, 2025, 44(9): 175–185.

基于 Transformer 及其耦合模型的分钟级用水量预测新方法

刘康乐^{1,2}, 林涛^{1,2,*}, 孙军益³, 张雪⁴, 沈月生⁵

(1. 河海大学浅水湖泊综合治理与资源开发教育部重点实验室, 江苏南京 210098; 2. 河海大学环境学院, 江苏南京 210098; 3. 江苏省城镇供水安全保障中心, 江苏南京 210036; 4. 苏州市自来水有限公司, 江苏苏州 215002; 5. 上海中韩杜科泵业制造有限公司, 上海 201619)

摘要 【目的】精准的短期用水量预测是供水优化调度的关键前提, 对实现供水行业碳减排具有重要意义。然而, 现有研究多集中于小时级或日级预测, 缺乏分钟级用水量预测的深入探讨。本研究旨在构建高精度的分钟级用水量预测模型, 以提高供水调度效率, 助力低碳供水。【方法】以苏州市某居民小区智能水表监测数据为基础, 选取 6 min 时间步长的用水量序列作为研究对象。采用 Transformer 深度学习模型进行预测, 并与反向传播神经网络 (BPNN)、长短期记忆网络 (LSTM) 进行对比分析。为进一步提升预测精度, 引入经验模态分解 (EMD) 数据预处理方法, 构建 EMD-Transformer 耦合预测模型。评价指标包括决定系数 (R^2)、平均绝对百分比误差 (MAPE) 和均方根误差 (RMSE)。【结果】Transformer 模型的 R^2 达到 0.925, 较 BPNN ($R^2=0.812$) 和 LSTM ($R^2=0.873$) 分别提升 13.9% 和 6.0%; MAPE 为 10.89%, 较 BPNN (MAPE=16.24%) 和 LSTM (MAPE=12.75%) 显著降低。EMD-Transformer 耦合模型进一步提升了预测性能, R^2 提高至 0.964, MAPE 降至 7.50%, RMSE 减少 23.6%。分析表明, EMD 分解有效提取了用水量序列的多尺度特征, 使模型能更好地捕捉不同时间尺度下的用水规律。【结论】Transformer 架构在分钟级用水量预测中具有优越性, EMD-Transformer 耦合模型将预测精度提升至实用化水平。该成果为供水系统的实时优化调度提供了可靠的技术支持, 其方法论框架可推广至其他城市公用事业领域的负荷预测。分钟级精准预测有望使供水系统能耗降低, 对推动水务行业绿色低碳发展具有重要实践价值。

关键词 分钟级用水量预测 长短期记忆网络 (LSTM) Transformer 模型 EMD-Transformer 耦合模型 预测性能

中图分类号: TU991 文献标志码: A 文章编号: 1009-0177(2025)09-0175-11

DOI: 10.15890/j.cnki.jsjs.2025.09.022

New Method for Minute-Level Water Consumption Prediction Based on Transformer and the Coupling Models

LIU Kangle^{1,2}, LIN Tao^{1,2,*}, SUN Junyi³, ZHANG Xue⁴, SHEN Yuesheng⁵

(1. Key Laboratory of Integrated Regulation and Resource Development on Shallow Lakes, Ministry of Education, Hohai University, Nanjing 210098, China;

2. School of Environment, Hohai University, Nanjing 210098, China;

3. Jiangsu Urban Water Supply Safety Guarantee Center, Nanjing 210036, China;

4. Suzhou Municipal Water Supply Co., Ltd., Suzhou 215002, China;

5. Shanghai Zhonghan Duke Pump Industry Manufacturing Co., Ltd., Shanghai 201619, China)

Abstract [Objective] Accurate short-term water consumption forecasting is a crucial prerequisite for optimized water supply scheduling and holds significant importance for achieving carbon emission reduction in the water supply industry. However, existing research predominantly focuses on hourly or daily-scale predictions, with limited in-depth exploration of minute-level water consumption

[收稿日期] 2024-01-30

[基金项目] 江苏省住房和城乡建设厅科技项目: 基于优化调度和智能运控算法的供水节能关键技术与示范 (2022ZD033)

[作者简介] 刘康乐 (1999—), 男, 硕士研究生, 研究方向为饮用水安全保障、智慧水务, E-mail: 1322428126@qq.com。

[通信作者] 林涛 (1978—), 男, 教授, 研究方向为饮用水安全保障、水厂节水与排泥水安全回用、突发水污染应急等方面的研究与应用工作, E-mail: hit_lintao@163.com。

forecasting. This study aims to develop a high-precision minute-level water consumption prediction model to enhance water supply scheduling efficiency and support low-carbon water management. [**Methods**] Based on smart water meter monitoring data from a residential community in Suzhou, this study selected water consumption time series with a 6-minute time step as the research subject. The transformer deep learning model was employed for prediction and comparatively analyzed with backpropagation neural network (BPNN) and long short-term memory (LSTM) networks. To further enhance prediction accuracy, the empirical mode decomposition (EMD) data preprocessing method was introduced to construct an EMD-Transformer hybrid prediction model. Evaluation metrics include the coefficient of determination (R^2), mean absolute percentage error (MAPE), and root mean square error (RMSE). [**Results**] The transformer model achieved an R^2 of 0.925, representing a 13.9% and 6.0% improvement over BPNN ($R^2=0.812$) and LSTM ($R^2=0.873$) respectively. Its MAPE reached 10.89%, showing significant reductions compared to BPNN (MAPE=16.24%) and LSTM (MAPE=12.75%). The EMD-Transformer hybrid model demonstrated further enhanced predictive performance, elevating R^2 to 0.964 while reducing MAPE to 7.50% and decreasing RMSE by 23.6%. Analytical result indicated that EMD decomposition effectively extracted multiscale features from water consumption time series, enabling the model to better capture usage patterns across different temporal scales. [**Conclusion**] The transformer architecture demonstrates superior performance in minute-level water consumption prediction, with the EMD-Transformer hybrid model elevating prediction accuracy to practical application standards. This achievement provides reliable technical support for real-time optimization of water supply systems, while its methodological framework can be extended to load forecasting in other urban utility domains. The implementation of minute-level precision forecasting is projected to significantly reduce energy consumption in water supply systems, offering substantial practical value for promoting green and low-carbon development in the water industry.

