

DOI:10.19853/j.zgjsps.1000-4602.2026.05.008

基于BiTCN和QRF的DMA用水量区间预测

黄海东¹, 刘世通², 吴美琼¹

(1. 广西水利电力职业技术学院 土木建筑工程学院, 广西 南宁 530023; 2. 北部湾大学
建筑工程学院, 广西 钦州 535011)

摘要: 独立计量区(DMA)用水量区间预测有助于供水管网管理者做出合理的漏损诊断。然而,如何保证较高预测区间覆盖率的同时具有较窄的平均宽度,是用水量区间预测需要解决的一个技术难点。为此,提出了一种融合双向时间卷积网络(BiTCN)和分位数回归森林(QRF)的区间预测模型。首先,基于用水量数据的特点,通过数据重构生成合适的样本集;然后,利用BiTCN模型提取用水量数据深层次的时序特征;接着,将所提取的特征输入QRF模型,初步构建BiTCN-QRF模型;最后,利用开普勒优化算法(KOA)优化BiTCN-QRF模型,并将优化后的模型用于区间预测。以广西都安县某DMA为例,对所提模型的有效性进行验证。结果表明,在95%、90%、85%、80%和70%置信水平下,所提模型均能实现高质量的区间预测,并且整体性能优于其他参与对比的模型。因此,该模型可作为基于流量的DMA实时漏损诊断的有效辅助工具。

关键词: 用水量区间预测; 双向时间卷积网络; 分位数回归森林; 开普勒优化算法; 独立计量区

中图分类号: TU991 **文献标识码:** A **文章编号:** 1000-4602(2026)05-0055-08

Interval Prediction for Water Demand in DMA Based on BiTCN and QRF

HUANG Haidong¹, LIU Shitong², WU Meiqiong¹

(1. School of Civil Engineering and Architecture, Guangxi Vocational College of Water Resources and Electric Power, Nanning 530023, China; 2. College of Civil Engineering and Architecture, Beibu Gulf University, Qinzhou 535011, China)

Abstract: Interval prediction for water demand in district metered area (DMA) is conducive to making appropriate leakage diagnosis for managers of water supply networks. However, ensuring higher coverage probability of the prediction interval while maintaining a lower average width remains a technical challenge in water demand interval prediction. To this end, an interval prediction model combining the bidirectional time convolutional network (BiTCN) with quantile regression forests (QRF) was proposed. First, based on the characteristics of water demand data, a suitable sample set was generated through data reconstruction. Then, the BiTCN model was used to extract deep-level temporal features of the data. Next, the extracted features were input into the QRF model to preliminarily construct the BiTCN-QRF model. Finally, the model was optimized using the Kepler optimization algorithm (KOA) and applied for interval prediction. The effectiveness of the proposed model was validated using data from a DMA located in Du'an County, Guangxi. The results showed that the proposed model could achieve

基金项目: 博士及高层次人才科研专项(2024kyzx01); 广西职业教育教学改革研究一般项目(GXGZJG2022B102)
通信作者: 吴美琼 E-mail: gxsl2023@163.com

high-quality interval prediction at 95%, 90%, 85%, 80%, and 70% confidence levels. Moreover, its overall performance was better than that of other comparative models. Based on these findings, the proposed model can serve as an effective auxiliary tool for flow-based real-time leakage diagnosis in DMA.

Keywords: interval prediction for water demand; bidirectional time convolutional network; quantile regression forests; Kepler optimization algorithm; district metered area

时间颗粒度为分钟级别的用水量预测是基于流量的独立计量区(DMA)实时漏损诊断的重要基础。为此,近年来分钟级别的DMA用水量预测受到科研人员的重视。Guo等^[1]将门控循环单元(GRU)和误差修正模块相结合对预测间隔为15 min的DMA用水量进行预测,发现所提方法的预测精度和稳定性均超过了季节性差分自回归滑动平均模型和传统的人工神经网络模型。Salloom等^[2]在上述模型的基础上提出了一种新的深度学习框架,获得了更高精度的预测结果。王晓婷等^[3]以长短时记忆网络(LSTM)为基本预测手段,实现了更低时间颗粒度的用水量预测。此外,一些学者尝试将卷积神经网络(CNN)与GRU进行融合以期达到更好的预测效果^[4]。

虽然分钟级别的DMA用水量预测已取得了一些研究成果,但已有研究大都集中在点预测。点预测会存在不同程度的误差,难以描述预测结果的不确定性,从而给实时漏损诊断工作带来一定困难。区间预测不仅能够提供预测范围,还能提供关于预测不确定性的信息。这有助于管理者更全面地理解预测信息,从而做出合理的漏损诊断。然而,在DMA用水量区间预测中,如何保证较高预测区间覆盖率的同时具有较窄的平均宽度,尚有待进行深入研究。

