

付桐林,金晶.以霜冰优化算法优化CNN-BiLSTM-Attention的参考蒸散量估算[J].中国沙漠,2025,45(3):302-312.

以霜冰优化算法优化CNN-BiLSTM-Attention 的参考蒸散量估算

付桐林,金晶

(陇东学院 数学与信息工程学院,甘肃 庆阳 745000)

摘要:有限气象参数条件下借助于深度学习实现蒸散量的准确估算对于旱区有限水资源的高效利用和管理具有重要意义。当前基于混合深度学习模型CNN-BiLSTM-Attention的蒸散发估算忽视了参数优化,导致估算精度难以契合实际应用需求。本文提出了一种新的霜冰优化算法(RIME)优化CNN-BiLSTM-Attention的超参数的混合模型RIME-CNN-BiLSTM-Attention,实现了有限气象参数条件下临泽县参考蒸散量(ET_0)的准确预测。与CNN-BiLSTM-Attention相比,混合模型RIME-CNN-BiLSTM-Attention的平均绝对百分比误差(MAPE)从14.56%下降到14.09%,可决系数从0.8654上升到0.8930。此外,数值结果表明混合模型RIME-CNN-BiLSTM-Attention的模型性能优于分别采用哈里斯鹰优化算法(HHO)、鱼鹰优化算法(OOA)、北方苍鹰算法(NGO)对CNN-BiLSTM-Attention进行优化的混合模型HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention,意味着所构建混合模型RIME-CNN-BiLSTM-Attention具有更加稳健的模型性能和更高的计算精度,能够实现研究区域 ET_0 的准确估算。

关键词:参考蒸散量;霜冰优化算法;卷积神经网络;双向长短期记忆网络;注意力机制

文章编号: 1000-694X(2025)03-302-11

DOI: 10.7522/j.issn.1000-694X.2025.00167

中图分类号: P246.2

文献标志码: A

0 引言

蒸散量是水资源循环的关键组成部分,准确预测蒸散量对于水资源的合理分配^[1]、节水灌溉的有效实施^[2]、农作物产量的提升^[3]及水资源利用效率的提高^[4]等具有重要意义。在完全数据下,采用传统的经验模型如Penman-Monteith公式能够快速且准确地计算出蒸散量^[5-6],但在有限气象参数条件下,经验模型的应用受到很大限制;借助于蒸渗仪长期观测蒸散量又存在代价大且耗时费力的缺陷^[6-8]。这突出了开发创新方法的必要性,以实现有限数据条件下蒸散量的准确估算。

随着人工智能技术的显著进步,一系列典型的浅层机器学习模型如极致梯度提升树(XGBoost)^[7]、支持向量机(SVM)^[9-10]、M5模型树(M5T)^[10]及自适应模糊神经推理系统(ANFIS)^[10]等和深度学习模型

如长短期记忆网络(LSTM)^[11]、卷积神经网络(CNN)^[12-15]、时间卷积网络(TCN)^[14]、双向长短期记忆网络(BiLSTM)^[10,15]等在水文建模中得到了广泛的应用。这些先进的模型为解决复杂的水文问题开辟了新的可能,特别是在有限气象参数条件下通过利用人工智能强大的数据处理能力实现蒸散发的准确估算^[12-13]。将粒子群算法(PSO)和XGBoost相结合,可在有限气象参数下估算摩洛哥西北部Bouregreg盆地参考蒸散量(ET_0),且PSO-XGBoost的模型性能优于经验模型和单个XGBoost模型^[7]。Farshad等^[9]采用智能水滴算法(IWD)优化SVM的建模方法,实现了不完全数据下伊朗Arak、Mashhad及Shiraz的月 ET_0 准确估算。Roy等^[10]分别采用BiLSTM、SVM、M5T及ANFIS等预测了孟加拉国Gazipur和Pabnad地区的 ET_0 。结果表明深度学习

收稿日期:2025-03-27; 改回日期:2025-06-06

资助项目:甘肃省自然科学基金项目(23JRRM734);陇东学院博士基金项目(XYBYZK2305);庆阳市联合科研基金专项一般项目(QY-STK-2024A-068)

作者简介:付桐林(1977—),男,甘肃民乐人,博士,教授,主要研究方向为干旱区生态水文学、机器学习及其应用。

E-mail: futonglin2008@163.com

模型LSTM的预测精度最高。Fu等^[11]采用GWO分别优化SVM和LSTM的超参数建模策略,构建了混合模型对中卫市沙坡头地区的蒸散量预测,结果表明,采用启发式算法优化机器学习模型的参数,能够显著提高预测精度,且深度学习模型LSTM与启发式算法相结合的混合模型计算精度最高。上述研究表明,将单个深度学习模型如LSTM、BiLSTM与数据预处理技术或启发式算法相集成,构建混合模型能够实现在不完全数据下不同研究区域蒸散发的准确估算,且具有比浅层机器学习模型和经验模型更高的估算精度。

随着深度学习技术在水文领域的深入应用,有学者着眼于模型集成策略,通过精心设计的集成框架,充分挖掘各单个深度学习模型在架构、算法原理上的独特优势,将其进行有机融合,以期实现在不完全数据条件下对蒸散发的高精度估算^[14-15]。Ferreira等^[14]评估了深度学习模型LSTM和TCN相集成的混合模型TNN-LSTM、浅层机器学习模型ANN和RF提前7天预测 ET_0 方面的性能。结果表明,深度学习模型LSTM、TCN、TNN-LSTM的预测性能优于浅层机器学习模型ANN和RF,且TNN-LSTM比LSTM和TCN的预测精度更高。Dong等^[15]将Attention机制、CNN及BiLSTM相融合构建了混合模型CNN-BiLSTM-Attention,对不完全数据下中国不同区域 ET_0 进行估算,并与CNN-BiLSTM、BiLSTM、LSTM相比较,结果表明MA-CNN-BiLSTM的估算精度最高。上述研究充分证实,通过集成多个深度学习模型,能够有效整合各模型的独特优势,在不完全数据条件下成功实现蒸散发的精确估算。这一集成策略打破了单一模型的局限性,利用不同模型在特征提取、时间序列分析等方面的优势,显著提升了估算的准确性与可靠性。

系统分析现有研究发现,将多个深度学习模型集成的混合模型CNN-BiLSTM-Attention在参数设置上仍存在较强的主观性,模型参数的选择往往依赖于研究者的经验和先验知识,导致模型无法达到最优性能,这在很大程度上限制了估算精度的进一步提升^[14-15]。启发式算法源于对自然界生物觅食行为的模拟,这些算法能够准确地搜索到最优解,为解决机器学习模型的参数优化问题提供了新的思路^[16-19]。通过启发式算法,可以自动寻找模型的最优参数组合,避免人为因素的干扰,提高模型的性能和稳定性。