Keywords minute-level water consumption prediction long short-term memory (LSTM) Transformer model EMD-Transformer coupling model predicted performance

供水行业是国家碳排放的重要环节^[1],随着国家“双碳”战略的提出,供水行业的碳减排行动迫在眉睫。供水水量的优化调度是供水行业节能减排的一个重要前提,然而,现阶段的供水优化调度主要靠工人经验调配来完成,这种调度方法并不是最优解,容易造成能量的浪费。水量的精准预测是优化调度的基础,它不仅可以准确指导泵送水量,而且为后续的泵组高效运行提供了理论依据。水量预测可以分为长期、中期和短期,其界限分别为24个月以上、3~24个月和3个月以内^[2]。对于水量优化调度,水量的短期预测更为关键,然而现有的文献缺少分钟级用水量预测。

由于用户用水量的非线性特征明显,传统的数学模型在预测用户水量时并不理想^[3],Mu等^[4]使用自回归积分滑动平均模型来预测不同时间步长的水量数据但是预测结果较差。神经网络是近些年兴起的强大的非线性黑箱模型,它可以在输入值与输出值之间建立起准确的非线性关系,从而作出预测。Bai等^[5]使用广义回归神经网络(GRNN)和前馈神经网络(FFNN)预测了水厂日用水量并且得到了较好的预测效果[平均绝对百分比误差(MAPE)=1.385%]。然而上述模型并未将时间因素考虑进模

型,Mu等^[4]使用长短时记忆网络(LSTM)模型对15 min级和24 h级用水量进行预测,其预测结果均强于传统的非时间序列型神经网络;Guo等^[6]使用门循环控制单元(GRU)对15 min和24 h级水量进行预测,结果优于人工神经网络(ANN)模型。上述循环神经网络虽然在水量预测领域取得了不错的效果,但是它自身仍存在缺陷如无法并行运行、捕捉长序列特征困难。Transformer是Google在2017年提出的强大的网络模型,已在自然语言处理领域广泛使用,如ChatGPT,其在处理序列式数据时可以并行运行且可以与前序很远的的数据建立联系^[7]。然而,现在仍然没有报道将Transformer模型应用于水量预测领域。

此外,水量短期预测时其水量数据波动较大,耦合模型较单一模型表现出越来越大的优势。Hu等^[8]采用CNN-BiLSTM耦合模型与传统神经网络模型对比对水量进行预测,结果表明,耦合模型预测性能更优;Pu等^[9]使用CNN-LSTM耦合模型来预测短期水量,耦合模型较单一模型表现出更优的预测精度。经验模态分解(EMD)算法在分解复杂时间序列数据方面的高效简单,考虑到用水数据的复杂性,本研究选择EMD算法与Transformer模型耦合对模

型性能作进一步优化。

本研究将采用 Transformer 模型对苏州某小区 6 min 时间步长分辨率用水量进行预测,同时与 LSTM、反向传播神经网络 (BPNN) 模型进行对比,最后使用 EMD-Transformer 模型对模型进行改进优化,探究该耦合模型的预测性能。

1 材料与方法

1.1 数据收集

苏州是三角重要中心城市之一,其供水水平提升必将为全国其他城市起到示范引领作用。以苏州市某小区为例开展用户水量预测研究,该小区建筑面积为 253 300 m²,约有房屋为 3 000 户,居住人口约为 9 000 人,每年的总需水量约为 34 4 692 m³。对该小区 2020 年 2 月 1 日—2020 年 2 月 29 日 6 min 时间步长的用水量数据进行收集,共收集了 6 959 条水量信息。

1.2 EMD

EMD 通过自身的时间尺度来进行信号分解^[10],用水量数据波动起伏较大,非线性特征明显,较难预测,利用 EMD 可以将用水量数据分解为更平稳、更有规律的易于预测的子序列。其具体过程如下。

步骤 1:找出用水量 $[x(t)]$ 的所有极值点,设本征模函数的总数为 i ,并令 $i=1$ 。

步骤 2:用插值法对用水量的极小值点和极大值点形成下包络 $[e_{\min}(t)]$ 和上包络 $[e_{\max}(t)]$ 。

步骤 3:计算下包络和上包络的均值 $m_i(t)$,并产生均值包络。

步骤 4:计算原始用水量数据与均值 $m_i(t)$ 的差,即 $h_i(t)$ 。

步骤 5:检验 $h_i(t)$ 是否满足本征模函数的 2 个条件。

步骤 5.1:若满足,则 $h_i(t)$ 为第 i 个本征模函数,记为 $c_i(t)$,并计算残余分量如式(1),令 $x(t)=r_i(t)$,重复以上过程, $i=i+1$ 。

$$r_i(t) = x(t) - h_i(t) \quad (1)$$

步骤 5.2:若不满足,则令 $x(t) = h_i(t)$,并返回步骤 1。

最后用水量数据将被分解为如式(2)。

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (2)$$

其中: $c_i(t)$ ——第 i 个本征模函数;

$r_n(t)$ ——残余分量,表示的是用水量数据 $x(t)$ 的一般趋势。

1.3 BPNN 模型

BPNN 是最基础的神经网络模型,由输入层、隐藏层和输出层构成。输入的数据经过隐藏层中的神经元进行处理后,再通过输出层进行输出,从而通过权值建立起输入与输出之间的非线性关系。值得注意的是,BPNN 并没有处理时间序列信息的能力,只能简单识别前序时间序列水量与预测水量之间的简单关系,而不能将前序时间序列水量中的时间信息进行提取,这可能会导致 BPNN 在水量预测中并不占优势。

1.4 LSTM 模型

LSTM 是一种循环神经网络,其可以捕捉前序序列的信息从而对后序序列进行预测^[11]。每个 LSTM 单元由输入门、遗忘门和输出门组成,如式(3)~式(5)。

遗忘门:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

输入门:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (4)$$

输出门:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (5)$$

其中: σ ——Sigmoid 函数,用于将计算结果映射到 0~1。

LSTM 模型的每个存储单元如式(6)~式(8)。

单元状态:

$$c_t = f_t c_{t-1} + i_t \bar{c}_t \quad (6)$$

$$\bar{c}_t = \sigma(W_c x_t + U_c h_{t-1} + b_c) \quad (7)$$

输出向量:

$$h_t = o_t \sigma(c_t) \quad (8)$$

其中: c_t ——当前时刻的单元状态;

$\bar{c}_t$ ——候选单元状态,它是当前时刻可能要存入单元状态的候选信息;

f_t, i_t, o_t ——LSTM 的遗忘门、输入门和输出门;

x_t —— t 时刻的输入向量;

c_{t-1}, h_{t-1} ——前一时刻的单元状态和输出向量;

h_t ——输出向量;

W_f, W_i, W_o, W_c ——遗忘门、输入门、输出门及单元状态对于输入的权重矩阵;

U_f, U_i, U_o, U_c ——遗忘门、输入门、输出门及单元状态对于上一时刻输出的权重矩阵;

b_f, b_i, b_o, b_c ——相应的偏差向量。

LSTM 通过增加“门”的概念解决了传统循环神经网络中存在的梯度消失问题,但在处理非常长的序列信息时,LSTM 可能仍然面临挑战,并且需要额外的技术或方法来提高其性能。