针对上述问题,笔者提出了一种融合双向时间卷积网络(BiTCN)和分位数回归森林(QRF)的区间预测模型——BiTCN-QRF模型。首先,利用BiTCN模块强大的特征提取能力对用水量数据深层次的时序特征进行提取;然后,将所提取的时序特征输入QRF模块以构建BiTCN-QRF模型;接着,利用开普勒优化算法(KOA)优化BiTCN-QRF模型;最后,通过优化后的BiTCN-QRF模型获取不同分位数下的预测结果,进而实现DMA用水量区间的准确预测。

1 研究方法

1.1 双向时间卷积网络

时间卷积网络(TCN)是一种强大且高效的深度学习模型,擅长处理具有时序特性的数据。结合卷积操作的优势,TCN能捕捉数据的长期依赖关系,在多种时序任务中均具有良好表现^[5-6]。TCN架构一般由膨胀因果卷积层、激活函数层、Dropout层、归一化层和残差连接层等组成,其中膨胀因果卷积层是TCN架构的核心组件。因果卷积确保TCN某一时刻的输出仅依赖于该时刻及其之前的输入;膨胀卷积则用于扩大卷积的感受野范围,从而在不增加网络深度的情况下覆盖更多的历史信息。

尽管TCN能很好地处理各种时序任务,但也存在局限性。例如,TCN只能从前一个时间点向后一个时间点进行卷积操作,无法获得时间序列的反向信息,这限制了模型对序列中长期依赖关系的捕获能力。针对上述问题,BiTCN构建了两个并行的TCN网络,一个用于处理正向信息,另一个用于处理反向信息,最后将捕获的正反信息进行融合后输出。这种信息传递的设计方式使得BiTCN能更好地理解序列数据的上下文信息,从而有效提取序列中的时序特征^[7]。对于一维数据 $\{x_0, x_1, \dots, x_t, x_{t+1}\}$,利用BiTCN提取时序特征的主要步骤如下:

① 进行正向时序特征提取。首先将一维数据输入因果卷积,经过一维卷积核训练后, t 时刻的隐藏层输出可用式(1)表示。

$$y_t = f(x_0, x_1, x_2, \dots, x_t) \quad (1)$$

式中: y_t 为隐藏层输出; $f()$ 为一维因果卷积核。

然后,将因果卷积的输出结果进行膨胀卷积处理,经过处理后 t 时刻的隐藏层输出见式(2)。

$$\vec{h}(t) = \sum_{i=0}^{\text{kernel}-1} f(i) \cdot x_{t-d+i} \quad (2)$$

式中: $\vec{h}(t)$ 为经过一次完整膨胀因果卷积后 t 时刻的正向序列隐藏层输出; $f(i)$ 为卷积核位置 i 的元素;kernel为卷积核的尺寸; d 为膨胀系数; x_{t-d+i} 为

对序列数据 $\{x_0, x_1, \dots, x_t, x_{t+1}\}$ 间隔采样后得到的元素。

② 进行反向时序特征提取。首先将一维数据 $\{x_0, x_1, \dots, x_t, x_{t+1}\}$ 进行翻转,生成反向序列数据 $\{x'_0, x'_1, \dots, x'_t, x'_{t+1}\}$,然后按照正向时序特征提取的方法依次进行因果卷积和膨胀卷积处理,得到 t 时刻的隐藏层输出,见式(3)。

$$\tilde{h}(t) = \sum_{i=0}^{\text{kernel}-1} f(i) \cdot x'_{t-d-i} \quad (3)$$

式中: $\tilde{h}(t)$ 为经过一次完整膨胀因果卷积后 t 时刻的反向序列隐藏层输出; x'_{t-d-i} 为对序列数据 $\{x'_0, x'_1, \dots, x'_t, x'_{t+1}\}$ 间隔采样后得到的元素。