然而,目前尚未有研究采用启发式算法优化集成模型CNN-BiLSTM-Attention的超参数并将其应用于有限气象参数条件下蒸散发的估算研究中。

针对以上问题,本文首先将深度学习模型CNN和BiLSTM相集成构建CNN-BiLSTM模型,进而在混合模型CNN-BiLSTM中引入Attention机制,建立了混合模型CNN-BiLSTM-Attention。为获得混合模型CNN-BiLSTM-Attention最优参数,分别采用霜冰优化算法(RIME)^[16]、哈里斯鹰优算法(HHO)^[17]、鱼鹰优化算法(OOA)^[18]及北方苍鹰优化算法(NGO)^[19]优化混合模型CNN-BiLSTM-Attention的关键参数,构建了混合模型RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention及NGO-CNN-BiLSTM-Attention,以期实现有限气象参数条件研究区 ET_0 的准确估算。为验证模型的有效性,选取位于黑河流域中游干旱区水资源利用和调控的核心区临泽县作为研究区域,基于部分气象观测数据对所构建的模型在不完全数据下进行 ET_0 估算的模型性能进行比较,旨在获得适应于该区域最优的估算模型,为区域农田灌溉的节水实施和水资源的合理分配提供依据。

1 相关理论及模型

1.1 霜冰优化算法(RIME)

RIME是新兴的元启发式优化算法,其灵感来源于自然界中霜冰的形成和发展过程。该算法将优化问题中的候选解类比为水汽粒子,解空间对应霜冰形成的环境,通过模拟水汽粒子在该环境中的运动、聚集和变化等行为,来搜索最优解^[16]。

记初始时刻的待选解为 $X_0^i = [x_0^1, x_0^2, \dots, x_0^N]$,其中 N 的取值与待优化参数个数一致,适应度函数为 $f(X_0^i)$,在第 t 次迭代中,第 i 个待选解位置更新公式为

$$X_t^{id} = X_{t-1}^{id} + r_t^{id} \Delta X_{t-1}^{id} \quad (1)$$

式中: X_t^{id} 表示第 i 个候选解在第 t 次迭代时第 d 维的位置; r_t^{id} 是 $[0, 1]$ 内的随机数; ΔX_{t-1}^{id} 表示相对上一次迭代时位置的变化量。当 X_t^{id} 趋向最优解区域时,引入全局最优解 X_{tgbest}^{id} 和 X_{tbest}^{id} 分别表示当前适应度函数的最优待选解和第 i 个候选解领域内的局部最

优解, X_t^{id} 的更新公式为

$$X_t^{id} = X_{t-1}^{id} + \alpha(X_{t_{gbest}}^{id} - X_{t-1}^{id}) + \beta(X_{t_{lbest}}^{id} - X_{t-1}^{id}) \quad (2)$$

式中: 参数 α 和 β 用来调节 $X_{t_{gbest}}^{id}$ 和 $X_{t_{lbest}}^{id}$ 聚集的程度。在 RIME 算法中, 适应度函数值越小意味着虚拟温度越低, 通常保留使得适应度函数最小的待选解进入下一次迭代。同时, 为了避免算法陷入局部最优, 以一定概率对待选解进行随机扰动。当适应度函数达到要求时迭代终止。

1.2 卷积神经网络(CNN)

CNN 是专门为处理具有网格结构数据而设计的深度学习模型, 通过卷积核在数据上滑动进行卷积操作, 自动提取数据中的局部特征, 进而池化层对数据进行下采样, 保留数据主要特征的同时降低维数, 从而达到降低模型复杂度, 提高模型鲁棒性的目的, 全连接层将池化层的特征向量进行整合, 输出层给出最终的预测结果^[12-15]。设时刻 t CNN 的输入矩阵用 $X' = [x'_1, x'_2, \dots, x'_n]$ 表示, 则第一个卷积层的输出为

$$F'_{k1} = I\left(b_{k1} + \sum_{i=1}^n x'_i \cdot W_{k1 \cdot i}\right) \quad (3)$$

式中: b_{k1} 表示第一卷积层卷积核的偏移项; $W_{k1 \cdot i}$ 表示权重项, 激活函数

$$I(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

第一个池化层选择最大池化, 可以有效降低网络复杂度, 保留卷积运算后的主要特征。第一次池化的输出结果为 $F'_{m1} = \max\{F'_{k1}(d_1)\}$, 其中, $F'_{k1}(d_1)$ 是第一个卷积层输出的第 d_1 个向量, $d_1 \in m_1$ 。经过第二次卷积后, 可以从第一池化层的输出结果得到特征向量

$$F'_{k2} = I\left(b_{k2} + \sum_{d_1=1}^{m_1} F'_{d_1} \cdot W_{k2 \cdot d_1}\right) \quad (5)$$

式中: b_{k2} 表示卷积核在第二卷积层的偏移项; $W_{k2 \cdot d_1}$ 是权重项。第二池化层的输出为 $F'_{m2} = \max\{F'_{k2}(d_2)\}$, $F'_{k1}(d_1)$ 是第二个卷积层输出的第 d_2 个向量, $d_2 \in m_2$ 。经过两次卷积和池化运算后输出的特征向量表示为 $F' = [F'_1, F'_2, \dots, F'_{m2}]$ 。

1.3 双向长短时记忆网络(BiLSTM)

LSTM 通过输入门 F_t 、遗忘门 I_t 和输出门 O_t 及

单元状态, 这些门机制决定着哪些信息被保留、哪些被遗忘以及如何更新信息, 使得 LSTM 能够有效捕捉并学习序列数据中的长期依赖关系。

$$F_t = \sigma(W_F x_t + U_F h_{t-1} + b_F) \quad (6)$$

$$I_t = \sigma(W_I x_t + U_I h_{t-1} + b_I) \quad (7)$$

$$C_t = F_t \otimes C_{t-1} + I_t \otimes \tilde{C}_t \quad (8)$$

$$\tilde{C}_t = \sigma(W_C x_t + U_C h_{t-1} + b_C) \quad (9)$$

$$O_t = \sigma(W_O x_t + U_O h_{t-1} + b_O) \quad (10)$$

$$h_t = O_t \otimes \tan h(C_t) \quad (11)$$

式中: $\sigma(x) = (1 + e^{-x})^{-1}$ 和 $\tan h(\cdot)$ 是激活函数; \otimes 是 Hadamard 乘积; W 是对应栅极的权重向量; b 表示偏置向量; h_{t-1} 是隐含层前一时刻的输出; h_t 是 LSTM 的新状态; C_{t-1} 和 C_t 是 LSTM 细胞前后时刻的状态向量; \tilde{C}_t 表示输入门 I_t 的备选信息状态。

