

基于时域卷积网络与 Transformer 的茶园蒸散量预测模型

赵秀艳¹ 王彬¹ 都晓娜² 王武闯³ 丁兆堂⁴ 周长安⁵ 张开兴⁵

(1. 山东农业大学信息科学与工程学院, 泰安 271018; 2. 潍柴雷沃智慧农业科技股份有限公司, 潍坊 261200;

3. 山东科润信息技术有限公司, 威海 264200; 4. 山东省农业科学院茶叶研究所, 济南 250100;

5. 山东农业大学机械与电子工程学院, 泰安 271018)

摘要: 在茶园水资源管理中, 蒸散量 (Evapotranspiration, ET) 是评估作物水分需求的关键指标, 由于茶园蒸散量预测具有时序性、不稳定性以及非线性耦合等特点, 目前的茶园蒸散量预测模型存在预测精度较低的问题, 针对此问题本文提出了一种新型的茶园蒸散量预测模型。首先使用互信息算法 (Mutual information, MI) 与主成分分析算法 (Principal component analysis, PCA) 相融合的数据处理算法 (MIPCA), 筛选强相关的特征并提取主成分; 其次将时域卷积网络 (Temporal convolutional network, TCN) 与 Transformer 融合, 利用灰狼算法 (Grey wolf optimization, GWO) 优化超参数, 捕捉茶园数据的全局依赖关系; 最后整合 2 个网络构建了 MIPCA-TCN-GWO-Transformer 模型, 通过消融试验和对比试验验证了模型性能, 并对模型在不同时间步长下的性能进行测试。结果表明, 该模型平均绝对百分比误差 (Mean absolute percentage error, MAPE)、均方根误差 (Root mean square error, RMSE) 和决定系数 (Coefficient of determination, R^2) 3 个评价指标分别为 0.015 mm/d、0.312 mm/d 和 0.962, 优于长短期记忆模型 (Long short term memory, LSTM) 等传统预测模型。在小时尺度、日尺度和月尺度下的 R^2 分别为 0.986、0.978 和 0.946, 在不同时间步长下展现了良好的适应性和准确性。本文构建的 MIPCA-TCN-GWO-Transformer 模型具有较高的预测精度和稳定性, 可为茶园水资源优化管理和灌溉制度制定提供科学参考。

关键词: 茶园; 蒸散量; 预测模型; 主成分分析; 互信息; 时域卷积网络

中图分类号: S161.4; S571.1 文献标识码: A 文章编号: 1000-1298(2024)09-0337-10

OSID:



Evapotranspiration Prediction Model of Tea Garden Based on Temporal Convolutional Network and Transformer

ZHAO Xiuyan¹ WANG Bin¹ DU Xiaona² WANG Wuchuang³ DING Zhaotang⁴

ZHOU Chang'an⁵ ZHANG Kaixing⁵

(1. College of Information Science and Engineering, Shandong Agricultural University, Taian 271018, China

2. Weichai Lovol Intelligent Agricultural Technology Co., Ltd., Weifang 261200, China

3. Shandong Kerun Information Technology Co., Ltd., Weihai 264200, China

4. Tea Research Institute, Shandong Academy of Agricultural Sciences, Ji'nan 250100, China

5. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China)

Abstract: In tea garden water resource management, accurately assessing crop water requirements is crucial, with evapotranspiration (ET) serving as a key indicator. The challenges posed by the time series nature, instability, and non-linear coupling in tea garden data were addressed by introducing a novel evapotranspiration prediction model. Firstly, a data processing algorithm, mutual information-principal component analysis (MIPCA), was employed to integrate mutual information (MI) and principal component analysis (PCA), facilitating the selection of features strongly correlated with tea garden transpiration and the extraction of principal components. Subsequently, the temporal convolutional

收稿日期: 2023-12-19 修回日期: 2024-03-19

基金项目: 山东省科技型中小企业创新能力提升工程项目 (2022TSGC2487、2023TSGC0557)、日照市重点研发计划项目 (2023ZDYF010129) 和泰安市科技创新重大专项项目 (2023NYLZ13)

作者简介: 赵秀艳 (1977—), 女, 副教授, 博士, 主要从事农业机械化与信息化研究, E-mail: zhaoxy@sdau.edu.cn

通信作者: 张开兴 (1984—), 男, 教授, 博士, 主要从事农业机械化与信息化研究, E-mail: kaixingzhang@139.com

networks (TCN) was integrated with Transformer to construct a new model. Specifically, the grey wolf optimization (GWO) algorithm was employed to optimize the hyperparameters of the TCN, followed by the utilization of the Transformer to capture global dependencies. Ultimately, the two networks were integrated to propose the hybrid model MIPCA – TCN – GWO – Transformer. The model performance was validated through ablation experiments and comparative analyses, while also examining the model's performance across different time scales. The results showed that the model's three evaluation indicators such as mean absolute percentage error (MAPE), root mean square error (RMSE) and coefficient of determination (R^2) were 0.015 mm/d, 0.312 mm/d and 0.962, respectively, which as better than that of traditional prediction models such as long short term memory (LSTM). R^2 at hourly scale, daily scale and monthly scale were 0.986, 0.978 and 0.946, respectively, showing good adaptability and accuracy at different time scales. The MIPCA – TCN – GWO – Transformer model constructed had high prediction accuracy and can provide scientific reference for the optimal management of tea garden water resources and the formulation of irrigation systems.

Key words: tea garden; evapotranspiration; prediction model; principal component analysis; mutual information; time convolutional network

0 引言

近年来,随着我国茶树栽培技术逐渐完善,北方茶区种植面积不断扩大,但由于北方降雨分配不均,季节性的干旱和水资源的不足成为茶叶产量和品质的主要限制因素^[1-2],因此北方茶园需要有效的水资源管理策略。根据联合国粮农组织 (Food and Agriculture Organization of the United Nations, FAO) 的研究,作物蒸散量是反映作物耗水量的重要指标^[3-4]。FANG 等^[5]利用机器学习方法构建茶树日参考蒸散量预测模型,提高了茶树的生产效率和适应性,证明蒸散量预测模型具有实用价值。

在茶园数据处理方面,为分析输入特征与输出特征的相关性,文献[6-7]使用互信息法评估变量间的非线性关系,确定了与参考蒸散量高度相关的输入特征,精简了模型输入集,提高了模型的效率和准确性;为过滤数据集噪声和提取重要特征,ZHAO 等^[8]使用 PCA 算法在 7 种气象因子中提取 5 个主成分作为支持向量机 (Support vector machine, SVM) 的输入,以降低数据冗余度;MANIKUMARI 等^[9]使用 PCA 算法对印度 Veeranam tank 地区的 6 个气象因子降维,并提取 3 个主成分用于蒸散量的预测,提高了预测精度。因此,使用互信息算法筛选与蒸散量呈强相关的输入特征,以主成分分析算法选择数据中的主成分,可提高模型的性能。

在蒸散量模型构建方面,目前机器学习模型在农作物蒸散量预测的方法及研究成果较多,其能够有效地捕捉作物生长过程中复杂的非线性特征,从而实现自动且准确的建模。谢家兴等^[10]构建了 LSTM 模型提取柑橘园时空序列特征,在日参考蒸散量的预测方面展现出了较好的性能;BABAK 等^[11]将支持向量回归 (Support vector regression,