1.5 Transformer 模型

Transformer 作为自然语言处理领域的新型模型^[7],一经提出就受到了广泛关注,其序列-序列结构允许数据进行并行处理,大大提高了模型的运行效率。Transformer 一共有 2 个部分,分别为编码器和解码器。编码器的实现过程就如同在编撰一本由数据构成的字典,而解码器的过程则是在查询字典中的内容,其具体的实现过程如图 1 所示。

左边为编码器部分,包括多头注意力机制、残差连接和层归一化,前馈神经网络、残差连接和层归一化。右边为解码器部分,解码器较编码器则多了掩码多头注意力机制和残差连接和层归一化,这是为了防止程序运行过程中使用已知水量来预测已知水量。编码器和解码器可以多层堆叠以增强模型的学习能力,本研究选用 6 层编码器和解码器堆叠。

多头注意力机制是 Transformer 模型的核心,其实现过程如下:首先,在自注意力的基础上,对查询向量(query, Q)、键向量(key, K)、值向量(value, V)根据注意力头数(head)进行拆分。 Q, K, V 都源于输入特征本身,是根据输入特征产生的向量。随后,将上标第二位参数为 1 的 Q, K, V (例如: Q^1, K^1, V^1)归为 Head1。Head2, ..., Headn 同理。每个 Head 的输出如式(9)。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (9)$$

其中: d_K —— K 的维度大小。

最后,对每个 Head 的输出进行拼接,然后将拼接后的输出乘以 $W^0 \in R^{h_{dv} \times d_{\text{model}}}$ 得到多头注意力机

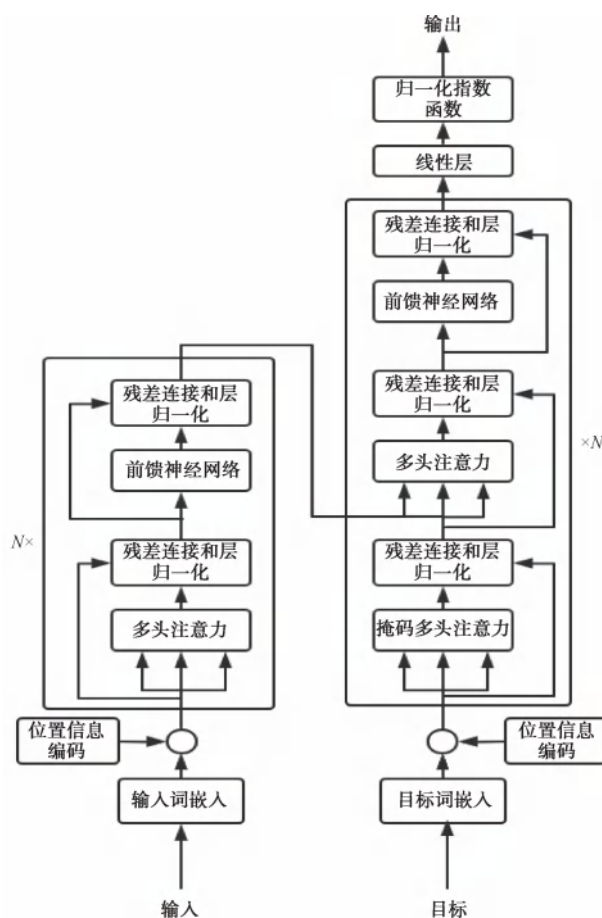


图 1 Transformer 模型结构

Fig. 1 Structure of Transformer Model

制的输出。这里有 $d_k = d_v = d_{\text{model}}/h$, h 为多头注意力机制的头数,本研究使用的多头注意力机制头数为 10,即 $h = 10$ 。Transformer 的自注意力机制可以计算输入序列中任意 2 个位置之间的依赖关系,而不仅仅是相邻位置。这使得 Transformer 能够更好地捕捉长距离依赖关系,尤其是在处理长序列时,这为使用 Transformer 模型来预测用户用水量打下了良好的基础。

1.6 模型比选方法

为了对比模型的预测性能,选用测试集均方根误差(RMSE)、平均绝对百分比误差(MAPE)和相关系数(R^2)作为性能指标来指示模型的预测优劣^[12-13]。3 个性能指标的计算方法如式(10)~式(12)。

$$R_{\text{RMSE}} = \sqrt{\frac{\sum (y_{\text{true},i} - y_{\text{pred},i})^2}{N}} \quad (10)$$

$$M_{\text{MAPE}} = \frac{\sum |y_{\text{true},i} - y_{\text{pred},i}|}{N} \times 100\% \quad (11)$$

$$R^2 = 1 - \frac{S_{SSE}}{S_{SST}} = 1 - \frac{\sum (y_{\text{true},i} - y_{\text{pred},i})^2}{\sum (y_{\text{pred},i} - \overline{y_{\text{true}}})^2} \quad (12)$$

其中: $y_{\text{true},i}$ ——测试集第 i 个用水量实际值, m^3/h ;
 $y_{\text{pred},i}$ ——测试集第 i 个用水量预测值, m^3/h ;
 N ——测试集的数据个数;
 $\overline{y_{\text{true}}}$ ——测试集用水量真实值的平均值, m^3/h ;
 R_{RMSE} ——RMSE 值;
 R_{MAPE} ——MAPE 值;
 S_{SSE} ——误差平方和;

S_{SST} ——总平方和。

RMSE 和 MAPE 值越小,说明模型的预测误差越小; R^2 越大,越接近 1,说明模型的拟合效果越好。

2 结果与讨论

2.1 数据分析

共收集了 6 959 条 6 min 时间步长用水量数据,图 2 为 2020 年 2 月用水量折线图和一周用水量折线图。总的来说,每日的用水量变化呈现出相同的变化趋势,每日有 2 个用水高峰,分别出现在 9:00 和 18:00 附近,这与人的作息规律相符合。整月的用水量高峰出现在 2020 年 2 月 27 日 19:47,流量为 $68.15 \text{ m}^3/\text{h}$;整月的用水量低峰则出现在 2020 年 2 月 14 日 3:33,流量为 $3.95 \text{ m}^3/\text{h}$ 。