BiLSTM 是特殊的 LSTM 结构, 通过结合正反向 LSTM 层来增强对时间序列信息的处理能力。正向 LSTM 层按照时间序列的自然顺序处理信息, 而反向 LSTM 层则逆向处理, 以捕获序列中早期信息对当前点的潜在影响。通过这种方式, BiLSTM 能够在每个时间点上综合正反向信息, 从而得到更加准确的预测结果^[10, 15]。图 1 显示了 BiLSTM 的结构和 LSTM 的细胞结构。

1.4 Attention 机制

Attention 机制最初被用于解决序列到序列(seq2seq)任务中的长距离依赖问题, 在计算能力有限的情况下, Attention 机制通过计算 Query(Q)、键 Key(K)和值 Value(V)之间的关系, 能够将输入变量的不同特征对预测结果影响的大小作为赋予神经网络隐藏单元权重大小的依据, 使对估算结果有突出影响的因素获得更高关注, 可以辅助提高模型的精度^[20]。

设 Q 为查询向量, 对于输入向量 $X = [x_1, x_2, \dots, x_n]$, 注意力变量 $z \in [1, n]$ 表示被选择的数据的下标, 则给定 Q 时输入向量 X 中第 i 个数据与查询 Q 的注意力得分为

$$\alpha_i = \frac{\exp(s(x_i, Q))}{\sum_{j=1}^n \exp(s(x_j, Q))} \quad (12)$$

式中: $s(x_i, Q) = x_i^T Q / \sqrt{d}$ 表示注意力得分函数; d 是维度。对 V 进行加权求和得 Attention 的输出

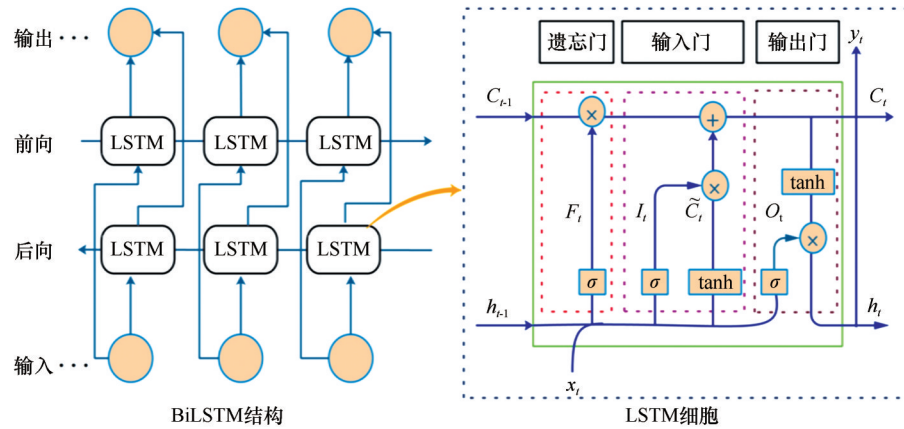


图1 BiLSTM的结构和LSTM的细胞

Fig.1 Structure of BiLSTM and cells of LSTM

$$a = \text{Attention}(K, V, Q) = \sum_{j=1}^n \alpha_j v_{j\odot} \quad (13)$$

1.5 RIME 优化 CNN-BiLSTM-Attention 的 ET_0 预测模型

本文构建的 RIME 优化 CNN-BiLSTM-Attention 的 ET_0 预测模型结构如图 2 所示,该模型主要有输入层、CNN 层(卷积层和池化层)、BiLSTM 层(前向 LSTM 层和后向 LSTM 层)、Attention 层及 RIME 算法构成。使用 CNN-BiLSTM-Attention 模型进行预测时,输入层将日时间尺度的气象要素如最高气温($^{\circ}\text{C}$)、最低气温($^{\circ}\text{C}$)、平均气温($^{\circ}\text{C}$)、平均风速($\text{m}\cdot\text{s}^{-1}$)及 ET_0 (mm)等划分为训练集和测试集并归一化后,进行数据平铺,实现多维数据的转换,并将相应的特征数据输入 CNN。CNN 中的卷积层对输入的数据进行特征提取,选取重要特征数据,池化层对特征数据进行降维;BiLSTM 层将数据分别通过前向和后向链式连接的多个 LSTM 单元进行计算;Attention 层计算不同时刻数据的特征状态对预测值的注意力值,得到时间序列各项数据与预测值的关联性,对 BiLSTM 的输出进行加权,进而得到估算值。为获得更加准确的估算结果,以 MAPE 为目标函数,采用 RIME 算法来优化混合模型 CNN-BiLSTM 中的学习率、BiLSTM 的神经元个数、正则化参数、注意力机制的键值等关键参数,从而构建了混合模型 RIME-CNN-BiLSTM-Attention 来实现不完全数据下研究区的 ET_0 估算。

为进一步探究不同优化算法对模型性能的影响,分别采用启发式算法 HHO、NGO 及 OOA 对 CNN-BiLSTM-Attention 模型的参数进行优化。在

此基础上,构建了混合模型 HHO-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention 及 OOA-CNN-BiLSTM-Attention,对研究区域的 ET_0 展开估算,以形成多维度的对比分析。上述混合模型的建模过程与混合模型 RIME-CNN-BiLSTM-Attention 的建模过程类似。HHO、NGO 及 OOA 的数学原理不再赘述,同时,单个的深度学习模型和浅层机器学习模型的估算性能也不再重复比较,因为单个深度学习模型具有比浅层机器学习模型更好的模型性能^[3-11]。

1.6 模型评价指标

采用纳什系数(NSCE)、均方误差(MSE)、均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE),可决系数 R^2 等模型精度评价指标对所提出模型的性能进行评价。

$$NSCE = 1 - \frac{\sum_{i=1}^n (o_i - \hat{e}_i)^2}{\sum_{i=1}^n (o_i - \bar{o})^2} \quad (14)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (o_i - \hat{e}_i)^2 \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (o_i - \hat{e}_i)^2} \quad (16)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |o_i - \hat{e}_i| \quad (17)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{o_i - \hat{e}_i}{o_i} \right| \times 100\% \quad (18)$$