SVR) 与鲸鱼优化算法 (Whale optimization algorithm, WOA) 相结合,对伊朗 3 个气象站的日参考蒸散量进行建模,验证了该模型的有效性和准确性;WU 等^[12]使用粒子群算法 (Particle swarm optimization, PSO) 优化的极限学习机 (Extreme learning machine, ELM) 估算中国不同气候带的参考蒸散量,试验结果表明该模型具有较好的预测效果;WU 等^[13]以华南鄱阳湖流域为研究区域,将 K 均值聚类 (K-means clustering, Kmeans) 和萤火虫算法 (Firefly algorithm, FFA) 引入核极限学习机 (Kernel extreme learning machine, KELM), 构建了 Kmeans – FFA – KELM 模型,与常用的机器学习模型对比效果更优;PETKOVIC 等^[14]构建了径向基函数网络 (Radial basis function network, RBFN), 并使用粒子群算法进行优化,来预测塞尔维亚 1980—2010 年期间的参考蒸散量,结果表明该模型具有良好的统计特性,有助于参考蒸散量估算;YU 等^[15]提出粒子群算法优化的极限梯度提升模型 (Extreme gradient boosting, XGBoost), 对太阳温室的参考蒸散量进行预测,结果表明该模型可以在各种数据模式下准确估计参考蒸散量。

为进一步提高模型对作物蒸散量的预测精度,研究者采用特征提取和超参数优化等策略。时序特性的准确提取对于模型训练至关重要,CHEN 等^[16]使用时域卷积网络对地膜覆盖滴灌情况下玉米长期序列特征进行提取,并证明该模型在估算参考蒸散量方面具有较好的性能。注意力机制能够调整模型对不同时间步或位置的关注度,从而使模型更好地捕捉时序数据中的重要模式。田原等^[17]提出 Transformer 与 TCN 结合的滑坡短期位移情况预测模型,Transformer 在序列建模中具有良好的并行性和全局注意力,弥补了 TCN 固定感受野和局部关注

的局限,试验结果具有较高的准确度。

对于复杂的数据模式,算法的参数会限制模型的准确程度,针对这一问题,研究者考虑通过优化算法调整模型的超参数来适应复杂的数据分布。PANGAM等^[18]使用灰狼优化算法(GWO)对机器学习算法的超参数进行优化,提高了模型的全局搜索能力和收敛速度,试验结果表明,优化后的机器学习模型在参考蒸散量的准确预测方面优于传统模型。灰狼优化算法具有收敛速度快、适用性高、全局搜索能力强的特点。因此针对TCN模型易陷入局部最优解且收敛速度慢的问题,使用GWO优化算法迭代搜索空间,优化模型超参数。

综上所述,本文针对茶园数据时序性、不稳定性以及非线性耦合等特点导致的预测精度低问题,拟提出一种茶园蒸散量预测模型。使用互信息算法与主成分分析算法相融合的数据处理算法,筛选强相关特征并提取主成分。将时域卷积网络与Transformer融合构建新模型,利用GWO优化TCN超参数,捕捉时序特征,进而与Transformer融合,考虑多参数间影响。最终将2个网络集成,构建MIPCA-TCN-GWO-Transformer模型,以期能为茶园蒸散量预测提供科学依据。

1 材料与方法

数据来源于泰安市生态种植茶园(117.43°E, 36.08°N),茶园地处暖温带半湿润大陆性季风气候,平均海拔135.2 m,年均降雨量664.5 mm,年均光照时数2 663.8 h,茶园位置及实地情况如图1所示。试验地土壤为微酸性红壤土,茶树品种为鸠坑茶,采用单栋双坡面式温室,以双行条植方式种植。使用综合气象监测站采集光照时长(L_d)、空气温度(A_t)、空气相对湿度(A_h)、二氧化碳含量(C_d)和降雨量(R_f)等环境数据,光照传感器、二氧化碳传感器和雨量采集器距地1.5 m,温湿度传感器距地1.8 m;使用土壤墒情监测仪采集土壤温度(S_t)、土壤含水率(S_m)、土壤EC值(S_e)和土壤pH值(S_p)等土壤数据,土壤传感器均放置于地下60 cm处。气象数据来源于中国气象数据网,包括日最高温度(M_t)、日最低温度(L_t)、日平均风速(A_s)、最大持续风速(M_s)、压强(P_s)和露点温度(D_t)等。

数据采集自2021年3月21日—4月30日,频率为30 min采集1次,剔除缺失数据及异常数据后,共采集33 276条有效数据。根据上述数据,采用王健等^[19]提出的计算公式,针对设施内风速为0的情况进行日参考蒸散量的计算,计算公式为

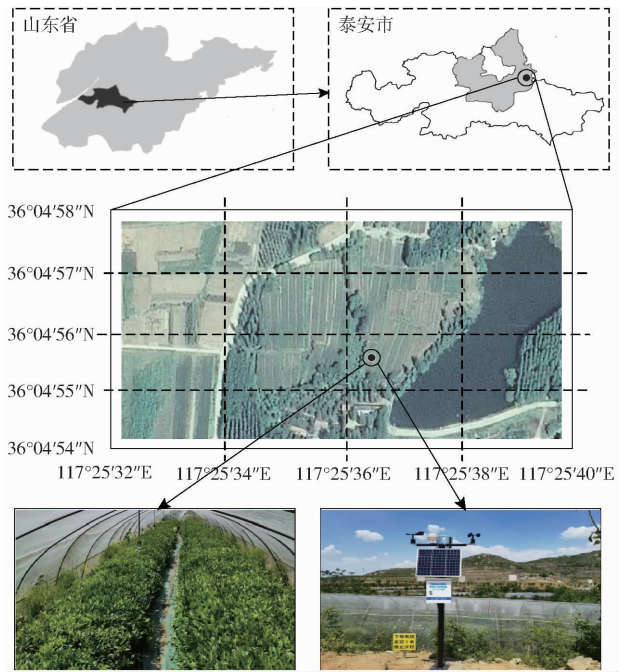


图1 茶园位置及实地图

Fig. 1 Location and actual map of tea garden

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{1\,713(e_s - e_a)}{T + 273}}{\Delta + 1.64\gamma} \quad (1)$$

式中 ET_0 ——日参考蒸散量,mm/d
 R_n ——净辐射量,MJ/($m^2 \cdot d$)
 G ——土壤热通密度,MJ/($m^2 \cdot d$)
 T ——2 m 高度平均空气温度,℃
 e_s ——饱和蒸汽压强,kPa
 e_a ——实际蒸汽压强,kPa
 Δ ——蒸汽压力曲线斜率,kPa/℃
 γ ——温湿度常数,kPa/℃