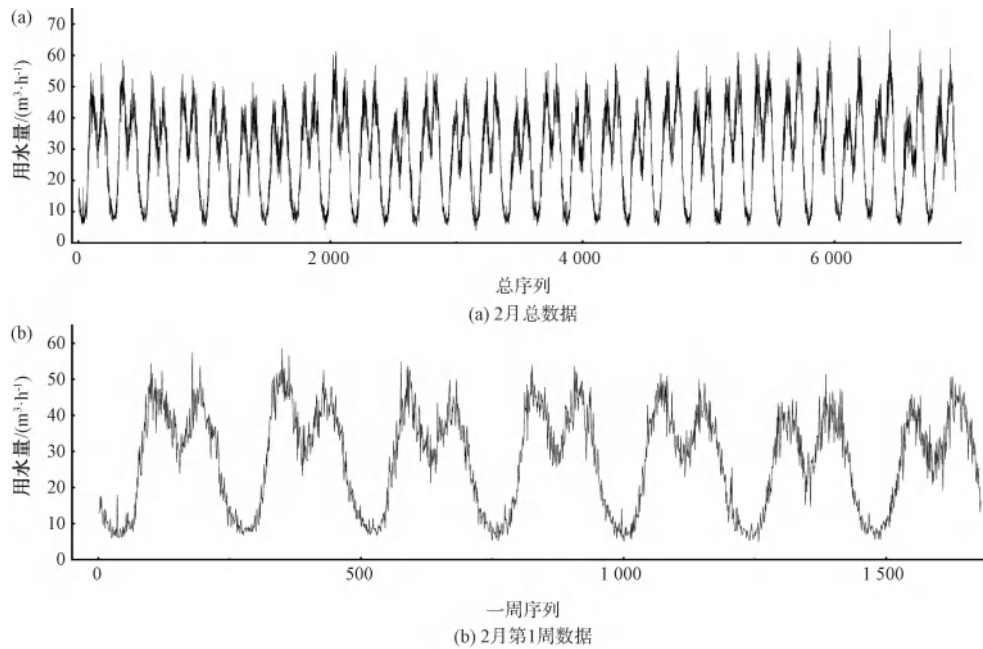


图2 用水量数据折线图

Fig. 2 Line Chart of Water Consumption Data

为了减少模型运行时间,增加模型收敛速度,通过式(13)对用水量数据进行归一化处理^[14]。归一化后的数据变为 0~1,数据经预测后再通过式(14)进行反归一化操作重新变为预测水量数据。

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (13)$$

$$y_{\text{pred}} = y_i(x_{\max} - x_{\min}) + x_{\min} \quad (14)$$

其中: x'_i ——归一化后的数据;
 x_i ——数据集的 i 个用水量数据, m^3/h ;
 $x_{\min}$ ——数据集最小值, m^3/h ;

$x_{\max}$ ——数据集最大值, m^3/h ;

y_{pred} ——反归一化后的预测用水量数据, m^3/h ;

y_i ——第 i 个预测数据, m^3/h 。

对水量数据进行分割,探究了 20~50 个输入值对用户水量预测性能的影响,最后选用前 40 个数据作为输入值,后一个数据作为输出值,对水量数据进行分组重构。将重新得到的数据进行分组,前 80% 的数据作为训练集,后 20% 的数据作为测试集^[15]。

2.2 LSTM 模型预测结果分析

通过 Matlab R2017b 软件使用 LSTM 模型对用

水量数据进行预测。隐藏层神经元数目、Dropout 率、学习率、梯度阈值是影响 LSTM 模型预测性能的关键因素^[16],为了提高模型性能,通过网格搜索寻找相应的参数最优值。此外,由于 Dropout 率的存在、初始权重的随机性等因素,LSTM 每次运行的结果可能存在一定的偏差,为了公平决策出最佳参数,对每个相应参数下的模型进行 50 次运行并取最优 1 次作为此参数下的模型性能。

LSTM 模型隐藏层神经元数目越多,模型的预测能力越强,但是相应的模型训练时间也会增加,可能会造成模型泛化能力下降,模型在测试集中的表现反而有所下降。图 3(a)为隐藏层神经元数目在 40~150 内模型的预测性能,预测误差总体呈现出先下降后上升的趋势,当隐藏层神经元数目为 60 时,模型的预测误差最小,此时的 R^2 、RMSE 和 MAPE 分别为 0.920、4.19 m^3/h 和 11.11%。

Dropout 率反映了在训练期间删除网络结点的概率,可以有效减少模型的过拟合现象。为了得到

最佳的 Dropout 率,尝试了以 0.05 为间隔,从 0.10~0.30 的 Dropout 率下的 LSTM 性能[图 3(b)]。当 Dropout 率为 0.10 时,模型表现最优,这可能是 Dropout 率增加虽然提升了模型的泛化性能,但是也导致了删除结点的概率增加,从而造成了网络信息传递的减少,影响了模型的预测性能。

图 3(c)和图 3(d)为学习率为 0.01~0.10 和梯度阈值为 0.01~0.10 LSTM 模型的预测性能变化情况。随着学习率的增加,模型预测误差呈现出先降低后增大的趋势,学习率过小可能会造成模型掉入局部最优解,从而降低模型性能,学习率过大可能会造成模型参数更新时跨度过大从而跨过最优解。综合分析,当学习率为 0.04 时模型最优。梯度阈值可以将 LSTM 模型的梯度更新限定在一定范围内,从而解决循环神经网络的梯度爆炸问题,随着梯度阈值的增大,LSTM 模型性能呈现出振荡趋势。当梯度阈值为 0.08 时,模型的预测性能最优。

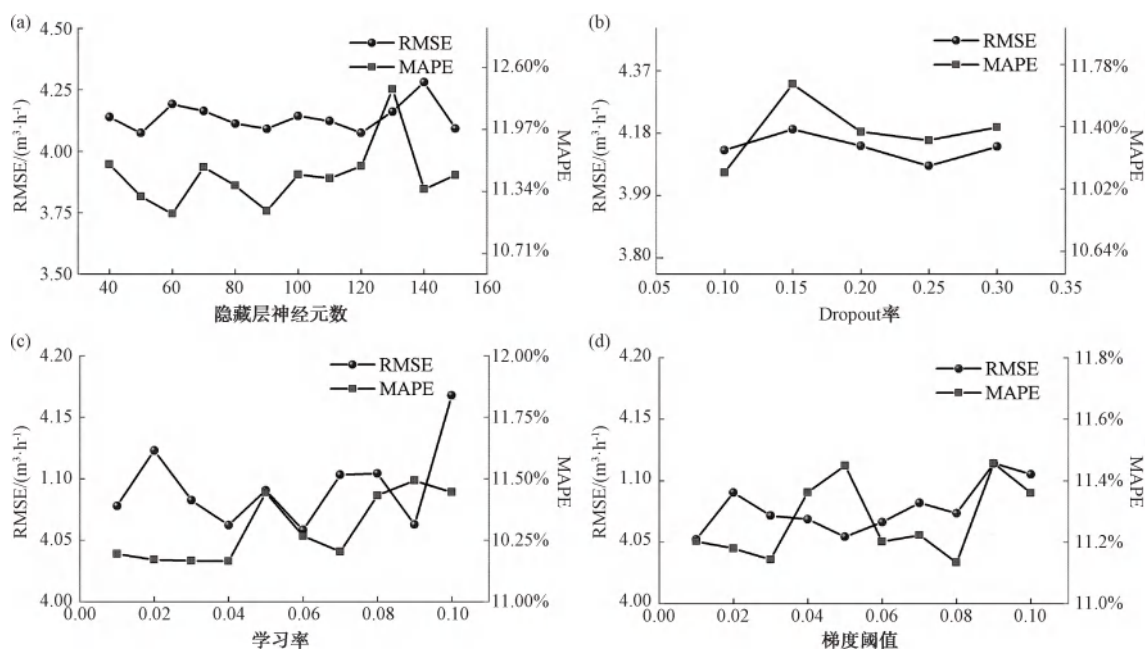


图 3 LSTM 模型参数优化

Fig. 3 Optimization of LSTM Model Parameters

综上所述,LSTM 模型的最优参数设置:隐藏层神经元数目为 60,Dropout 率为 0.1,学习率为 0.04,梯度阈值为 0.08。使用最优参数对用户水量进行预测,最终得到的 R^2 、RMSE、MAPE 分别为 0.922、4.12 m^3/h 、11.07%。 R^2 非常接近于 1,说明模型拟合效果较好,通过 LSTM 模型,用水量的前 40 个数