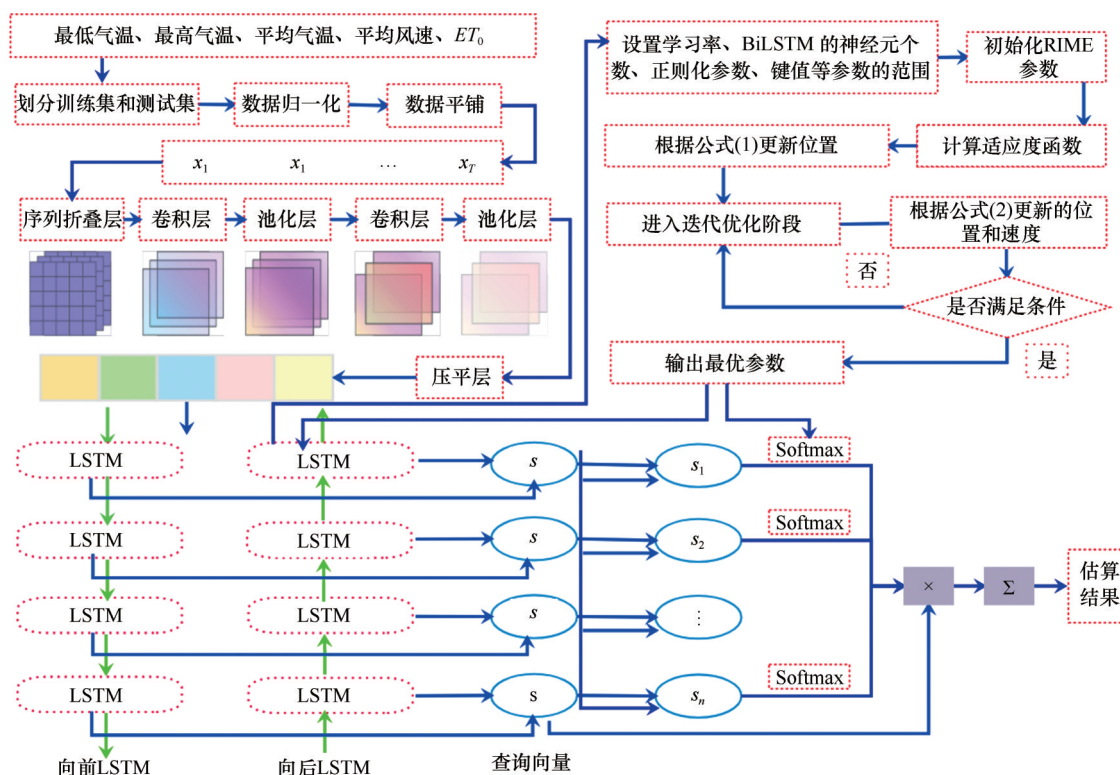


图2 RIME-CNN-BiLSTM-Attention的流程图

Fig.2 The flow chart of RIME-CNN-BiLSTM-Attention

$$R^2 = \frac{\sum_{i=1}^n (o_i - \bar{o})(\hat{e}_i - \bar{e})}{\sqrt{\sum_{i=1}^n (o_i - \bar{o})^2} \sqrt{\sum_{i=1}^n (\hat{e}_i - \bar{e})^2}} \quad (19)$$

式中: o_i 表示期望值; \hat{e}_i 表示模型的输出值。上述评价指标中, MSE 反映预测值与真实值偏差平方和的平均水平, $RMSE$ 是 MSE 的平方根, 用于直观展现预测误差的平均程度, MAE 体现预测值与真实值误差绝对值的平均值, $MAPE$ 则以百分比形式呈现预测值与真实值之间的相对误差。 $NSCE$ 和 R^2 的取值范围是 $[0, 1]$, 用于评估观测值和模拟值之间的拟合程度, 其值越大意味着拟合优度越高。

2 研究区域和数据

临泽县地处河西走廊中部 ($38^{\circ}57' - 39^{\circ}42'N$, $99^{\circ}51' - 100^{\circ}30'E$), 属温带大陆性干旱气候, 多年来平均年气温维持在 $7.6^{\circ}C$, 平均年降水量为 117 mm 。该区域主要景观类型涵盖绿洲、沙漠以及戈壁, 境内分布着 8 座中小型水库, 总库容达 $3\ 350\text{ 万 m}^3$ 。该地区农业灌溉高度依赖黑河中游的水资源, 是干旱地区水资源利用与调控的核心区域^[21-22]。绿洲的动态演变过程与下游生态进程紧密相连^[23]。因此, 精

确预测 ET_0 能够为水资源的高效利用与管理提供极具价值的参考依据。

本研究基于临泽 2011 年 1 月 1 日至 2020 年 12 月 31 日期间的部分逐日气象观测数据, 用于开展日尺度 ET_0 的估算研究。已获得的气象观测要素包含最高气温 ($^{\circ}C$)、最低气温 ($^{\circ}C$)、平均气温 ($^{\circ}C$)、平均风速 ($\text{m} \cdot \text{s}^{-1}$)。其中, 温度数据借助 HMP450 型温度计进行精确观测, 该温度计的测量范围为 $-40 \sim 60^{\circ}C$, 精度 $\pm 0.13^{\circ}C$ 。

3 结果

本研究将 2011 年 1 月 1 日至 2019 年 12 月 31 日的数据设定为训练集和验证数据集, 采用 5 折时间序列交叉验证 (TS-CV) 方法, 对每次划分得到的训练集和验证集, 使用构建的混合模型进行训练和验证。通过多轮数据划分和模型训练, 减少单次划分带来的随机性偏差, 同时在交叉验证过程中进行超参数优化, 找到更优的模型参数组合, 确保模型在不同数据分布下都能展现出稳定的性能。2020 年的数据作为独立测试集, 用于最终评估模型泛化能力, 训练集与独立测试集的比例为 9:1。

为实现数据的同维度化, 并有效提升模型的训

练效果,分别运用 Min-Max 归一化方法对训练集和测试集分别进行归一化处理。随后,将归一化后的数据平铺并依次输入至构建的混合模型,包括 RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention 及 CNN-BiLSTM-Attention,分别对研究区域的 ET_0 进行预测。

在本研究中,所有模型延时步长均设置为 10,这意味着模型将使用前 10 个历史数据作为自变量。在混合模型 CNN-BiLSTM-Attention 中梯度阈值为 1,初始学习速率为 0.01,L2 正则化参数取值为 0.001,交叉验证最大训练次数为 1 000。在混合模型 RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention 中,最大训练次数调