作物实际需水量 ET_c 由日参考蒸散量 ET_0 与作物系数 K_c 相乘而得,即

$$ET_c = K_c ET_0 \quad (2)$$

文献[20-21]提出温室环境内作物系数有低于大田环境内作物系数的趋势,因此本文使用较低的植物系数值($K_c = 0.95$)。为降低各输入特征数据量纲不同带来的影响,对所有茶园数据进行归一化处理,将训练样本的所有数据转换到[0,1]区间内,归一化计算公式为

$$X_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

式中 X_i ——归一化值
 x_i ——样本值
 $x_{\max}$ ——样本最大值
 $x_{\min}$ ——样本最小值

本文采用时间序列划分法对数据集进行划分,按照时间先后顺序划分训练集和测试集。将2021

年 3 月 21 日—2022 年 4 月 30 日的茶园数据作为训练集,2022 年 5 月 1 日—7 月 31 日的茶园数据作为测试集,划分比例约为 8:2。并将 2022 年 5 月 1 日—2023 年 4 月 30 日的茶园数据用于不同时间步长下模型性能的测试。

2 茶园蒸散量预测模型

2.1 模型构建

茶园蒸散量预测模型的构建分为 4 个主要步骤,即茶园特征处理、局部特征提取、超参数优化和全局关联学习。茶园特征处理涵盖特征选择和降维,以减少数据噪声和复杂性;局部特征提取使用 TCN 模型提取茶园特征中的时序信息;超参数优化通过 GWO 优化 TCN 中的超参数,提高收敛速度,优化模型性能;全局关联学习则通过 Transformer 捕捉序列中不同位置的依赖关系,提高模型的全局理解能力。

2.1.1 茶园特征处理

茶园特征处理中,使用互信息算法评估原始茶园参数之间的关系,并筛选出与输出参数相关性较高的特征。使用主成分分析算法对筛选后的特征降维,以提取数据中的主要信息。

首先,计算原始数据中各个特征与输出参数之间的互信息。互信息用于分析变量之间的信息传递程度,从而评估它们之间的相关性和非线性关系,互信息计算公式为

M_I(X,Y)=\int\int dx dy \mu(x,y) \lg \frac{\mu(x,y)}{\mu(x)\mu(y)} \tag{4}

式中 $\mu()$ ——概率分布函数

通过计算互信息,得到各个特征与输出参数的相关性分数,基于互信息分数,选择与输出参数相关性较高的特征,形成一个新的特征子集。将原始数据进行中心化,即减去均值,得到中心化数据。然后,计算中心化后的样本在各个特征下的协方差矩阵。这个协方差矩阵表示样本特征之间的相关性。协方差矩阵计算公式为

C=\frac{1}{i}(\overline{X}\overline{X}^T) \tag{5}

式中 i ——样本数量

$\overline{X}$ ——中心化后的样本

对协方差矩阵 C 进行特征值分解,得到特征值 λ 和对应的特征向量 μ 。特征值表示投影后数据在特征向量方向上的方差,特征向量则表示投影方向,特征值分解计算公式为

C\mu=\lambda\mu \tag{6}

通过将原始数据 X 与特征向量 μ 正交,将原始

数据映射到一个新的特征空间 X' ,计算公式为

X'=X\mu \tag{7}

式中 X' ——降维后的新特征空间,每一列表示一个主成分

2.1.2 局部特征提取

由于茶叶生长与气候、季节等时间因素密切相关,使用 TCN 对输入特征中的时间序列信息进行提取。TCN 在卷积神经网络 (Convolutional neural network,CNN) 基础上加入了因果卷积和扩张卷积,并额外引入一层扩张卷积,确保输入输出为形状相同的张量,TCN 卷积网络结构如图 2 所示,图中 d 为膨胀系数, k 为卷积核大小。

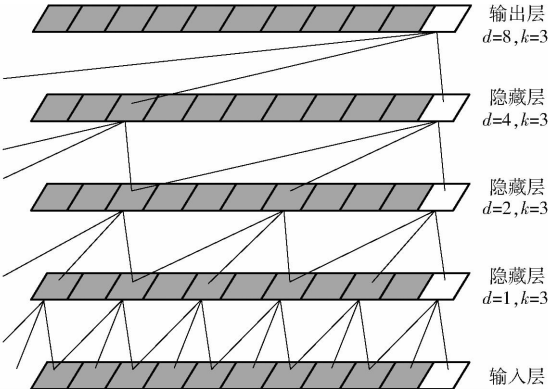


图 2 TCN 卷积网络结构图

Fig. 2 TCN convolutional network structure diagram

将经过特征筛选的数据 X' 作为 TCN 的输入,通过卷积层捕捉序列中的局部特征,卷积操作生成特征映射,然后应用非线性激活函数,堆叠多个卷积层以增强模型复杂度。最后一个卷积层的输出 H_{tcn} 作为模型最终输出。

因果卷积为单向结构,各个时间步仅考虑当前和过去信息,避免了未来信息泄露,确保模型符合真实数据的因果关系^[22]。同时,茶树的生长过程会受到较远时间范围内的因素影响,传统的卷积操作难以捕捉这种远距离关系,针对此问题 TCN 引入扩张卷积,通过增加卷积核的感受野,提高模型捕捉长距离依赖关系的能力,在不加深网络层次的同时提高其计算效率^[23]。

对于序列中的元素 s ,扩张卷积操作 F 定义为

F(s)=(x*f)(s)=\sum_{i=0}^{k-1} f(i)x_{s-di} \tag{8}

式中 $*$ ——卷积运算符号

x_{s-di} ——扩张卷积的第 i 个数据

$f(i)$ ——卷积核中的第 i 个元素

图 2 中 TCN 网络设置 4 个膨胀系数 d 为 1、2、4、8 的卷积层,膨胀系数 $d=1$ 时,表示对各输入点采样,当 $d=2$ 时,表示对下层数据每隔 1 个数据点

采样,以此类推,当 $d = n$ 时,表示对下层数据隔 $n - 1$ 个数据点采样。扩张卷积的感受野指数级增长能够更好地捕捉茶园特征的时序模式和变化趋势。本文扩张卷积中的卷积核大小 k 设置为 3,各卷积核同时考虑当前时间步以及前后各一个时间步的信息。在不增加参数和计算量的情况下增加卷积层的感受野,更好地捕捉茶园特征的长程依赖关系。

为解决网络深度增加时的梯度消失问题,TCN 采用残差链接策略。各残差块包含多个具有相同卷积核大小的卷积层,卷积层之间通过 ReLU 非线性激活函数连接。在每个残差块内部,输入数据经过一系列卷积和激活操作,最后通过跳跃式连接与原始输入相加,形成了残差链接,残差模块结构如图 3 所示。

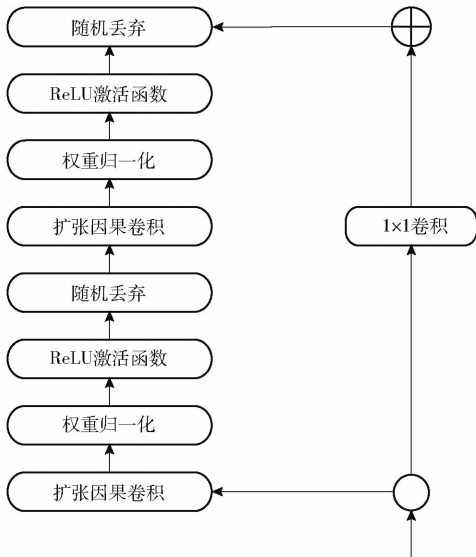


图3 残差模块结构图

Fig.3 Residual module structure diagram

2.1.3 超参数优化

采用灰狼优化算法对 TCN 模型进行超参数优化,包括卷积核数量、学习率、训练轮数和批量大小等。卷积核数量设置在 16 ~ 256 之间,学习率范围在 0.001 ~ 0.1 之间,训练轮数在 50 ~ 500 之间,批量大小在 16 ~ 128 之间。