据可以较好地解释后一时刻数据。预测误差也较小,仅为 11.07%,这为以后的供水优化调控提供了基础。

2.3 Transformer 模型预测结果分析

Transformer 模型基于 Python 3.8 通过 PyCharm Community Edition 平台进行实现。与 LSTM 类似,

Transformer 模型的性能也受相应参数的影响;输入数据个数、学习率、初始范围、缩放因子(γ)、优化器以及特征尺寸^[17]。为了提高模型性能,通过网格搜索寻找相应的参数最优值。

输入数据个数即时间步数,时间步数过少会导致模型对水量特征提取不充分,过多则会导致模型的收敛时间增加与泛化能力下降。图 4(a)列举了不同的时间步数下模型的 RMSE 和 MAPE 值。RMSE 整体呈现出下降趋势,而 MAPE 则呈上升趋势,这是由于随着时间步数的增加,模型提取到的特征更多,部分极端用水条件下的用水量预测更加准确, RMSE 降低,但是模型的泛化能力下降,整体的预测准确性反而下降, MAPE 升高。综合考虑来看,选取 40 为输入数据个数。

与 LSTM 模型类似, Transformer 的学习率也是影响预测性能的关键因素,试验学习率从 0.001 ~ 0.048,共 48 个 Transformer 模型。由图 4(b)可知,随着学习率的增大,模型整体性能呈下降趋势,误差呈上升趋势,当学习率为 0.001 时模型预测性能最

优,此时 RMSE 和 MAPE 分别为 $4.22 \text{ m}^3/\text{h}$ 和 11.87%。

初始范围和 γ 是 Transformer 的另外 2 个重要参数,初始范围的正确选取可以保证模型的稳定性, γ 则控制了注意力的强度,保证了模型信息传递的可靠性。图 4(c)和图 4(d)给出了不同初始范围和 γ 值下模型的预测性能。当初始范围为 0.01, γ 为 0.95 时,模型的预测性能最优。

Transformer 中优化器决定了模型训练时参数更新的频次,优化器越小,模型的更新参数频率越高。而特征尺寸指的是输入特征的维度大小,对模型的性能和计算效率都有影响。特征尺寸越大,模型捕捉数据的信息越多,但是相应的模型效率会降低,而特征尺寸越小,捕捉信息越模糊,但是计算量会相应减少,模型效率提升。当优化器为 1,特征尺寸为 100 时,模型的误差最小,预测性能最优。

综上所述, Transformer 模型的最优参数设置:时间步数为 40,学习率为 0.001,初始值为 0.01,

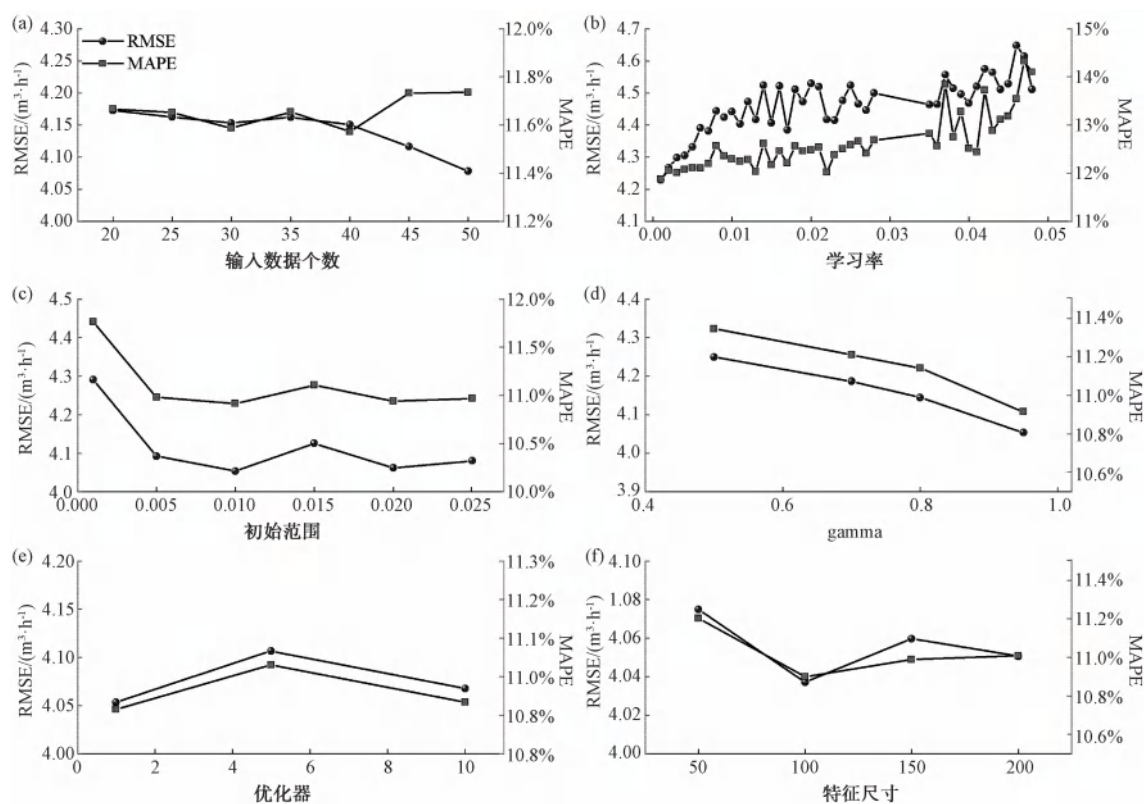


图 4 Transformer 模型参数优化

Fig. 4 Optimization of Transformer Model Parameters

gamma 为 0.95, 优化器 1, 特征尺寸为 100。使用最优参数对用户水量进行预测, 最终得到的 R^2 、RMSE、MAPE 分别为 0.925、4.03 m^3/h 、10.89%。 R^2 较高且预测误差较小, Transformer 模型可以较好地进行用户水量预测。