③ 将正向和反向的时序特征进行融合,融合后得到的时序特征见式(4)。

$$h(t) = \text{concat}[\vec{h}(t), \tilde{h}(t)] \quad (4)$$

式中: $h(t)$ 为融合后 t 时刻的时序特征; concat 为拼接操作。

通过在正向和反向操作中堆叠多层膨胀因果卷积,能够有效提取序列数据更深层次的时序特征。图1为标准的BiTCN架构及多个BiTCN模块串联组成的BiTCN网络。

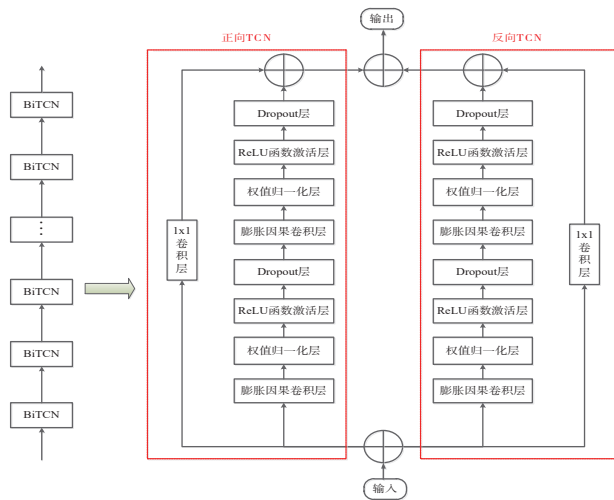


图1 BiTCN网络

Fig.1 BiTCN network

1.2 分位数回归森林

QRF是一种基于随机森林(RF)的衍生算法,它结合了分位数回归(QR)和RF两种算法的优势,具有鲁棒性强、预测效果好等特点,已成功用于水库调度、风力发电等领域^[8-9]。利用QRF算法进行区间预测的具体步骤如下^[10]:

① 根据RF算法,生成 k 棵决策树 $T(\theta_t)(t = 1, 2, \dots, k)$ 。

② 给定自变量 $D=d$,遍历 k 棵决策树,得到每棵决策树的叶子节点 $\ell(d, \theta_t)$ 。每棵决策树叶子节点 $\ell(d, \theta_t)$ 中,每个观测样本 $Y_i(i = 1, 2, \dots, n)$ 的权重见式(5)。

$$w_i(d, \theta_t) = \begin{cases} 1/m & d_i \in R_{\ell(d, \theta_t)} \\ 0 & d_i \notin R_{\ell(d, \theta_t)} \end{cases} \quad (5)$$

式中: $w_i(d, \theta_t)$ 为样本权重; $R_{\ell(d, \theta_t)}$ 为第 t 棵决策树叶子节点对应的特征空间; m 为叶子节点上的特征数量。

③ 对于RF算法而言,每个观测样本 $Y_i(i = 1, 2, \dots, n)$ 在 k 棵决策树上的权重均值见式(6)。

$$w_i(d) = k^{-1} \sum_{t=1}^k w_i(d, \theta_t) \quad (6)$$

式中: k 为决策树的数量; $w_i(d)$ 为观测样本 Y_i 的权重均值; $w_i(d, \theta_t)$ 为第 t 棵决策树叶子节点 $\ell(d, \theta_t)$ 上观测样本 Y_i 的权重。

④ 由 k 棵决策树上相应叶子中观测样本的加权平均近似即可得到RF算法对条件均值 $E(Y|D = d)$ 的估计,见式(7)。

$$\hat{\mu}(d) = \sum_{i=1}^n w_i(d) Y_i \quad (7)$$

式中: $\hat{\mu}(d)$ 为条件均值 $E(Y|D = d)$ 的估计; n 为观测样本的数量。

⑤ 通过观测样本的权重均值不断逼近目标值的条件分布,可得到目标值 y 的条件分布函数,其数学表达式见式(8)。

$$\hat{F}(y|D = d) = \sum_{i=1}^n w_i(d) 1_{\{Y_i \leq y\}} \quad (8)$$

式中: $\hat{F}(y|D = d)$ 为目标值 y 的条件分布函数; $1_{\{Y_i \leq y\}}$ 为所有观测值。

⑥ 根据式(8)可获得在 α 分位下的条件分位数估计,见式(9)。

$$Q_\alpha(d) = \inf \{ y: \hat{F}(y|D = d) \geq \alpha \} \quad (9)$$

式中: $Q_\alpha(d)$ 为 α 分位下的条件分位数估计; $\inf$ 为最大下界。

⑦ 给定不同的分位点 α ,根据上一步骤可获得多个条件分位数估计结果。利用这些结果可构造不同置信水平下的预测区间,从而实现区间预测。对于置信水平 β ,可通过式(10)构造出相应的

预测区间。

$$PI(d) = [Q_{0.5(1-\beta)}(d), Q_{0.5(1+\beta)}(d)] \quad (10)$$

式中: $PI(d)$ 为预测区间; $Q_{0.5(1-\beta)}$ 和 $Q_{0.5(1+\beta)}$ 分别为分位数 $0.5(1-\beta)$ 和 $0.5(1+\beta)$ 对应的条件分位数估计。

1.3 开普勒优化算法

KOA 算法是一种新型的基于天体运动原理的智能优化算法,通过模拟开普勒定律中行星围绕太阳的运动来解决优化问题。KOA 算法具有快速收敛、适应性强等优点^[10],主要算法步骤如下:

① 随机生成一组行星的初始位置和速度作为种群初始解。

② 根据每个星体(包括太阳和行星)的位置计算其相应的适应度。在 KOA 算法中,选择全局最优解作为太阳的位置。

③ 基于行星的当前位置和速度,模拟开普勒定律和万有引力来更新它们的新位置和速度。

④ 为获得最佳位置, KOA 算法实施精英策略,若位置更新后适应度更佳则更新位置,否则保留原位置。

⑤ 重复执行步骤②~④,直到满足终止条件。当算法达到终止条件时,最终的太阳位置即为算法的最优解。

1.4 评价指标

评价模型区间预测性能的常用指标有两种:预测区间覆盖率(PICP)和预测区间平均宽度(PINAW)。PICP是指真实值落入所给预测区间的概率,可以用来检验区间预测的可靠性。PINAW则反映了区间预测的准确度。在覆盖率一定的情况下, PINAW 越小,表示预测区间越窄,预测效果越好。因此,良好的区间预测应具有较大的 PICP 和较小的 PINAW。上述两个评价指标分别按式(11)和(12)计算^[11]。

$$PICP = \frac{1}{n} \sum_{i=1}^n k_i \quad \begin{cases} k_i = 1 & y_i \in [l_i, u_i] \\ k_i = 0 & y_i \notin [l_i, u_i] \end{cases} \quad (11)$$

$$PINAW = \frac{1}{nR} \sum_{i=1}^n (u_i - l_i) \quad (12)$$

式中: l_i 和 u_i 分别为第 i 个预测区间的下限和上限; R 为预测目标值的变化范围,用于对平均带宽进行归一化处理; n 为测试样本的数量。

在实际应用中,想要提高预测的可靠性往往要

以牺牲准确度为代价,因此时常会出现 PICP 和 PINAW 相冲突的现象。针对上述问题,引入综合评价指标 F 。该综合指标同时兼顾区间预测的可靠性和准确度, F 值越大,表明区间预测的质量越高。综合评价指标 F 按式(13)计算^[12]。

$$F = \frac{2 \times PICP \times \frac{1}{PINAW}}{PICP + \frac{1}{PINAW}} \quad (13)$$

1.5 基于 BiTCN-QRF 模型的用水量区间预测

针对 DMA 用水量区间预测问题,提出一种基于 BiTCN-QRF 模型的预测方法。下面详述所提方法应用于 DMA 用水量预测区间的主要步骤。

① 对原始用水量数据进行预处理,包括数据清洗和归一化处理。数据清洗的具体措施如下:采用前两天同一时刻用水量的平均值进行缺失值填充;对于异常值,则采用前两个用水量数据的平均值进行替换。数据清洗完成后,采用离差标准化方法对数据进行归一化处理。

② DMA 用水量数据具有短期相关性与周期性,因此为了能够有效刻画上述特性,对数据进行重构。通过用水量数据重构,一方面能够提高数据的信息利用度;另一方面可以避免过多数据参与建模带来的冗余信息。用水量数据重构后生成三个数据序列,分别用于刻画用水量数据的短期相关性、以 24 h 为周期的周期性及以 7 d 为周期的周期性。在数据重构的基础上,构建合适的样本。样本的构造方式如下:样本输出为 i 时刻的用水量数据 x_i ,其对应的样本输入包括三部分,分别为 i 时刻之前 n 个用水量数据点 $\{x_{i-1}, x_{i-2}, \dots, x_{i-n}\}$,用于描述短期相关性;前 j 天 i 时刻及其之前 $n-1$ 个用水量数据点 $\{x_{i-24*m*1}, \dots, x_{i-(24*m*j+n-1)}; \dots; x_{i-24*m*j}, \dots, x_{i-(24*m*j+n-1)}\}$,用于描述以 24 h 为周期的周期性, m 为每小时数据采样次数;前 q 周同一天 i 时刻及其之前 $n-1$ 个用水量数据点 $\{x_{i-168*m*1}, \dots, x_{i-(168*m*q+n-1)}; \dots; x_{i-168*m*q}, \dots, x_{i-(168*m*q+n-1)}\}$,用于描述以 7 d 为周期的周期性。参数 n 可以根据自相关性分析选取,而参数 j 和 q 可以结合试凑法进行确定。

③ 将样本集按一定比例划分为训练集、验证集和测试集。将 BiTCN 和 QRF 模型进行有效融合,利用训练集初步构建 BiTCN-QRF 模型。BiTCN 与 QRF 模型融合的实质其实就是利用 BiTCN 模型提