初始随机生成一组灰狼,每只代表一个潜在的超参数解,利用灰狼优化算法的机理找到最佳超参数组合,以获得最佳模型性能。灰狼优化算法优化过程将包围、追捕、攻击 3 个阶段的任务分配给各等级的灰狼群来完成捕食行为,从而实现全局优化的搜索过程^[24]。

在包围阶段,灰狼群通过计算与猎物的距离来确定猎物的位置。距离 D 计算公式为

$$\mathbf{X}_d(t+1) = \mathbf{X}_p(t+1) - \mathbf{A}\mathbf{D} \quad (9)$$

$$D = |\mathbf{C}_p \mathbf{X}_p(t) - \mathbf{X}_d(t)| \quad (10)$$

其中

$$\mathbf{A} = 2a\mathbf{r}_1 - a \quad (11)$$

$$\mathbf{C}_p = 2\mathbf{r}_2 \quad (12)$$

式中 $\mathbf{X}_d$ ——灰狼位置

t ——当前迭代次数

$\mathbf{X}_p$ ——猎物的位置

$\mathbf{A}$ 、 $\mathbf{C}_p$ ——协同系数向量

$\mathbf{r}_1$ 、 $\mathbf{r}_2$ —— $[0, 1]$ 内的随机一维分量

a ——收敛因子,取值随着迭代次数的增大从 2 到 0 线性递减

在追捕阶段,灰狼个体根据距离来调整其位置,分别计算与 α 、 β 和 δ 狼的距离 D_α 、 D_β 和 D_δ ,更新位置的计算公式为

$$D_\alpha = |\mathbf{C}_1 \mathbf{X}_\alpha - \mathbf{X}_d| \quad (13)$$

$$D_\beta = |\mathbf{C}_2 \mathbf{X}_\beta - \mathbf{X}_d| \quad (14)$$

$$D_\delta = |\mathbf{C}_3 \mathbf{X}_\delta - \mathbf{X}_d| \quad (15)$$

$$\mathbf{X}_1 = \mathbf{X}_\alpha - \mathbf{A}_1(D_\alpha) \quad (16)$$

$$\mathbf{X}_2 = \mathbf{X}_\beta - \mathbf{A}_2(D_\beta) \quad (17)$$

$$\mathbf{X}_3 = \mathbf{X}_\delta - \mathbf{A}_3(D_\delta) \quad (18)$$

式中 $\mathbf{C}_1$ 、 $\mathbf{C}_2$ 、 $\mathbf{C}_3$ ——随机数

$\mathbf{X}$ ——灰狼个体当前位置

$\mathbf{X}_\alpha$ 、 $\mathbf{X}_\beta$ 、 $\mathbf{X}_\delta$ —— α 狼、 β 狼、 δ 狼当前位置

$\mathbf{X}_1$ 、 $\mathbf{X}_2$ 、 $\mathbf{X}_3$ —— ω 狼分别受 α 狼、 β 狼、 δ 狼影响后的位置

在攻击猎物阶段,灰狼个体根据追捕阶段的调整来确定最终位置,通常取位置平均值^[25],计算公式为

$$\mathbf{X}_d(t+1) = \frac{\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{X}_3}{3} \quad (19)$$

2.1.4 全局关联学习

全局关联学习主要由 Transformer 编码器实现,将经过 GWO 优化的 TCN 在茶园数据集上的训练结果 $\mathbf{H}_{\text{tcn}}$ 输入到 Transformer 编码器中。首先在 $\mathbf{H}_{\text{tcn}}$ 的基础上加入正余弦位置编码,位置编码通过使用不同频率的正弦、余弦函数生成,然后和对应位置的输入向量相加,位置编码计算公式为

$$E_{PE}(\mathbf{p}_{\text{pos}}, 2i) = \sin \frac{\mathbf{p}_{\text{pos}}}{10\,000^{\frac{2i}{d_{\text{model}}}}} \quad (20)$$

$$E_{PE}(\mathbf{p}_{\text{pos}}, 2i+1) = \cos \frac{\mathbf{p}_{\text{pos}}}{10\,000^{\frac{2i}{d_{\text{model}}}}} \quad (21)$$

$$\mathbf{H}'_{\text{tcn}} = \mathbf{H}_{\text{tcn}} + E_{PE} \quad (22)$$

式中 $\mathbf{p}_{\text{pos}}$ ——输入向量在 $\mathbf{H}_{\text{tcn}}$ 矩阵中的绝对位置

d_{model} ——输入向量维度

$2i$ 、 $2i+1$ ——奇偶性标识

E_{PE} ——位置编码

$\mathbf{H}'_{\text{tcn}}$ ——加入位置编码后的训练结果

然后通过 3 个权值矩阵 $\mathbf{W}_Q$ 、 $\mathbf{W}_K$ 、 $\mathbf{W}_V$,将 $\mathbf{H}'_{\text{tcn}}$ 转

换成为计算注意力权重所需的 Query、Keys、Values 向量 Q 、 K 、 V 。通过计算各输入向量的相关性得分并进行归一化,使用 softmax 函数将每个单词之间的得分向量转换成 $[0,1]$ 之间的概率分布。然后,根据输入向量的概率分布乘以对应的 Values 值得到注意力权重 Z ,再通过前馈神经网络来叠加操作得到 Transformer 的输出 Y ,计算公式为

$$[Q,K,V] = H'_{\text{ten}}[W_Q,W_K,W_V]$$

(23)

$$s_{\text{core}} = QK^T$$

(24)

$$Z = \text{softmax}(s_{\text{core}}/\sqrt{d_k})V$$

(25)

$$Y = \text{FFN}(Z) = \max(0,ZW_1 + b_1)W_2 + b_2$$

(26)

式中 s_{core} ——相关性得分
 W_1 、 W_2 ——权值矩阵
 b_1 、 b_2 ——偏置项
FFN()——前馈神经网络
 d_k ——向量 Q 、 K 、 V 的维度,通过数值缩放 $\sqrt{d_k}$,控制相关性得分,避免维度增加导致的梯度消失