整为 30,自注意力层设置为单头,键数为 2,在每次训练迭代时,Dropout 层的舍弃概率固定为 0.1。学习率、BiLSTM 的神经元个数、正则化参数、注意力机制的键值等参数的取值区间分别为:学习率在 $[0.001, 0.1]$,BiLSTM 的神经元个数是 $[10, 50]$,注意力机制的键值为 $[2, 50]$,正则化参数则在 $[0.0001, 0.001]$ 。对于启发式算法 RIME、HHO、OOA 及 NGO,初始种群规模均设置为 5,最大进化代数数为 10。图 3 直观地呈现了混合模型 RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention 及 CNN-BiLSTM-Attention 的估算结果。表 1 显示了启发式算法 RIME、HHO、OOA 及 NGO 分别对混合模型 CNN-BiLSTM-Attention 进行优化后所获取的最优参数。

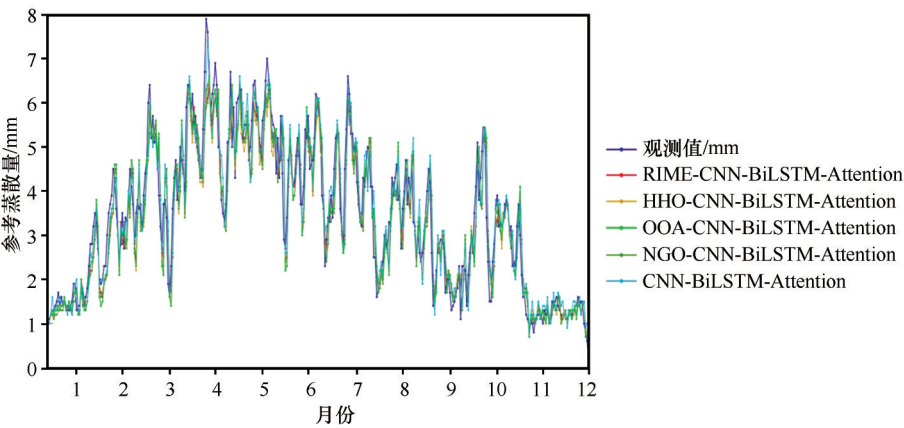


图3 RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention 及 CNN-BiLSTM-Attention 的估算结果

Fig.3 The forecasting results of RIME-CNN-BiLSTM-Attention, HHO-CNN-BiLSTM-Attention, OOA-CNN-BiLSTM-Attention, NGO-CNN-BiLSTM-Attention, and CNN-BiLSTM-Attention

表 1 启发式算法优化 CNN-BiLSTM-Attention 所得的最优参数

Table 1 The optimal parameters of CNN-BiLSTM-Attention

模型	学习率	神经元个数	键值	正则化参数
RIME-CNN-BiLSTM-Attention	0.0096	18	13	0.0005
HHO-CNN-BiLSTM-Attention	0.0063	16	8	0.0003
OOA-CNN-BiLSTM-Attention	0.0096	18	13	0.0005
NGO-CNN-BiLSTM-Attention	0.0078	20	23	0.0006

由图 3 可知,所有模型均能对研究区域的 ET_0 进行有效预测,在大部分数据点上,模型的预测值与实际观测值基本相符,且预测结果与观测值变化趋

势一致,仅在端点和极值点处存在一定偏差。值得注意的是,所有模型在极值点处均出现了系统性的高估或低估现象。这主要归因于气象观测要素本身存在极值,各要素与 ET_0 之间存在复杂的非线性关系,致使机器学习模型在局部范围内难以精准捕捉其变化规律,进而导致估算结果出现偏差。后续研究可考虑针对极值点的特殊情况,优化模型结构或调整参数,以进一步提升模型的预测精度。

为了全面准确地评估所构建模型的性能表现,分别计算了各模型的 $NSCE$ 、 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 及 R^2 等模型性能评价指标,具体计算结果如表 2 所列。做出各评价指标的柱形图如图 4 所示。将 CNN、BiLSTM 及 Attention 相集成的混合模型

表2 各模型在测试阶段的性能评价指标

Table 2 The performance metrics of each model in testing stage

模型	$NSCE$	MSE	$RMSE$	MAE	$MAPE/\%$	R^2
RIME- CNN-BiLSTM-Attention	0.8830	0.2879	0.5366	0.4013	14.09	0.8930
OOA- CNN-BiLSTM-Attention	0.8816	0.3249	0.5700	0.4349	14.34	0.8816
NGO- CNN-BiLSTM-Attention	0.8796	0.3303	0.5747	0.4352	14.28	0.8796
HHO- CNN-BiLSTM-Attention	0.8769	0.3378	0.5812	0.4482	14.38	0.8769
CNN-BiLSTM-Attention	0.8754	0.3548	0.5957	0.4761	14.56	0.8654