取用水量数据的时序特征,然后将所提取的特征作为QRF模型的输入,在此基础上得出预测结果。

④ 选取综合评价指标 F 作为适应度函数,利用KOA算法在验证集上优化BiTCN-QRF模型各模块的重要超参数。最大 F 值对应的超参数组合即为最优组合。涉及BiTCN模块需要进行寻优的重要超参数包括初始学习率、正则化系数、一维卷积层的过滤器宽度和个数等;对于QRF模块,需要优化的重要超参数有树的数量、每棵树的最大深度等。

⑤ 将优化后的BiTCN-QRF模型用于获取各分位数对应的最佳预测结果,在此基础上根据式(10)构造出不同置信水平下用水量的预测区间。

2 案例分析

2.1 数据描述

为了验证BiTCN-QRF模型的有效性,选取广西都安县某DMA的用水量数据作为数据集。该DMA地处新建开发区,采用单进水口无出水口的布置形式,主要向居民住宅区和少量工业企业供水,服务面积达1.8 km²,2023年12月—2024年2月平均日供水量约为1 200 m³。用水量数据通过进水口处的电磁流量计进行记录并上传至智慧水务管控云平台,采集时间为2023年12月1日—2024年2月29日,采集时间间隔为15 min,共获得8 736条数据。该DMA典型的一周用水量曲线如图2所示。

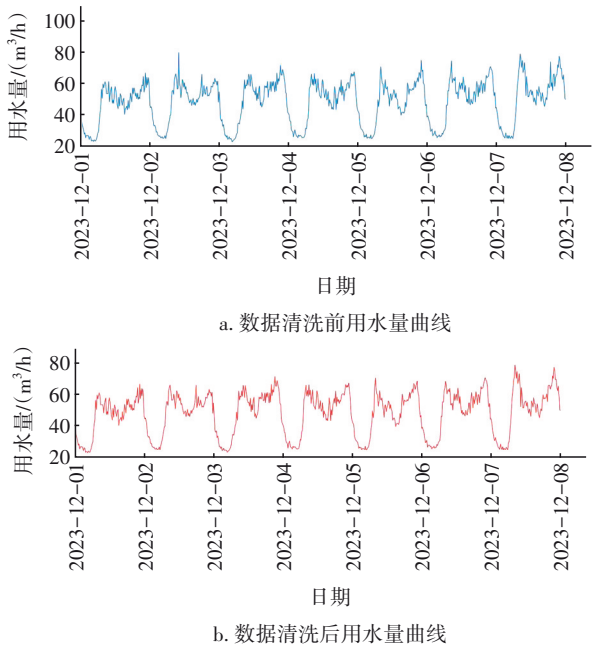


图2 一周用水量曲线

Fig.2 Water demand curves for a week

2.2 样本构建

分别在工作日和休息日随机选取某天用水量数据进行自相关性分析。根据自相关性分析结果,自相关系数大于0.5对应的滞后数为8,本研究将其作为 n 值,因此 $n=8$ 。由于数据的采样时间间隔为15 min,因此 $m=60/15=4$ 。根据自相关性分析结果和试凑法,在本研究中 j 和 q 的取值均为1。参数 n 、 m 、 j 和 q 确定后,即可获得样本的具体数据结构,进而生成合适的样本。

参照前述样本构造方式,共生成8 057个样本。将样本集按照6:2:2划分为训练集、验证集和测试集,可得到4 833个训练样本、1 612个验证样本和1 612个测试样本。

2.3 参数设置

为获得预测性能较好的BiTCN-QRF模型,需要找到相对较优的重要超参数组合。各模块的参数设置细节见表1。表1中没有涉及的超参数均采用默认值,本研究所有模型的测试均在Matlab 2023a环境下进行。

表1 各模块参数设置

Tab.1 Parameter settings of each module

模块	参数名称	设置值
BiTCN	初始学习率	[0.001,0.05]
	正则化系数	[0.000 1,0.000 5]
	一维卷积层过滤器个数	[16,64]
	一维卷积层过滤器宽度	[4,32]
	堆叠层数	1
	最大迭代次数	50
	网络训练优化器	Adam
QRF	梯度阈值	1
	树的数量	[10,200]
KOA算法	树的最大深度	[2,40]
	种群规模	5
	迭代次数	6

2.4 结果与分析

为全面评估所提模型的预测性能,设置两组对比实验。在第一组对比实验中,对比所提模型与QRF、TCN-QRF两个模型,旨在检验所提模型结构设计的有效性;在第二组对比实验中,选取高斯过程回归(GPR)和QR-LSTM两种常见的区间预测模型与所提模型进行对比,以此验证所提模型在区间预测上的优越性。为保证合理性,所有参与对比模型的重要超参数均采用相应优化算法进行寻优,且