最后对茶园蒸散量进行预测,Transformer 结构如图 4 所示。

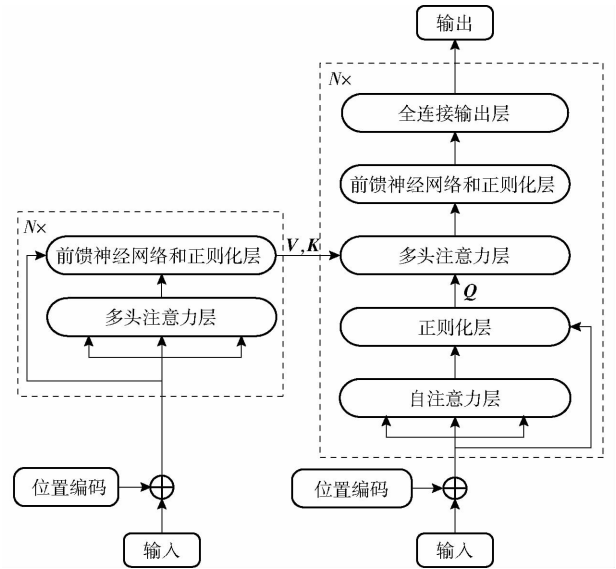


图 4 Transformer 结构图

Fig. 4 Transformer structure diagram

2.2 模型评价指标和参数设置

通过茶园数据构建蒸散量预测模型,利用灰狼优化算法对模型中超参数进行优化,来提高模型的预测精准度。经迭代优化后得到 TCN 模型最优参数: K_n 为 256, L_r 为 0.01, E_n 为 50, B_s 为 64。

为了评估混合模型对蒸散量的预测性能,使用 RMSE、MAPE 和 R^2 等 3 个评价指标评价模型预测性能, R^2 越大, RMSE 与 MAPE 越小表示模型预测性能越好。

3 结果与分析

3.1 特征筛选

对茶园环境、土壤以及气象等数据进行分析,计算得到了茶园特征的均值和标准差,详细描述如表 1 所示。为降低茶园冗余度,对归一化后的茶园数据使用互信息法分析各输入特征与标签值的相关性,分析结果如图 5 所示。

表 1 茶园数据均值及标准差

Tab. 1 Mean and standard deviation of tea garden data

指标	均值	标准差
L_d/h	7.274	1.696
$A_t/^\circ\text{C}$	15.9	8.786
$A_h/\%$	68.178	18.65
$M_t/^\circ\text{C}$	22.071	8.105
$L_t/^\circ\text{C}$	7.964	7.235
$S_t/^\circ\text{C}$	15.797	5.156
$S_m/\%$	26.409	8.16
$A_s/(\text{m}\cdot\text{s}^{-1})$	5.798	2.286
$M_s/(\text{m}\cdot\text{s}^{-1})$	12.649	6.529
R_f/mm	2.101	7.364
P_s/hPa	986.402	61.456
$D_t/^\circ\text{C}$	9.969	3.294
$C_d/(\text{mg}\cdot\text{m}^{-3})$	606.92	42.831
$S_e/(\mu\text{S}\cdot\text{cm}^{-1})$	195.294	20.067
S_p	6.045	0.276
$ET_c/(\text{mm}\cdot\text{d}^{-1})$	6.549	2.519

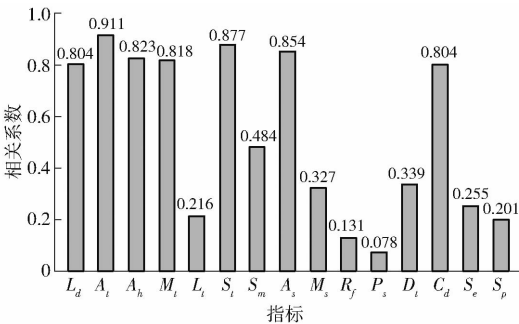


图 5 相关性分析图

Fig. 5 Correlation analysis diagram

由分析结果可知, ET_c 与 L_d 、 A_t 、 A_h 、 M_t 、 S_t 、 A_s 、 C_d 等相关系数大于 0.8, 呈强相关性。为降低预测模型的复杂度,利用 PCA 算法对茶园特征进行降维,茶园特征的主成分系数矩阵如表 2 所示,从主成分系数矩阵分析得出,前 5 个主成分累计贡献率为 88.81%,其能够反映原始数据的绝大部分信息,将以上主成分作为模型的输入。

3.2 消融试验

为验证模型有效性,以 2022 年 5 月 1 日—7 月 31 日的数据集作为测试集,将 MIPCA-TCN-GWO-Transformer 模型与其他 4 组消融试验模型进行对

表 2 成分矩阵					
Tab.2 Component matrix					
指标	成分 1	成分 2	成分 3	成分 4	成分 5
光照时长	0. 673	0. 812	-0. 080	0. 573	-0. 058
空气温度	0. 1245	-0. 857	0. 116	0. 284	-0. 089
空气相对湿度	0. 809	-0. 412	-0. 003	0. 219	0. 201
最高温度	-0. 229	0. 421	0. 864	-0. 348	-0. 283
土壤温度	-0. 129	0. 056	0. 545	-0. 406	0. 842
平均风速	0. 839	-0. 443	0. 069	-0. 052	-0. 097
二氧化碳含量	0. 730	-0. 761	-0. 551	-0. 971	0. 603
特征值	3. 840	2. 821	2. 172	1. 291	0. 483
累计贡献率/%	32. 81	55. 46	74. 03	85. 05	88. 81