2.4 模型比选

对 LSTM 和 Transformer 模型的预测性能进行比较, 结果如表 1 所示, 同时, 本研究也探究了传统的 BPNN 对用户水量的预测性能, 结果也如表 1 所示。

表 1 模型性能对比

Tab. 1 Comparison of Models Performance

模型名称	R^2	RMSE/ $(\text{m}^3 \cdot \text{h}^{-1})$	MAPE
BPNN	0.916	4.26	11.63%
LSTM	0.922	4.12	11.07%
Transformer	0.925	4.03	10.89%

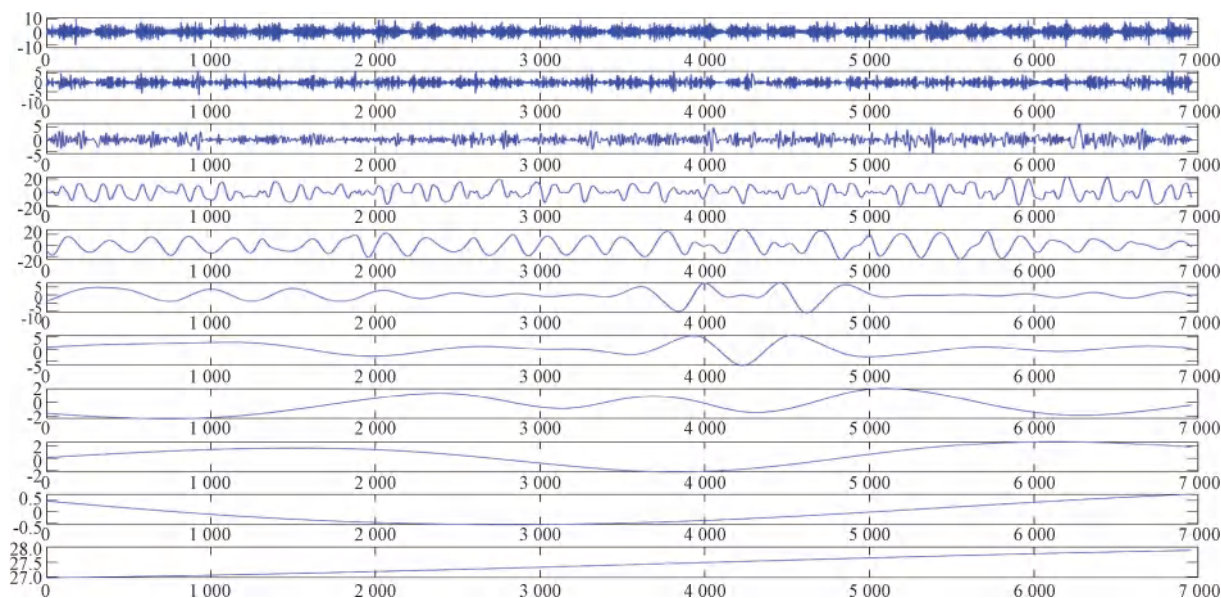
综合分析这 3 个模型的预测性能, BPNN 最差, 这是因为 BPNN 通常使用固定长度的窗口来处理时间序列数据, 这会导致网络难以捕捉长期依赖关系。对于长期依赖关系比较重要的时间序列数据, 这种方法可能效果不佳, 在本研究中 BPNN 模型并不能将时间序列的信息考虑在内, 其会同时对 40 个输入数据进行分析而忽略其包含的时间信息, 所以预测效果并不理想。而 LSTM 较 BPNN 则优异很多, LSTM 的关键在于它引入了称为“门”的结构, 包括

输入门、遗忘门和输出门, 这些门控制着信息的流动, 使 LSTM 能够有选择地记住或忘记输入数据中的信息。这使得 LSTM 可以通过循环用到上一次运行的信息, 将水量数据中的时间信息充分利用。然而, LSTM 在处理学习长期序列关系的时候仍然存在着局限性: 例如长期依赖问题、参数量大问题、序列并行受限问题, Transformer 则很好地解决了这些问题, 相较于 LSTM, Transformer 采用了注意力机制来捕捉序列中的长期依赖关系, 它不需要通过递归的方式来处理序列, 因此能够更好地处理长距离的依赖关系。此外, Transformer 还具有更高的并行性, 能够更有效地利用计算资源, 其预测性能在 3 个模型中也是最优, R^2 最高, 为 0.925, RMSE 和 MAPE 最小, 分别为 4.03 m^3/h 和 10.89%。综合来看, 在预测水量数据时, Transformer 模型是最佳选择。

2.5 EMD-Transformer 耦合模型

尽管 Transformer 模型表现出了比其他 2 个模型更优的预测性能, 然而其 MAPE 仍然大于 10%, 较难运用到实际生活中。考虑到水量数据波动较大, 非线性较强, 非平稳性特征明显, 而 EMD 可以将非平稳性数据转化为较为平稳易于预测的数据序列^[18]。为此, 开发 EMD-Transformer 耦合模型以减少模型的预测误差。

如图 5 所示, 在 Matlab 中使用 emd 函数将用水



注: x 方向为数据集个数; y 方向为用水量, 单位为 m^3/h 。

图 5 用水量数据 EMD 分解图

Fig. 5 EMD Decomposition Diagram of Water Consumption Data

量数据分解为 11 个子序列,相较于图 2 的原序列数据,图 5 中分解后的子序列的波动性明显减弱,平稳性增强。通过对每个子序列分别使用 Transformer 模型并在最后对每个子序列的预测值进行加和得到 EMD-Transformer 模型的预测值。同样地,探究 EMD-Transformer 模型最佳运行参数,步骤如前所述,最终 EMD-Transformer 模型的最佳参数:学习率为 0.001,初始值为 0.006, γ 为 0.95,优化器 1,特征尺寸为 100,使用最佳参数对用水量数据进行预测,耦合模型的 R^2 、RMSE 和 MAPE 分别为 0.964、 $2.81 \text{ m}^3/\text{h}$ 和 7.50%。

EMD-Transformer 耦合模型的预测结果如图 6 所示。图 6(a) 为用水量实际值与预测值之间的散点关系图,空心点代表训练集数据,实心点代表测试集数据。图中虚线为斜率为 1 的参考线,即预

测值与实测值相等时连成的线。当散点越接近虚线时,模型的预测性能越好。图中的点均分布在虚线附近,这说明耦合模型的预测值与真实值接近,这与耦合模型的 R^2 为 0.964 相吻合。图 6(b)~图 6(d) 则分别为全部数据、训练集数据和测试集数据实际用水量与预测用水量折线图,实线为实际用水量,虚线为预测用水量。实际用水量与预测用水量折线基本重合,这说明耦合模型不仅在训练集上有着良好的预测性能,其在测试集上的预测也非常优异。与单一的 Transformer 模型相比,耦合模型的 R^2 有着明显的提高,误差也有明显下降,MAPE 由 10.89% 降至 7.50%。这说明 EMD-Transformer 耦合模型可以显著改善单一模型预测精度不够高的问题,进一步增强模型的预测性能,更好地预测用水量。