相同超参数的寻优范围一致。

2.4.1 对比实验一

三种对比模型在不同置信水平下的区间预测结果见表2。

表2 不同置信水平下区间预测结果

Tab.2 Interval prediction results at different confidence levels

置信水平/%	模型	PICP	PINAW	F
95	QRF	0.982	0.317	1.497
	TCN-QRF	0.975	0.272	1.541
	BiTCN-QRF	0.983	0.279	1.542
90	QRF	0.954	0.259	1.530
	TCN-QRF	0.937	0.224	1.549
	BiTCN-QRF	0.956	0.241	1.553
85	QRF	0.922	0.224	1.529
	TCN-QRF	0.897	0.194	1.528
	BiTCN-QRF	0.921	0.210	1.543
80	QRF	0.891	0.197	1.515
	TCN-QRF	0.858	0.172	1.495
	BiTCN-QRF	0.883	0.186	1.518
70	QRF	0.798	0.158	1.417
	TCN-QRF	0.762	0.139	1.379
	BiTCN-QRF	0.796	0.150	1.422

在95%置信水平下,QRF、TCN-QRF和BiTCN-QRF模型的PICP分别达到0.982、0.975和0.983,明显高于给定的95%置信水平。在其他置信水平下,上述模型的PICP同样超过了相应置信水平。这表明以QRF为基础的预测模型可靠性都很高,可以提供有效的区间预测。

在85%、80%和70%置信水平下,QRF模型的PICP最大,而TCN-QRF模型的PICP最小。对于平均宽度指标,TCN-QRF模型的PINAW最小,QRF模型的PINAW最大。这说明QRF模型同时具有最高的可靠性和最差的准确度,而TCN-QRF模型的情况刚好相反。这种现象表明QRF和TCN-QRF模型存在以牺牲某一性能来提高另一性能的情况。在95%和90%置信水平下,PICP最大的是BiTCF-QRF模型,最小的是TCN-QRF模型。但在平均宽度指标上,TCN-QRF模型的PINAW最小,但最大的PINAW并非BiTCN-QRF模型,而是QRF模型。从总体上看,BiTCN-QRF模型的PICP始终超过给定的置信水平,但与此同时PINAW维持在相对较低的

水平,这一点与其他模型不同。这说明BiTCN-QRF模型能够较好地兼顾覆盖率和平均带宽,因此在充分保证可靠性的同时具有较高的准确度。从模型结构的角度看,出现上述现象的主要原因是BiTCN模型发挥了积极作用。与TCN模型相比,BiTCN模型提取深层次时序特征的能力更强,可以更好地弥补QRF模型在特征提取能力上的不足,这为后续获得更准确的预测结果提供了坚实基础。

从综合评价指标 F 来看,在五个不同置信水平下,BiTCN-QRF的 F 值普遍高于其他模型。在85%、80%和70%置信水平下,TCN-QRF模型表现最差;而在90%和95%的置信水平下,QRF模型的 F 值最小。基于以上分析可知,BiTCN-QRF模型区间预测质量最好,由此进一步验证了BiTCN-QRF模型在结构设计上的有效性。

为了更直观地观察BiTCN-QRF模型的预测性能,给出了不同置信水平下连续3 d共288个样本点的预测区间曲线,如图3所示。

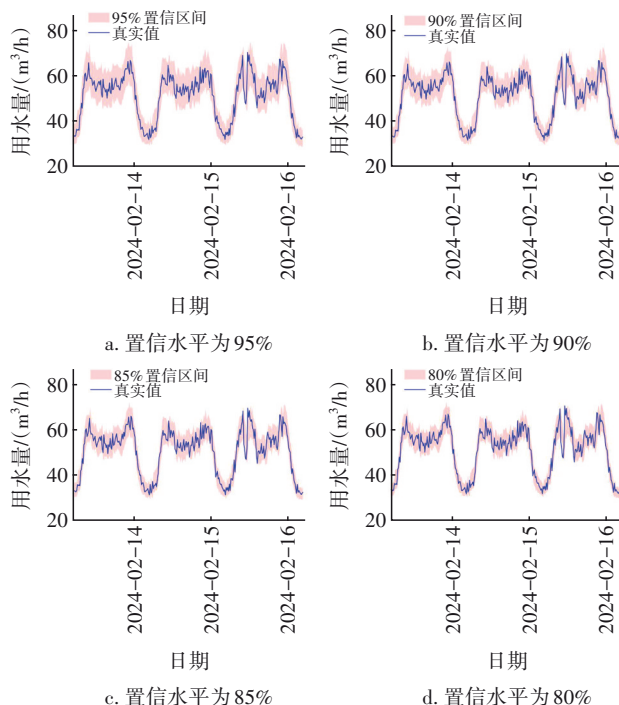


图3 不同置信水平下预测区间曲线

Fig.3 Prediction interval curves at different confidence levels

从图3可以明显看出,BiTCN-QRF模型的预测区间宽度较小,区间覆盖率较高,没有出现真实值明显偏离预测区间的现象。再次说明所提模型能够较好地同时兼顾可靠性和准确度,可提供可靠、