比,其拟合图及误差图如图 6 所示,消融试验结果如

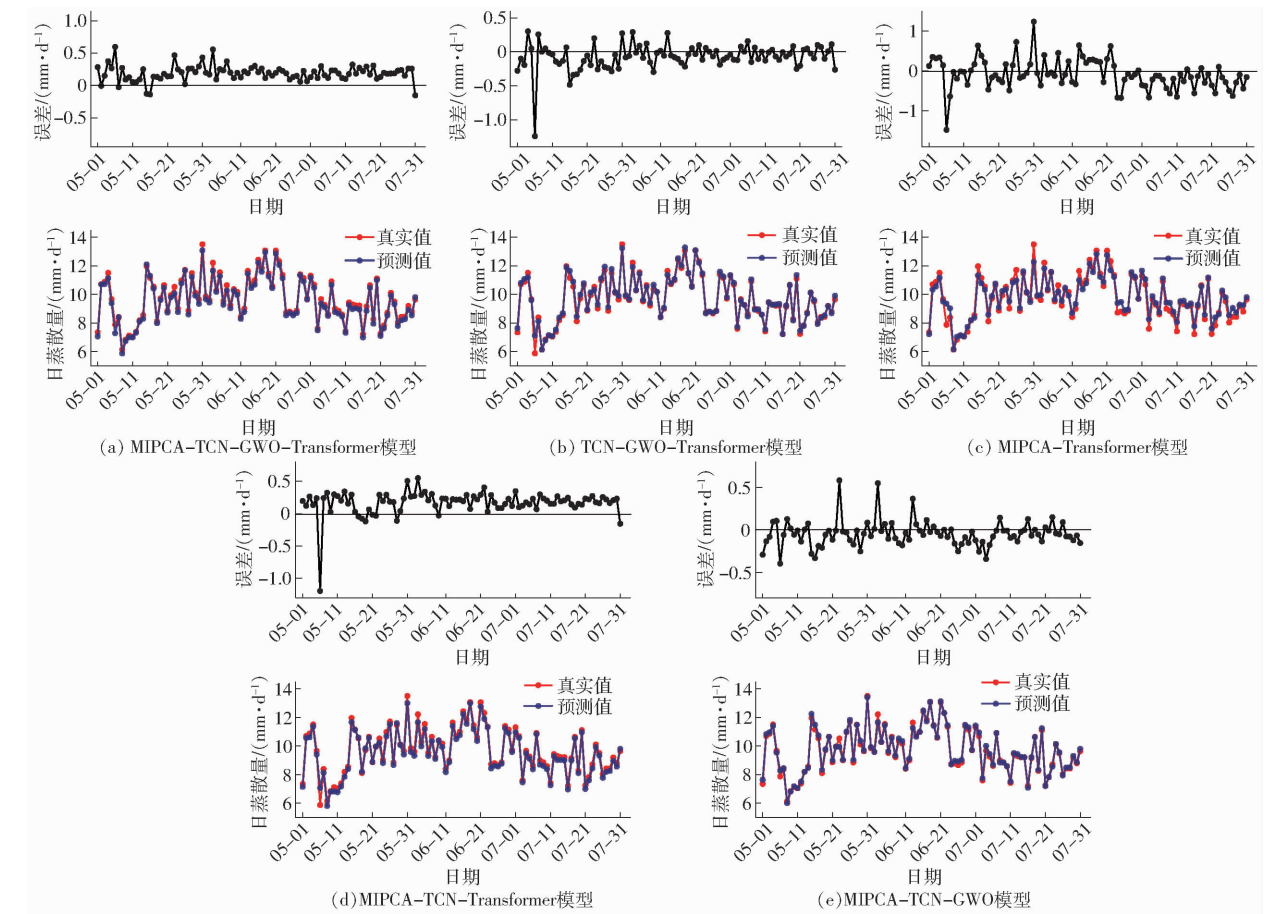


图 6 消融试验中预测值与真实值间的拟合曲线与误差曲线

Fig. 6 Fitting curves and error values between predicted and actual values in ablation experiment

表 3 消融试验结果			
Tab.3 Ablation experiment results			
模型	MAPE/ (mm·d ⁻¹)	RMSE/ (mm·d ⁻¹)	R ²
MIPCA-TCN-GWO-Transformer	0. 015	0. 312	0. 962
TCN-GWO-Transformer	0. 021	0. 379	0. 944
MIPCA-Transformer	0. 037	0. 585	0. 865
MIPCA-TCN-Transformer	0. 022	0. 406	0. 935
MIPCA-TCN-GWO	0. 029	0. 473	0. 913

综上所述,本文提出的茶园蒸散量预测模型与

表 3 所示。与 TCN-GWO-Transformer 预测模型对比,MAPE、RMSE 和 R^2 分别提升 0.6%、6.7% 和 1.8%;与 MIPCA-Transformer 预测模型对比,MAPE、RMSE 和 R^2 分别提升 2.2%、27.3% 和 9.7%;与 MIPCA-TCN-Transformer 预测模型对比,MAPE、RMSE 和 R^2 分别提升 0.7%、9.4% 和 2.7%;与 MIPCA-TCN-GWO 预测模型对比,MAPE、RMSE 和 R^2 分别提升 1.4%、16.1% 和 4.9%。其中,MIPCA-TCN-GWO-Transformer 模型相较于 MIPCA-Transformer 模型和 MIPCA-TCN-GWO 模型性能提升最显著,表示 TCN 与 Transformer 的融合策略能更有效地捕捉茶园蒸散量预测中的复杂时空关系。

以上 4 种模型对比,在各评价指标上均有不同程度的提高,模型在特征处理、参数优化以及特征提取等方面的性能明显提升。

3.3 对比试验

进一步在测试集上对本文模型和文献[9-14]所提出的模型以及利用 GA、PSO、WOA 等算法替换 GWO 算法的相关模型进行测试,验证优化算法对蒸散量预测模型的影响以及模型有效性,图 7 为各个模型在验证集上的拟合图和误差图,试验结果如表 4 所示。

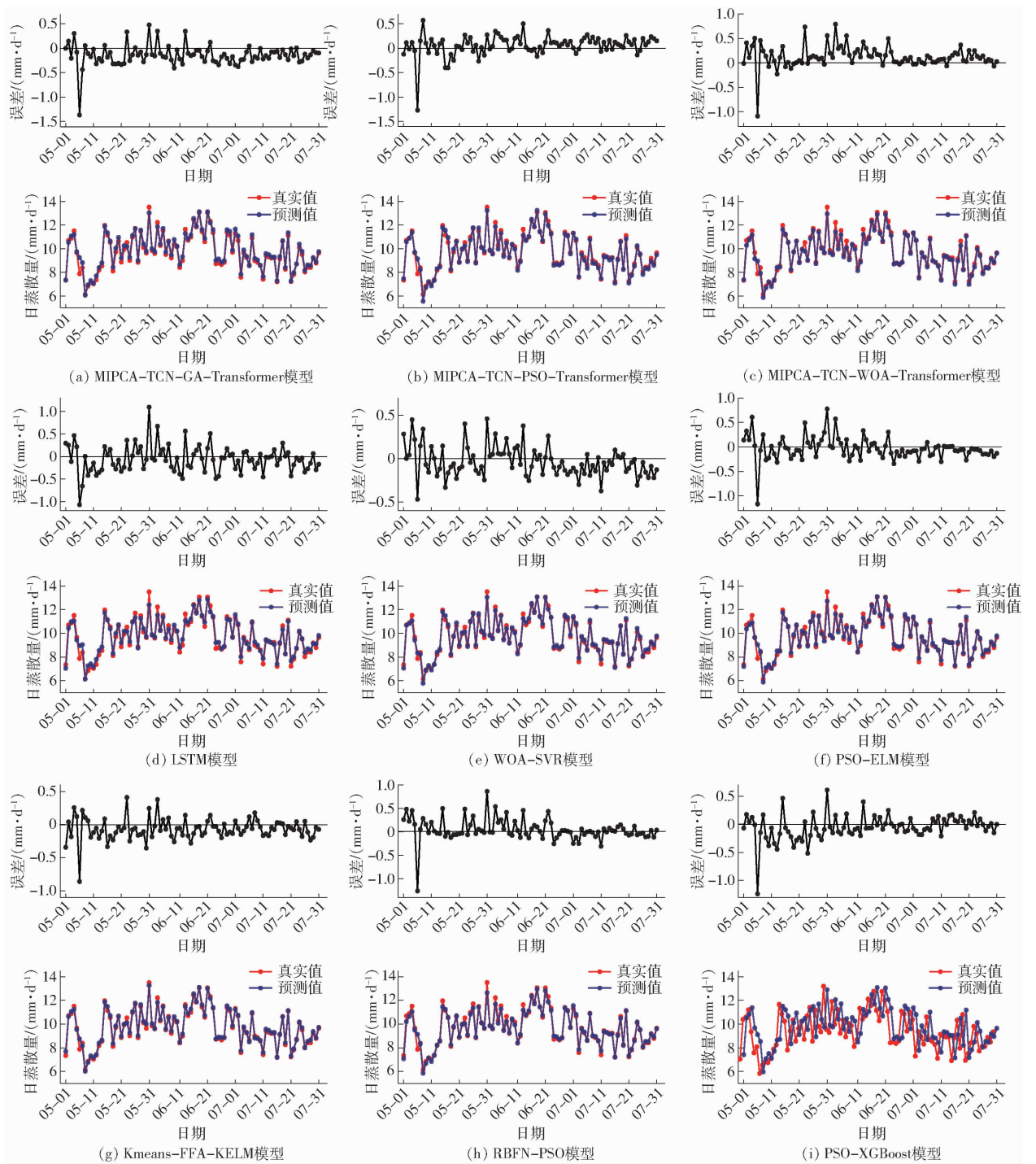


图 7 对比试验中预测值与真实值间的拟合曲线与误差曲线

Fig. 7 Fitting curves and error values between predicted and actual values in comparative experiment