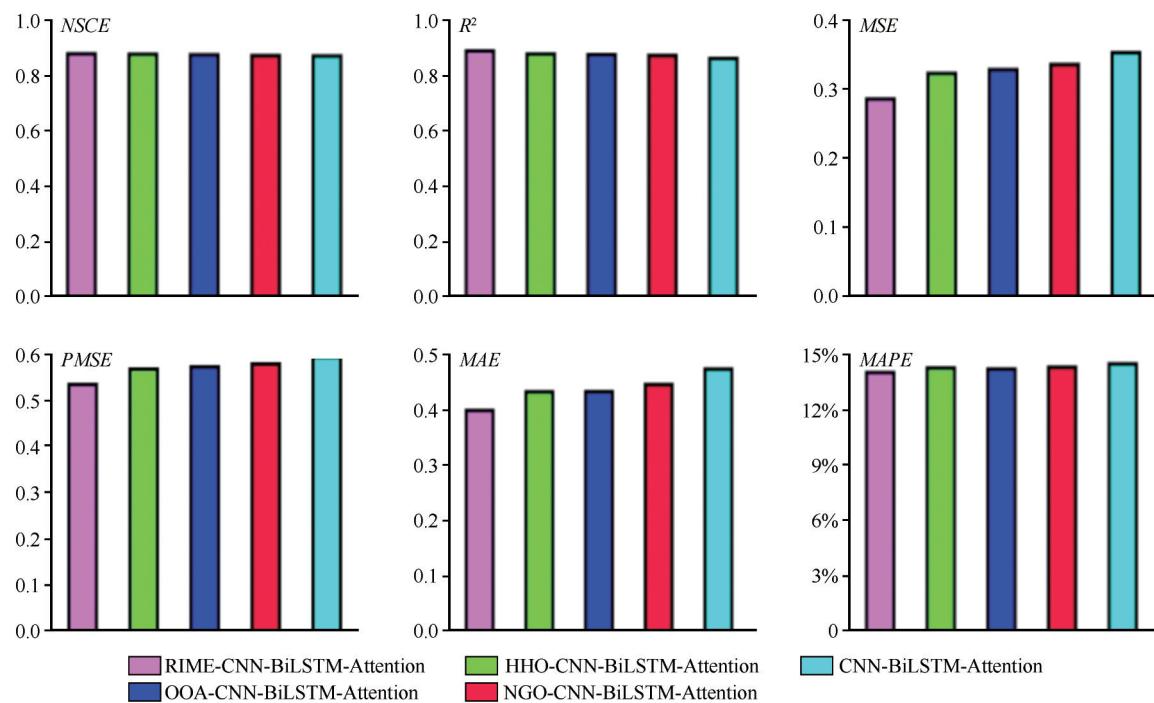


图4 误差评价指标的柱形图

Fig.4 Column chart of error evaluation indicators

CNN-BiLSTM-Attention 的误差评价指标 $NSCE$ 、 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 及 R^2 分别为 0.8754、0.3548、0.5957、0.4761、14.56% 及 0.8654, 基于启发式算法优化的混合模型 RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention 等的误差评价指标 $NSCE$ 、 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 及 R^2 的变化区间分别为 [0.8769, 0.8830]、[0.2879, 0.3378]、[0.5366, 0.5812]、[0.4013, 0.4482]、[14.09%, 14.38%] 及 [0.8769, 0.8930]; 对比发现, 采用启发式算法 RIME、HHO、OOA 及 NGO 优化混合模型 CNN-BiLSTM-Attention 的误差评价指标 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 均小于 CNN-BiLSTM-Attention 的, $NSCE$ 和 R^2 呈相反的变化趋势, 意味着优化后混合模型的预测效果均显著优于未优化的混

合模型 CNN-BiLSTM-Attention。这表明引入启发式算法的集成深度学习模型, 能够通过将不同算法的优势相整合, 构建出更为复杂且有效的模型结构, 实现研究区域 ET_0 的准确估算, 且表现出更高的预测精度和更强的稳定性。

将多个深度学习模型相集成, 能够充分利用各自的优点, 实现蒸散发的准确估算, 提高模型的预测精度。然而, 集成深度学习模型的参数决定了最终的预测结果和精度。伴随着各种新的启发式算法的提出, 将其应用于集成深度学习模型参数优化并实现蒸散发预测的研究有待开展。然而, 单独选择某种算法来优化继承深度学习模型的参数, 必然面临模型选择的风险。因此, 本文选择了近年来提出的新的启发式算法 RIME、HHO、OOA 及 NGO 分别优化集成深度学习模型 CNN-BiLSTM-Attention。

图4直观地显示出在相同的条件下,采用启发式算法RIME优化CNN-BiLSTM-Attention的参数构建混合模型RIME-CNN-BiLSTM-Attention,能够获得比其他3种启发式算法优化集成深度学习模型CNN-BiLSTM-Attention更好的模型预测性能,各项误差评价指标均显示混合模型RIME-CNN-BiLSTM-Attention的预测精度要比混合模型HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention、NGO-CNN-BiLSTM-Attention的预测精度高。这表明,本文所提出的集成模型RIME-CNN-BiLSTM-Attention具有很好的模型适应性和最小的计算误差。

4 讨论

在水文学领域,如何在不完全数据条件下实现不同生物气候带蒸散发的精准估算,一直是备受关注的键问题。机器学习模型,特别是深度学习模型,凭借强大的数据特征挖掘能力,在水文建模中得到广泛应用^[1-2]。这类模型在训练与测试过程中,能够精准估算有限气象参数条件下的 ET_0 ,不仅估算精度高,而且成本低廉,优势显著^[3-11]。由于不同的深度学习模型具有各自的优缺点,例如,CNN凭借独特的卷积层和池化层结构,在时间序列数据的特征提取环节表现卓越,能够高效捕捉数据中的局部特征和模式^[12-15]。然而,由于其缺乏对时间序列前后关联的有效建模机制,难以发掘时间序列数据中的长依赖关系^[13,15]。LSTM则是针对标准循环神经网络(RNN)在处理长时间序列数据时极易出现的梯度消失或梯度爆炸问题而设计的。LSTM通过引入门控机制,包括输入门、遗忘门和输出门,能够有效保存和传递长期信息,在处理长期依赖任务时性能优异。但LSTM在处理数据时,严格遵循输入时间序列的先后顺序,无法并行处理数据,这在一定程度上限制了其处理效率和对复杂数据模式的挖掘能力^[11-12,15]。双向长短期记忆网络(BiLSTM)作为LSTM的变体,创新性的同时从正向和反向两个方向对时间序列进行处理,这种结构使其能够深度提取时间序列的深层时间特征,充分利用过去和未来的信息^[10,15]。在不完全数据条件下估算蒸散发时,BiLSTM相较于LSTM展现出了更优的模型性能^[10,15]。然而,BiLSTM本质上仍属于递归神经网络结构,不可避免地存在梯度问题,在训练过程中

可能导致梯度消失或梯度爆炸,影响模型的收敛速度和稳定性。因此,将CNN和BiLSTM相耦合的混合模型CNN-BiLSTM相耦合能够充分利用各自的优点实现不完全数据下蒸散发的准确估算^[15]。但当利用输入项繁多、数据量庞大的时间序列数据进行模型训练时,该模型存在一定局限性。研究表明CNN-BiLSTM可能会忽视部分时刻数据所蕴含的重要特征信息使得模型学习能力降低,进而对模型预测精度产生负面影响。这是因为在处理复杂时间序列数据时,模型难以全面捕捉和处理所有信息,导致关键特征被遗漏。而注意力机制(Attention)则提供了有效的解决方案^[20]。注意力机制能够依据输入变量不同特征对预测结果影响程度的大小,为神经网络隐藏单元分配权重^[15,20,24-25]。这意味着,对 ET_0 估算结果影响显著的因素将得到更关注,有助于模型更精准地捕捉关键信息,从而辅助提升模型的估算精度。基于此,在CNN-BiLSTM模型中引入Attention机制,能够很好地解决原有模型容易忽略部分重要特征信息的问题,进而显著提高模型预测精度。因此,本研究将CNN-BiLSTM-Attention作为研究区域 ET_0 估算的主模块。