准确的用水量变化信息。

2.4.2 对比实验二

为进一步验证所提模型的优越性,首先结合分位数回归方法构建QR-LSTM组合模型,然后连同GPR模型与所提模型进行对比。表3给出了相同条件下三种模型的测试结果。

表3 相同条件下不同模型的测试结果

Tab.3 Test results of different models under same conditions

置信水平/%	模型	PICP	PINAW	F
95	GPR	0.981	0.380	1.434
	QR-LSTM	0.988	0.334	1.485
	BiTCN-QRF	0.983	0.292	1.527
90	GPR	0.956	0.319	1.465
	QR-LSTM	0.986	0.309	1.512
	BiTCN-QRF	0.956	0.241	1.553
85	GPR	0.931	0.279	1.478
	QR-LSTM	0.960	0.267	1.528
	BiTCN-QRF	0.921	0.210	1.543
80	GPR	0.896	0.248	1.466
	QR-LSTM	0.896	0.202	1.517
	BiTCN-QRF	0.883	0.186	1.518
70	GPR	0.819	0.201	1.406
	QR-LSTM	0.810	0.161	1.402
	BiTCN-QRF	0.796	0.150	1.422

从表3可以看出,在覆盖率指标上,三种模型的PICP均以较大幅度超过了给定的置信水平,说明这些模型能够提供可靠的预测区间。各置信水平下,BiTCN-QRF模型在覆盖率上均不占优势,但与其他模型的差距较小。在平均宽度指标上,BiTCN-QRF模型的效果最好。这说明BiTCF-QRF模型区间预测的准确度优于其他模型。

由于PICP和PINAW的结果相互冲突,因此结合综合指标F评价模型的区间预测性能。从F来看,BiTCN-QRF模型在各置信水平下始终表现最优,而GPR模型的整体效果较差。这是因为与GPR模型相比,BiTCN-QRF和QR-LSTM模型能够更有效提取用水量数据的深层次特征,从而获得更准确的预测结果。上述分析结果验证了BiTCN-QRF模型的优越性。

为了更好地观察评价指标的变化趋势,图4给出了三种模型不同评价指标随置信水平变化的趋势。

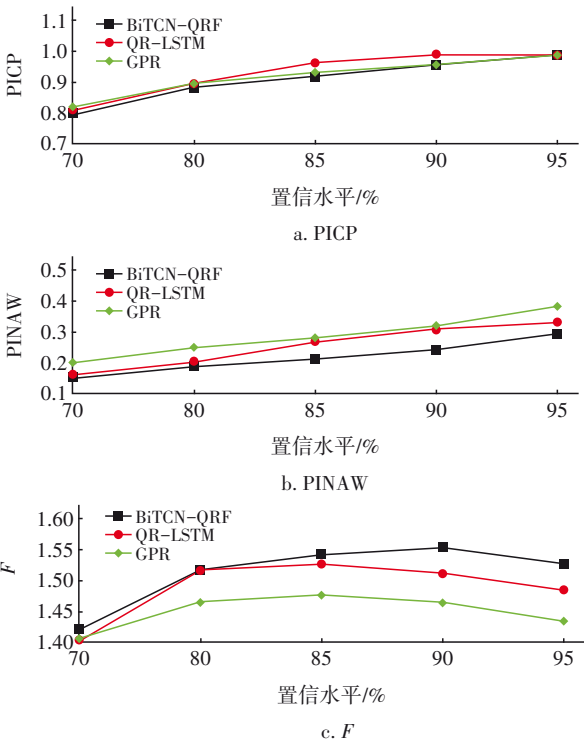


图4 不同置信水平下的评价指标
Fig.4 Evaluation indicators at different confidence levels

从图4(a)可以看出,随着置信水平的不断增加,BiTCN-QRF模型与其他模型的覆盖率差距不断缩小,在95%置信水平下,覆盖率的差距非常小。与此同时,BiTCN-QRF模型的平均带宽变化幅度明显小于其他模型,如图4(b)所示,而其他模型的平均带宽随置信水平的提高迅速增大。这表明BiTCN-QRF模型并没有以减小准确度为代价提高覆盖率,而其余模型保持较高的覆盖率是以牺牲准确度的代价的。由此可知,BiTCN-QRF模型区间预测的质量更高,这一点可以从图4(c)得到进一步印证。