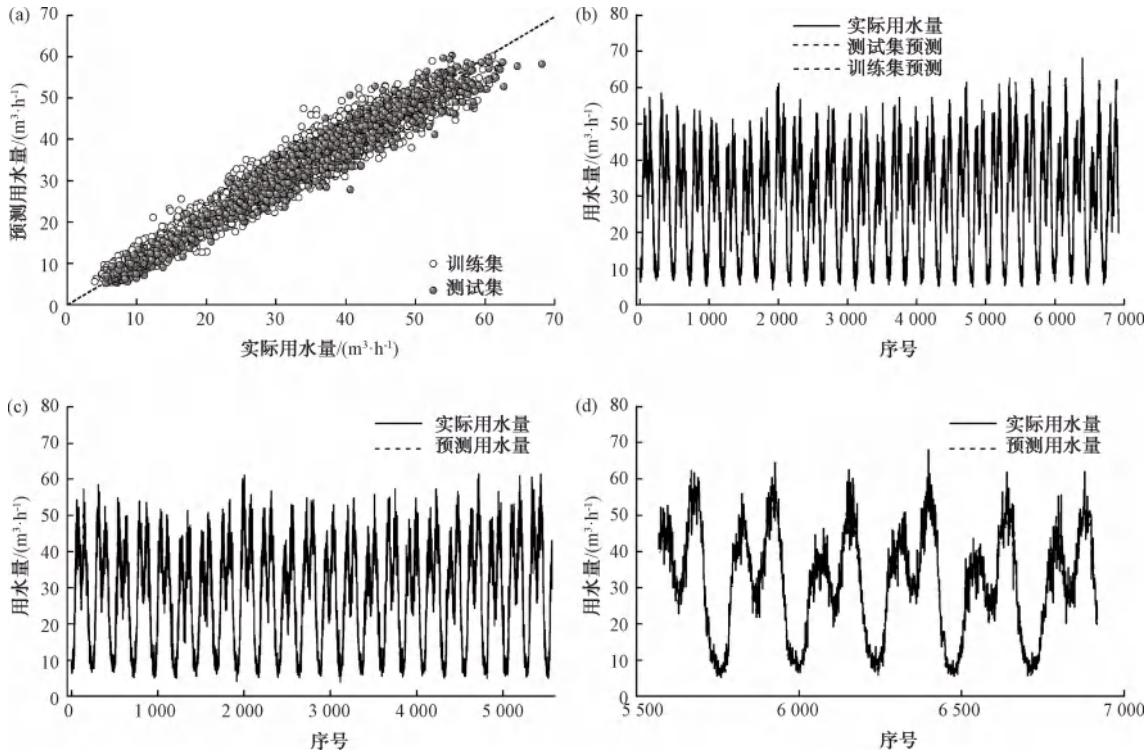


图 6 EMD-Transformer 预测结果
Fig. 6 Prediction Results of EMD-Transformer

由表 2 可知,将 BPNN 和 LSTM 与 EMD 耦合同样能够改善单一模型的预测性能,耦合模型的 MAPE 比单一模型显著降低。但是在 3 种耦合模型中,EMD-Transformer 模型的预测性能最优,误差最小,因此,在后续水量预测中,更推荐 EMD-Transformer 模型。

表 2 耦合模型性能对比
Tab. 2 Comparison of Coupling Models Performance

模型名称	R^2	RMSE/($\text{m}^3 \cdot \text{h}^{-1}$)	MAPE
EMD-BPNN	0.931	3.44	9.13%
EMD-LSTM	0.964	2.95	7.81%
EMD-Transformer	0.964	2.81	7.50%

3 结论与展望

3.1 结论

(1) 对苏州某小区部分区域 6 min 时间步长用水量数据进行收集,共收集了 6 959 条数据。分别用 LSTM、Transformer 建立用户用水量预测模型,并综合对比了 BPNN、LSTM 和 Transformer 的预测性能,最终得到 Transformer 的预测性能最优,相关系数最高 ($R^2 = 0.925$), 预测误差最小 ($\text{MAPE} = 10.89\%$)。

(2) Transformer 模型的预测性能仍然不能满足实际调度需求,需要进一步降低。本研究选用经验模态分解将用水量数据分解为更平稳、易于预测的多个子序列并分别预测,从而建立了 EMD-Transformer 耦合模型,预测效果显著提高,相关系数由 0.925 提升至 0.964, MAPE 由 10.89% 降低至 7.50%。

(3) 对比 3 种耦合策略, EMD-Transformer 耦合模型表现最优,该模型在分钟级用水量预测上的优异表现为现实中后续实现水量优化调度提供了重要保障,有利于实现供水行业的碳减排。