在 3 种优化算法中,PSO 算法对本文模型表现提升最为显著,将本文使用的 GWO 算法与之对比,MAPE、RMSE 和 R^2 分别提升 0.6%、7.1% 和 2%,证明 GWO 算法在克服 TCN 模型易陷入局部最优解问题上更具优势。在预测性能方面,文献[11-13]提出的预测模型表现较好,将本文模型与 Kmeans-FFA-KELM 预测模型对比,MAPE、RMSE 和 R^2 分别提升 0.2%、2.2% 和 0.6%;与 PSO-ELM 预测模型对比,MAPE、RMSE 和 R^2 分别提升 0.3%、7.4% 和 2.1%;与 WOA-SVR 预测模型对比,MAPE、

RMSE 和 R^2 分别提升 0.6%、8.9% 和 2.6%。上述模型均考虑了特征提取和优化算法等方面。由于茶园蒸散量预测受茶园环境、土壤以及气象等多方面因素影响,本文额外引入了茶园土壤和环境特征并使用 MIPCA 算法进行特征处理,在提升性能的同时降低复杂度,使得本文模型在综合性能方面更具优势。

综上所述,本文模型相较于传统模型在评价指标上均有所提升,且在参数优化与特征处理方面表现卓越,证明了其在茶园蒸散量预测中的优越性。

表 4 对比试验结果

Tab.4 Comparative experiment results

模型	MAPE/	RMSE/	R^2
	($\text{mm}\cdot\text{d}^{-1}$)	($\text{mm}\cdot\text{d}^{-1}$)	
MIPCA - TCN - GA - Transformer	0.024	0.409	0.934
MIPCA - TCN - PSO - Transformer	0.021	0.383	0.942
MIPCA - TCN - WOA - Transformer	0.022	0.393	0.939
LSTM	0.030	0.511	0.897
WOA - SVR	0.021	0.401	0.936
PSO - ELM	0.018	0.386	0.941
Kmeans - FFA - KELM	0.017	0.334	0.956
RBFN - PSO	0.021	0.454	0.918
PSO - XGBoost	0.022	0.459	0.916

3.4 不同时间步长下的模型性能分析

为深入了解模型在不同时间步长下的适用性，利用训练好的 MIPCA - TCN - GWO - Transformer 模型对不同时间步长的数据集进行分析,不同预测步长下模型的拟合图及误差分布图如图 8 所示。

首先,本文针对日尺度数据进行逐日分析,选用 2022 年 8 月 1 日—31 日期间的日蒸散量数据集。图 8a 中预测曲线与真实值紧密拟合,其 R^2 达到 0.978,预测误差整体维持在较小范围内,表明模型在日时间尺度上有较高的准确性。其次,使用 2022 年 5 月至 2023 年 4 月的月累计蒸散量数据进行逐月分析。模型成功拟合了茶园蒸散量的季节性和年

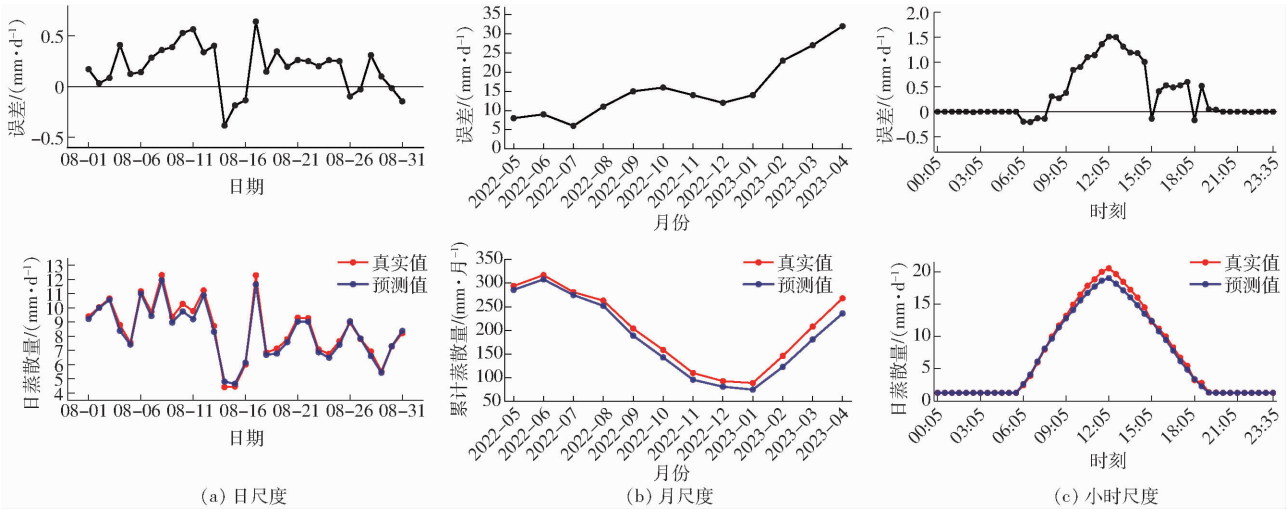


图 8 3 种时间步长下的模型性能分析

Fig.8 Model performance analysis under three time steps

度整体变化, R^2 达到 0.946。图 8b 中曲线后半段出现拟合程度下降和误差上升趋势,表明随着预测步长增加,模型预测误差逐渐增大。主要是由于时间跨度增加导致输入数据与输出数据联系减弱,进而影响了预测的准确性。随着数据量增加,模型对更长预测步长的预测性能也将逐渐提升。最后,对小时尺度数据进行蒸散量日变化分析,选取 2022 年 8 月 18 日的 48 组小时尺度数据。3 种时间尺度下的模型性能如表 5 所示,结果表明模型在高频率的时间尺度上表现出色, R^2 为 0.986。图 8c 中大多数时间点的误差较小,误差主要集中在峰值附近,整体平均绝对百分比误差为 0.608 mm/d。

表 5 3 种时间步长下模型性能比较

Tab.5 Comparison of model performance on three time steps

步长	MAPE/($\text{mm}\cdot\text{d}^{-1}$)	RMSE/($\text{mm}\cdot\text{d}^{-1}$)	R^2
0.5 h	3.387	0.608	0.986
1 d	0.031	0.295	0.978
30 d	9.576	17.410	0.946

综上所述,MIPCA - TCN - GWO - Transformer

模型在小时和日尺度上具有高度精确的预测能力。虽然模型在月尺度上误差相对较大,但仍能捕捉到蒸散量的整体趋势,可以满足不同时间尺度下的需求。

4 结论

(1)采用互信息法筛选重要特征,结合 PCA 降维来减少冗余数据。使用 TCN 对处理后的数据进行特征提取,通过 GWO 优化算法对网络中卷积核数量、学习率、训练轮数和批量大小等超参数进行优化,最后通过 Transformer 的全局注意力机制,考虑多参数间影响,来构建茶园蒸散量预测模型。

(2)消融试验和对比试验结果表明,本文模型决定系数达到 0.962,均方根误差为 0.312 mm/d,平均绝对百分比误差为 0.015 mm/d,预测性能优于 LSTM 等传统预测模型。

(3)针对模型在不同时间步长上的预测性能进行了分析,结果显示,模型在各个步长上的决定系数均高于 0.94,验证了模型具有较高的适用性和准确性。

参 考 文 献

[1] LI Xiaolin,ZHANG Xiaotao,NIU Jun, et al. Irrigation water productivity is more influenced by agronomic practice factors than by climatic factors in Hexi Corridor, Northwest China[J]. Scientific Reports,2016,6(1):37971.