在集成深度学习模型CNN-BiLSTM-Attention中,学习率、BiLSTM的神经元个数、注意力机制的键值以及正则化参数等超参数,对模型最终的输出结果和预测精度起着决定性作用。手动调整这些超参数不仅耗时费力,而且在实际操作中往往难以实现,因为面对复杂的模型结构和大量的数据,人为判断最优超参数组合几乎是不可能完成的任务。因此,引入启发式算法来快速、准确地确定最优超参数^[16-19]。如表2所列,采用启发式算法优化集成深度学习模型CNN-BiLSTM-Attention,并用以计算研究区域 ET_0 ,结果显示,启发式算法能够自动寻找模型的最优参数组合,有效提高了模型的性能和稳定性。

为了更直观地展示本文所构建的模型在各项指标上的差异,以MSE、RMSE、MAE和MAPE作为4个维度的坐标轴,绘制出雷达图如图5所示,每个模型对应一个多边形,多边形的顶点位置代表该模型在对应指标上的数值。RIME-CNN-BiLSTM-Attention模型的多边形在雷达图中最靠近中心,因为其MSE、RMSE、MAE和MAPE的值都最小。这直观地

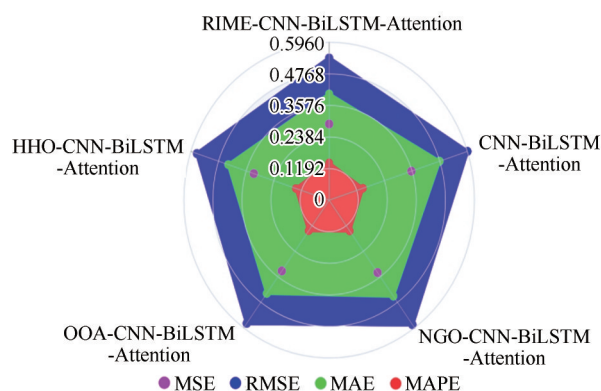


图5 各模型评价指标的雷达图

Fig.5 Radar diagrams of each model evaluation index

表明该模型在预测 ET_0 时,各方面误差指标都控制得最好,预测精度最高;OOA-CNN-BiLSTM-Attention 模型多边形相对离中心稍远,在 MSE 、 $RMSE$ 、 MAE 和 $MAPE$ 指标上,数值都比 RIME-CNN-BiLSTM-Attention 略大,说明它在预测误差控制方面稍逊一筹;HHO-CNN-BiLSTM-Attention 和 NGO-CNN-BiLSTM-Attention、CNN-BiLSTM-Attention 离多边形离中心要比 OOA-CNN-BiLSTM-Attention 稍远,其 MSE 、 $RMSE$ 、 MAE 和 $MAPE$ 均比混合模型 OOA-CNN-BiLSTM-Attention 的相应误差指标要大,但比混合模型 CNN-BiLSTM-Attention 的相应误差指标要小。这说明这两个混合模型其误差控制能力相对较弱。

综上所述,混合模型 RIME-CNN-BiLSTM-Attention 表现出比其他混合模型显著更低的误差水平,这一结果有力地证实了 RIME-CNN-BiLSTM-Attention 具备更高的预测精度,在研究区域 ET_0 预测任务中,具有更强的可靠性和精准性,能够为相关研究和实际应用提供更为准确的预测结果。

5 结论

有限气象参数条件下借助于深度学习实现 ET_0 的准确估算对于旱区有限水资源的高效利用和管理具有重要意义。本文将 CNN、BiLSTM 及 Attention 相集成,构建了混合模型 CNN-BiLSTM-Attention,进而采用 RIME、OOA、HHO 及 NGO 等启发式算法分别对混合模型 CNN-BiLSTM-Attention 中的关键参数进行优化,构建了多个混合模型等实现了不完全数据条件下黑河流域中游干旱区水资源利用和调控的核心区域临泽县 ET_0 的准确估算。与混合模型 CNN-BiLSTM-Attention ($NSCE=0.8754$, $R^2=$

0.8654 , $MSE=0.3548$, $RMSE=0.5957$, $MAE=0.4761$, $MAPE=14.56\%$) 相比,混合模型 RIME-CNN-BiLSTM-Attention、HHO-CNN-BiLSTM-Attention、OOA-CNN-BiLSTM-Attention 及 NGO-CNN-BiLSTM-Attention 的误差评价指标 $NSCE$ 和 R^2 的变化区间为 $[0.8769, 0.8830]$ 和 $[0.8769, 0.8930]$ 。 MSE 、 $RMSE$ 、 MAE 、 $MAPE$ 的变化区间分别为 $[0.2879, 0.3378]$ 、 $[0.5366, 0.5812]$ 、 $[0.4013, 0.4482]$ 、 $[14.09\%, 14.38\%]$,意味着采用启发式算法优化混合模型 CNN-BiLSTM-Attention 的关键参数能够有效提升模型的预测精度;整体上,混合模型 RIME-CNN-BiLSTM-Attention 的估算精度最高,其误差评价指标分别为 $NSCE=0.8830$ 、 $MSE=0.2879$ 、 $RMSE=0.5366$ 、 $MAE=0.4013$ 、 $MAPE=14.09\%$ 、 $R^2=0.8930$,与混合模型 CNN-BiLSTM-Attention 相比, $MAPE$ 从 14.56% 下降到 14.09% ,可决系数 R^2 从 0.8654 上升到 0.8930 ,意味着混合模型 RIME-CNN-BiLSTM-Attention 具有更好的模型性能,能够实现不完全气象数据下研究区 ET_0 的准确预测。本研究为不完全数据下实现干旱区 ET_0 的准确估算提供了新方法,也为干旱区水资源的调配提供了参考。

参考文献:

- [1] 任珩,赵文智,杨荣,等.河西走廊绿洲农业水生产力提升的途径与对策[J].中国沙漠,2024,44(5):217-224.
- [2] 邢铭强,马可,陈彩亮,等.河西地区农业绿色发展水平测度及耦合协调提升路径[J].中国沙漠,2024,44(6):207-219.
- [3] Fu T L, Li X R, Jia R L, et al. A novel integrated method based on a machine learning model for estimating evapotranspiration in dryland[J]. Journal of Hydrology, 2021, 603: 126881.
- [4] Fu T L, Li X R. Hybrid the long short-term memory with whale optimization algorithm and variational mode decomposition for monthly evapotranspiration estimation [J]. Scientific Reports, 2022, 12: 20717.
- [5] Chen S, He C, Huang Z, et al. Using support vector machine to deal with the missing of solar radiation data in daily reference evapotranspiration estimation in China[J]. Agricultural and Forest Meteorology, 2022, 316: 108864.
- [6] Bilali E A, Hadri A, Taleb A, et al. A novel hybrid modeling approach based on empirical methods, PSO, XGBoost, and multiple GCMs for forecasting long-term reference evapotranspiration in a data scarce-area[J]. Computers and Electronics in Agriculture, 2025, 232: 110106-110116.
- [7] Dong J, Xing L W, Cui N B, et al. Estimating reference crop evapotranspiration using optimized empirical methods with a novel improved Grey Wolf Algorithm in four climatic regions of China[J]. Agricultural Water Management, 2024, 108620.