3.2 展望

通过对本研究水量预测方法的迁移,可以为不同区域分钟级用户水量进行精准预测。在实际应用中具有广泛的后续应用意义。

(1) 水量预测可以帮助水资源管理部门更准确地了解未来一段时间内的水量供需情况,从而制定合理的水资源调度和管理策略。这有助于提高水资源的利用效率,确保供水安全。

(2) 水量预测可以为泵组优化调度提供重要的指导,帮助泵组管理者合理安排泵站运行,提高供水效率,降低运行成本,实现节能减排和资源有效利用的目标。

参考文献

- [1] 赵平伟, 蓝杨, 冯偲慙, 等. 基于预测的城市供水管网运行评估方法[J]. 净水技术, 2023, 42(2): 147-155.
ZHAO P W, LAN Y, FENG S M, et al. Evaluating method of prediction-based for operation of urban water supply pipelines distribution network [J]. Water Purification Technology, 2023, 42 (2): 147-155.
- [2] WONG J S, ZHANG Q, CHEN Y D. Statistical modeling of daily urban water consumption in Hong Kong: Trend, changing patterns, and forecast [J]. Water Resources Research, 2010, 46(3). DOI: 10.1029/2009WR008147.
- [3] CHEN L, YAN H, YAN J, et al. Short-term water demand forecast based on automatic feature extraction by one-dimensional convolution [J]. Journal of Hydrology, 2022, 606: 127440. DOI: 10.1016/j.jhydrol.2022.127440.
- [4] MU L, ZHENG F, TAO R, et al. Hourly and daily urban water demand predictions using a long short-term memory based model [J]. Journal of Water Resources Planning and Management, 2020, 146(9): 05020017. DOI: 10.1061/(ASCE)WR.1943-5452.0001276.
- [5] BAI Y, WANG P, LI C, et al. A multi-scale relevance vector regression approach for daily urban water demand forecasting [J]. Journal of Hydrology, 2014, 517: 236 - 245. DOI: 10.1016/j.jhydrol.2014.05.033.
- [6] GUO G, LIU S, WU Y, et al. Short-term water demand forecast based on deep learning method [J]. Journal of Water Resources Planning and Management, 2018, 144(12): 04018076. DOI: 10.1061/(ASCE)WR.1943-5452.0000992.
- [7] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. arXiv, 2017. DOI: 10.48550/arXiv.1706.03762.
- [8] HU P, TONG J, WANG J, et al. A hybrid model based on CNN and Bi-LSTM for urban water demand prediction [J]. IEEE, 2019. DOI:10.1109/CEC.2019.8790060.
- [9] PU Z, YAN J, CHEN L, et al. A hybrid Wavelet-CNN-LSTM deep learning model for short-term urban water demand forecasting [J]. Frontiers of Environmental Science & Engineering, 2023, 17(2): 97-110.
- [10] 汪岗, 马亮, 陈奕霖. 基于 EMD 与 Attention-LSTM 的铁路货运站短期装车量预测研究[J]. 铁道货运, 2023, 41(12): 60-69.
WANG G, MA L, CHEN Y L. Short-term prediction of car loading quantities for railway freight station based on EMD and attention-LSTM [J]. Railway Freight Transport, 2023, 41 (12): 60-69.
- [11] 张薇薇, 赵平伟, 王景成. 基于长短时神经网络的城市需水量预测应用[J]. 净水技术, 2019, 38(s1): 257-260, 286.
ZHANG W W, ZHAO P W, WANG J C. Application of long and short time neural network in urban water demand forecasting [J]. Water Purification Technology, 2019, 38(s1): 257-260, 286.
- [12] 刘志壮, 吕谋, 周国升. 基于小波组合模型的短期城市用水量预测[J]. 给水排水, 2020, 46(10): 110-114, 131.
LIU Z Z, LÜ M, ZHOU G S. Short term water supply load forecasting based on wavelet combination model [J]. Water & Wastewater Engineering, 2020, 46(10): 110-114, 131.
- [13] 朱铭江, 裴娅, 张祖鹏. 基于机器学习的城市用水量预测模型研究[J]. 浙江水利科技, 2022, 50(4): 103-107.
ZHU M J, QIU Y, ZHANG Z P. Research on urban water

- consumption prediction model based on machine learning [J]. Zhejiang Water Resources Science and Technology, 2022, 50 (4): 103-107.
- [14] 顾杰, 王嘉, 邓俊晖, 等. 基于 ARIMA 模型与 BP 神经网络算法的水质预测[J]. 净水技术, 2020, 39(6): 73-82.
- GU J, WANG J, DENG J H, et al. Water quality prediction based on ARIMA model and BP neural network algorithm [J]. Water Purification Technology, 2020, 39 (6): 73-82.
- [15] 高赫余, 王圣, 吴潇勇. 基于机器学习 XGBoost 集成模型的边界流量计日流量预测方法[J]. 净水技术, 2020, 39(10): 181-186.
- GAO H Y, WANG S, WU X Y. Daily flow prediction method of boundary flowmeter based on machine learning XGBoost integrated model [J]. Water Purification Technology, 2020, 39 (10): 181-186.
- [16] 朱健. LSTM 模型在海绵城市建设中的应用[J]. 净水技术, 2022, 41(2): 103-110.
- ZHU J. Application of LSTM model in sponge city construction [J]. Water Purification Technology, 2022, 41(2): 103-110.
- [17] 申志豪, 李娜, 尹世豪, 等. 基于 TPA-Transformer 的机票价格预测[J]. 数据与计算发展前沿, 2023, 5(6): 115-125.
- SHEN Z H, LI N, YIN S H, et al. Airfare price prediction based on TPA-transformer [J]. Frontiers of Data & Computing, 2023, 5(6): 115-125.
- [18] FV G T, JIN Y W, SUN S, et al. The role of deep learning in urban water management: A critical review[J]. Water Research, 2022, 223: 118973. DOI: 10.1016/j.watres.2022.118973.

(上接第 174 页)

(4) 层级化的同步采集流程设计实现了 NTP、GNSS 授时和激励信号同步 3 种技术的有机结合, 确保在各种环境条件下均能维持高精度泄漏定位。

在未来的工作中, 还应结合管网材质特性和环境条件合理选择激励信号特征参数, 进一步优化小波变换中的参数选择方法, 并探索将该技术拓展至其他类型管网泄漏检测领域的可能性。

参考文献

- [1] 胡文俊, 孙岩. 世界主要国家的节水措施及其启发与思考[J]. 水利经济, 2024, 42(2): 1-6.
- HU W J, SUN Y. Inspiration and reflection on review of water saving measures in major countries [J]. Journal of Economics of Water Resources, 2024, 42(2): 1-6.
- [2] 刘书明, 徐强, 信昆仑, 等. 供水管网漏损系统性管控技术与装备[J]. 建设科技, 2022(7): 115-117.
- LIU S M, XU Q, XIN K L, et al. Technologies and equipment for systematic leakage control of water distribution networks [J]. Construction Science and Technology, 2022(7): 115-117.
- [3] KOTHANDARAMAN M, LAW Z, EZRA M A G, et al. Adaptive independent component analysis-based cross-correlation techniques along with empirical mode decomposition for water pipeline leakage localization utilizing acoustooptic sensors [J]. Journal of Pipeline Systems Engineering and Practice, 2020, 11 (3): 04020027. DOI: 10.1061/(ASCE)PS.1949-1204.0000471.
- [4] KNAPP C H, CARTER G C. Generalized correlation method for estimation of time-delay [J]. IEEE Transactions on Acoustics Speech and Signal Processing, 1976, 24(4): 320-327.
- [5] GUO G C, YU X P, LIU S M, et al. Novel leakage detection and localization method based on line spectrum pair and cubic interpolation search [J]. Water Resources Management, 2020, 34(12): 3895-3911.
- [6] KANG J, PARK Y J, LEE J, et al. Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems [J]. IEEE Transactions on Industrial Electronics, 2018, 65(5): 4279-4289.
- [7] 杏运, 徐滢, 石琪琦. 分布式高精度时间同步装置专利技术现状[J]. 中国科技信息, 2024(11): 17-19.
- XING Y, XU Y, SHI Q Q. Current status of patent technology for distributed high-precision time synchronization devices [J]. China Science and Technology Information, 2024(11): 17-19.
- [8] 许龙霞, 刘娅, 李孝辉. GNSS 授时与时间同步在基础设施行业的应用[J]. 航空学报, 2024, 45(s1): 91-101.
- XU L X, LIU Y, LI X H. Applications of GNSS timing and time synchronization in infrastructures [J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(s1): 91-101.