[2] OTLES Z, GUTOWSKI W J. Atmospheric stability effects on Penman – Monteith evapotranspiration estimates[J]. Pure and Applied Geophysics, 2005, 162: 2239 – 2254.

[3] LI Jiusheng, ZHANG Jianjun, RAO Minjie. Wetting patterns and nitrogen distributions as affected by fertigation strategies from a surface point source[J]. Agricultural Water Management, 2004, 67(2): 89 – 104.

[4] GERARDO D,ALEJANDRO M,ABEL Q, et al. Estimation of reference evapotranspiration in a semi-arid region of Mexico[J]. Sensors,2023,23(15): 282 – 300.

[5] FANG Ruiming, SONG Shijie. Daily reference evapotranspiration prediction of Tieguanyin tea plants based on mathematical morphology clustering and improved generalized regression neural network[J]. Agricultural Water Management,2020,236: 116 – 135.

[6] FANG Wei,HUANG Shengzhi,HUANG Qiang, et al. Reference evapotranspiration forecasting based on local meteorological and global climate information screened by partial mutual information[J]. Journal of Hydrology,2018,561: 764 – 779.

[7] ARMAN A, ANDRE D, SNYDER R, et al. Meteorological driving forces of reference evapotranspiration and their trends in California[J]. The Science of the Total Environment,2022,25: 823 – 849.

[8] ZHAO Long,ZHAO Xinbo,ZHOU Hanmi, et al. Prediction model for daily reference crop evapotranspiration based on hybrid algorithm and principal components analysis in Southwest China[J]. Computers and Electronics in Agriculture,2021,190(1/2): 424 – 433.

[9] MANIKUMARI N,VINODHINI G,MURUGAPPAN A. Prediction of reference evapotranspiration for irrigation scheduling using machine learning[J]. Hydrological Sciences Journal,2020,65(16):2669 – 2677.

[10] 谢家兴,高鹏,孙道宗,等. 基于长短期记忆的柑橘园蒸散量预测模型[J]. 农业机械学报,2020,51(增刊2):351 – 356.

XIE Jiaying, GAO Peng, SUN Daozong, et al. Prediction model of evapotranspiration in citrus orchards based on long short-term memory[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51 (Supp. 2): 351 – 356. (in Chinese)

[11] BABAK M,SAEID M. Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm[J]. Agricultural Water Management, 2020, 237: 106 – 125.

[12] WU Zongjun, CUI Ningbo, HU Xiaotao, et al. Optimization of extreme learning machine model with biological heuristic algorithms to estimate daily reference crop evapotranspiration in different climatic regions of China[J]. Journal of Hydrology, 2021,603: 223 – 245.

[13] WU lifeng, PENG youwen, FAN junliang, et al. A novel kernel extreme learning machine model coupled with K-means clustering and firefly algorithm for estimating monthly reference evapotranspiration in parallel computation[J]. Agricultural Water Management, 2021, 245: 624 – 641.

[14] PETKOVIC D,GOCIC M,SHAMSHIRBAND S, et al. Particle swarm optimization-based radial basis function network for estimation of reference evapotranspiration[J]. Theoretical and Applied Climatology,2016,125(3 – 4):555 – 563.

[15] YU Jingxin,ZHENG Wengang, XU Linlin, et al. A PSO-XGBoost model for estimating daily reference evapotranspiration in the solar greenhouse[J]. Intelligent Automation & Soft Computing,2020,26(5):989 – 1003.

[16] CHEN Zhijun, SUN Shijun, WANG Yixin, et al. Temporal convolution-network-based models for modeling maize evapotranspiration under mulched drip irrigation[J]. Computers and Electronics in Agriculture,2020,169:206 – 217.

[17] 田原,庞骁,赵文祎,等. 基于 Transformer 的滑坡短期位移预测模型[J]. 北京大学学报(自然科学版),2023,59(2): 197 – 210.

TIAN Yuan, PANG Xiao, ZHAO Wenyi, et al. Short-term landslide displacement prediction model based on Transformer[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2023, 59(2): 197 – 210. (in Chinese)

[18] PANGAM H, RAMANA R K, SUBEESH A, et al. Predictive modelling of reference evapotranspiration using machine learning models coupled with grey wolf optimizer[J]. Water,2023,15(5):856.

[19] 王健,蔡焕杰,李红星,等. 日光温室作物蒸发蒸腾量的计算方法研究及其评价[J]. 灌溉排水学报,2006,25(6):11 – 14.

WANG Jian, CAI Huanjie, LI Hongxing, et al. Study on calculation method and evaluation of evapotranspiration of crops in solar greenhouse[J]. Journal of Irrigation and Drainage,2006,25(6):11 – 14. (in Chinese)

[20] 邱让建,杜太生,陈任强. 应用双作物系数模型估算温室番茄耗水量[J]. 水利学报,2015,46(6):678 – 686.

QIU Rangjian, DU Taisheng, CHEN Renqiang. Estimation of water consumption of greenhouse tomato by dual-crop coefficient model[J]. Journal of Hydraulic Engineering,2015,46(6):678 – 686. (in Chinese)

[21] GOYAL P, SHARDA R, SIAG M, et al. Development of the crop coefficient for vertically trained cucumber vines grown in soilless media under naturally ventilated greenhouse conditions[J]. Irrigation and Drainage, 2023, 72(2): 377 – 389.

[22] ZHANG Di,WEI Kai,YAO Yi, et al. Capture and prediction of rainfall-induced landslide warning signals using an attention-based temporal convolutional neural network and entropy weight methods[J]. Sensors, 2022,22(16):6240.

[23] LIMOUNI T, YAAGUBI R, BOUZIANE K, et al. Accurate one step and multistep forecasting of very short-term PV power using LSTM – TCN model[J]. Renewable Energy, 2023, 205: 1010 – 1024.

[24] KHAIRAN H E, ZUBAIDI S L, MUHSEN Y R, et al. Parameter optimisation-based hybrid reference evapotranspiration prediction models: a systematic review of current implementations and future research directions [J]. Atmosphere, 2022, 14(1): 77.

[25] ELKENAWY E M, BILEL Z, NADJEM B, et al. Improved weighted ensemble learning for predicting the daily reference evapotranspiration under the semi-arid climate conditions [J]. Environmental Science and Pollution Research International, 2022,29(54):81279 – 81299.