- [8] Wu L F, Huang G M, Fan J L, et al. Hybrid extreme learning machine with meta-heuristic algorithms for monthly pan evaporation prediction[J]. *Computers and Electronics in Agriculture*, 2020, 168: 105–115.
- [9] Farshad A, Saeid M, Babak M, et al. Application of an artificial intelligence technique enhanced with intelligent water drops for monthly reference evapotranspiration estimation[J]. *Agricultural Water Management*, 2021, 244: 106622.
- [10] Roy D K, Sarkar T K, Kamar S S A, et al. Daily prediction and multi-step forward forecasting of reference evapotranspiration using LSTM and Bi-LSTM models [J]. *Agronomy*, 2022, 12: 594.
- [11] Fu T L, Li X R. Estimating the monthly pan evaporation with limited climatic data in dryland based on the extended long short-term memory model enhanced with meta-heuristic algorithms[J]. *Scientific Reports*, 2023, 13(1): 5960–5969.
- [12] Farzad R, Ahmadi F, Sharafati A, et al. Investigating hybrid deep learning models and meta-heuristic algorithms in predicting evaporation from a reservoir: a case study of Dez Dam[J]. *Earth Science Informatics*, 2023, 16: 3597–3618.
- [13] Elzain E H, Abdalla A O, Abdallah M, et al. Innovative approach for predicting daily reference evapotranspiration using improved shallow and deep learning models in a coastal region: a comparative study [J]. *Journal of Environmental Management*, 2024, 354: 120246.
- [14] Ferreira L B, Cunha F. Multi-step ahead forecasting of daily reference evapotranspiration using deep learning [J]. *Computers and Electronics in Agriculture*, 2020, 178: 105728.
- [15] Dong J, Xing L, Cui N, et al. Estimating reference crop evapotranspiration using improved convolutional bidirectional long short-term memory network by multi-head attention mechanism in the four climatic zones of China[J]. *Agricultural Water Management*, 2024, 292: 108665.
- [16] Su H, Zhao D, Heidari A A, et al. RIME: a physics-based optimization[J]. *Neurocomputing*, 2023: 183–214.
- [17] Heidari A A, Mirjalili S, Faris H, et al. Harris hawks optimization: algorithm and applications[J]. *Future Generation Computer Systems*, 2019, 97: 849–872.
- [18] Dehghani M, Trojovský P. Osprey optimization algorithm: a new bio-inspired metaheuristic algorithm for solving engineering optimization problems [J]. *Frontiers in Mechanical Engineering*, 2023.
- [19] Dehghani M, Hubálovský Š, Trojovský P. Northern goshawk optimization: a new swarm-based algorithm for solving optimization problems[J]. *Ieee Access*, 2021, 9: 162059–162080.
- [20] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [Z]. 10.48550/arXiv.1706.03762, 2017.
- [21] Ji X B, Zhao W Z, Jin B W, et al. Seasonal variations in energy exchange and evapotranspiration of an oasis-desert ecotone in an arid region[J]. *Hydrological Processes*, 2021, 35(9): 1–16.
- [22] 许澳康, 吉天琪, 石晶. 甘肃黑河流域生态网络构建与优化[J]. *中国沙漠*, 2025, 45(2): 61–70.
- [23] 雷军, 程新平, 薛春, 等. 黑河流域中游北部荒漠区植物群落特征与稳定性[J]. *中国沙漠*, 2024, 44(6): 187–194.
- [24] Wang S X, Shi J R, Yang W Y, et al. High and low frequency wind power prediction based on Transformer and BiGRU-Attention[J]. *Energy*, 2024, 288: 129753.
- [25] Wang W H, Xi C P. Short-term electricity load forecasting based on TPE-optimized CNN-BiGRU-attention [J]. *Journal of Physics: Conference Series*, 2023, 2496: 1.

Estimating reference evapotranspiration using CNN-BiLSTM-Attention enhanced by RIME optimization algorithm

Fu Tonglin, Jin Jing

(College of Mathematics and Information Engineering, Longdong University, Qingyang 745000, Gansu, China)

Abstract: Accurately estimating evapotranspiration using deep learning models under limited meteorological parameter conditions is of great significance for the efficient utilization and management of limited water resources in arid regions. However, current deep-learning-based hybrid model CNN-BiLSTM-Attention neglect the parameter optimization, making it difficult to meet the requirements of practical application. In this study, a novel hybrid model RIME-CNN-BiLSTM-Attention was proposed to achieve accurate prediction of daily reference evapotranspiration (ET_0) in Linze County under limited meteorological parameter conditions by using the RIME Optimization Algorithm to optimize the hyper-parameter of CNN-BiLSTM-Attention. Compared with CNN-BiLSTM-Attention, the mean absolute percentage error ($MAPE$) of RIME-CNN-BiLSTM-Attention decreases from 14.56% to 14.09%, and the coefficient of determination increases from 0.8654 to 0.8930; Furthermore, the numerical results show that the hybrid model RIME-CNN-BiLSTM-Attention outperformed than that of HHO-CNN-BiLSTM-Attention, OOA-CNN-BiLSTM-Attention, and NGO-CNN-BiLSTM-Attention, where CNN-BiLSTM-Attention was separately optimized by the Harris Hawks Optimization (HHO), Osprey Optimization Algorithm (OOA), and Northern Goshawk Optimization (NGO), suggesting that the proposed hybrid model RIME-CNN-BiLSTM-Attention has more robust model performance and higher calculation accuracy, and can achieve accurate estimation of ET_0 in the study area.

Key words: reference evapotranspiration; RIME optimization algorithm; convolutional neural network; bidirectional long short-term memory network; attention mechanism