3 结论

① BiTCN模型的融入有助于提高QRF模型的区间预测性能。BiTCN-QRF模型利用BiTCN模型在特征提取时的优势弥补QRF模型的不足,能够同时兼顾可靠性和准确度,实现稳定可靠的DMA用水量区间预测。在95%、90%、85%、80%和70%置信水平下,BiTCN-QRF模型的区间预测质量均优于QRF、TCN-QRF、GPR和QR-LSTM等模型。

② 本研究所提出的方法可以对DMA用水量的变化趋势和波动范围进行准确预测,这将为基于流量的DMA实时漏损诊断提供重要参考依据,具有

较高的实用价值。因此,所提方法有潜力成为DMA漏损管理的一种有效辅助手段。

参考文献:

- [1] GUO G C, LIU S M, WU Y P, et al. Short-term water demand forecast based on deep learning method [J]. Journal of Water Resources Planning and Management, 2018, 144(12): 4018076.
- [2] SALLOOM T, KAYNAK O, HE W. A novel deep neural network architecture for real-time water demand forecasting[J]. Journal of Hydrology, 2021, 599: 126353.
- [3] 王晓婷, 吴以朋, 刘书明, 等. 基于FFB-LSTM的供水计量区超短时水量预测方法研究[J]. 给水排水, 2023, 49(3): 133-139.
- WANG X T, WU Y P, LIU S M, et al. A flow fusion-based long short-term memory neural network for ultra-short-term water demand forecasting in district metering areas[J]. Water & Wastewater Engineering, 2023, 49(3): 133-139 (in Chinese).
- [4] CHEN L, YAN H X, YAN J R, et al. Short-term water demand forecast based on automatic feature extraction by one-dimensional convolution [J]. Journal of Hydrology, 2022, 606(1): 127440.
- [5] 刘少卿, 李帅, 苗建国, 等. 基于TCN-BiGRU的锂离子电池健康状态评估[J]. 电子测量技术, 2023, 46(23): 68-76.
- LIU S Q, LI S, MIAO J G, et al. Lithium-ion battery state of health estimation based on TCN-BiGRU [J]. Electronic Measurement Technology, 2023, 46(23): 68-76 (in Chinese).
- [6] SHENG W X, LIU K Y, JIA D L, et al. Short-term load forecasting algorithm based on LST-TCN in power distribution network[J]. Energies, 2022, 15(15): 5584.
- [7] 孙红哲, 王坚, 王鹏, 等. 基于Attention-BiTCN的网络入侵检测方法[J]. 信息安全, 2024, 24(2): 309-318.
- SUN H Z, WANG J, WANG P, et al. Network intrusion detection method based on Attention-BiTCN[J]. Netinfo Security, 2024, 24(2): 309-318 (in Chinese).
- [8] 戴领, 骆光磊, 周建中. 基于分位数回归森林的水库调度滚动模拟方法研究[J]. 人民长江, 2023, 54(7): 218-224.
- DAI L, LUO G L, ZHOU J Z. Rolling simulation method of reservoir operation based on quantile regression forest[J]. Yangtze River, 2023, 54(7): 218-224 (in Chinese).
- [9] 孙国强, 梁智, 俞娜燕, 等. 基于EWT和分位数回归森林的短期风电功率概率密度预测[J]. 电力自动化设备, 2018, 38(8): 158-165.
- SUN G Q, LIANG Z, YU N Y, et al. Short-term wind power probability density forecasting based on EWT and quantile regression forest[J]. Electric Power Automation Equipment, 2018, 38(8): 158-165 (in Chinese).
- [10] ABDEL-BASSET M, MOHAMED R, AZEEM S A A, et al. Kepler optimization algorithm: a new metaheuristic algorithm inspired by Kepler's laws of planetary motion [J]. Knowledge-based System, 2023, 268(3): 110454.
- [11] 陈静杰, 梁国栋, 刘家学. 航段截尾油耗数据的区间估计方法[J]. 计算机工程与设计, 2023, 44(10): 3103-3109.
- CHEN J J, LIANG G D, LIU J X. Interval estimation method for flight segment truncated fuel consumption data [J]. Computer Engineering and Design, 2023, 44(10): 3103-3109 (in Chinese).
- [12] DU B G, HUANG S, GUO J, et al. Interval forecasting for urban water demand using PSO optimized KDE distribution and LSTM neural networks[J]. Applied Soft Computing, 2022, 122(3): 108875.

作者简介: 黄海东(1982—), 男, 广西钦州人, 博士, 副教授, 主要研究方向为水资源管理与给水排水管网设计优化。

E-mail: 2447012621@qq.com

收稿日期: 2025-04-28

修回日期: 2025-06-27

(编辑: 任莹莹)