

基于LSTM-Attention-GAN的潜在蒸散量预测研究

孙国荣, 杨时雨

北华大学计算机科学技术学院, 吉林 吉林

收稿日期: 2026年4月22日; 录用日期: 2026年5月19日; 发布日期: 2026年5月27日

摘要

全球月尺度潜在蒸散量(PET)预测面临驱动因素多源耦合、时序依赖复杂等挑战; 同时, 传统以MSE、MAE等点对点损失为主的回归方法容易出现均值收缩, 进而影响预测分布对尾部波动的表达。基于1981~2020年、 $1^\circ \times 1^\circ$ 统一网格数据, 文章构建了“过去24个月多源气象-水文变量预测下一月PET”的序列学习框架。模型以长短期记忆网络(LSTM)为主干, 引入时间注意力与特征注意力以强化关键月份与关键驱动因子的表征; 进一步结合生成对抗网络(GAN), 在分布层面对预测结果施加约束, 以促进预测分布与观测分布的一致性。实验结果表明, 引入注意力机制后RMSE显著降低且系统偏差减小; 在此基础上加入对抗学习后, 模型在分布形态相似性、综合一致性及高值区偏差控制方面进一步改善, 整体表现更为稳健。

关键词

潜在蒸散量(PET), 全球格点预测, 长短期记忆网络(LSTM), 注意力机制, 生成对抗网络(GAN), 对抗学习, 分布对齐

Prediction of Potential Evapotranspiration Based on LSTM-Attention-GAN

Guorong Sun, Shiyu Yang

College of Computer Science and Technology, Beihua University, Jilin Jilin

Received: April 22, 2026; accepted: May 19, 2026; published: May 27, 2026

Abstract

The prediction of global monthly potential evapotranspiration (PET) faces challenges such as multi-

source coupling of driving factors and complex time series dependence. At the same time, the traditional regression methods based on point-to-point losses such as MSE and MAE are prone to mean shrinkage, which in turn affects the expression of tail fluctuations in the prediction distribution. Based on the $1^\circ \times 1^\circ$ unified grid data from 1981 to 2020, this paper constructs a sequence learning framework of “multi-source meteorological-hydrological variables in the past 24 months predict PET in the next month”. The model takes the long-term and short-term memory network (LSTM) as the backbone, and introduces time attention and feature attention to strengthen the representation of key months and key driving factors. Furthermore, the generative adversarial network (GAN) is combined to impose constraints on the prediction results at the distribution level to promote the consistency between the prediction distribution and the observation distribution. The experimental results show that the RMSE is significantly reduced, and the system deviation is reduced after the attention mechanism is introduced. On this basis, after adding adversarial learning, the model is further improved in terms of distribution similarity, comprehensive consistency, and deviation control in high-value areas, and the overall performance is more robust.

Keywords

Potential Evapotranspiration (PET), Global Grid Prediction, Long Short-Term Memory Network (LSTM), Attention Mechanism, Generative Adversarial Network (GAN), Opposing Learning, Distribution Alignment

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来, 全球气温持续上升, 导致大气的潜在蒸散需求不断增加, 从而引发了陆 - 气系统水分平衡的显著变化, 也让干旱风险随之上升。在这个过程中, 潜在蒸散量(Potential Evapotranspiration, PET)逐渐成为气候因素、水文循环和生态响应之间的重要纽带。如今, PET 的变化趋势和空间分布, 越来越成为研究干旱发生和发展的关键因素[1]-[3]。

随着遥感技术的进步, 以及高时空分辨率的再分析数据集、多源观测资料和机器学习、深度学习等先进建模方法的发展, 潜在蒸散量(PET)的估算和预测已经取得了显著进展。遥感数据的应用, 可以获取大范围、连续的土地表面温度、植被指数和辐射收支等信息, 从而为构建 PET 的长时间序列提供了坚实的数据支持[4]。再分析数据集, 如 ERA5-Land 和 GLEAM-ET 等, 为全球和区域尺度上 PET 的研究提供了丰富的多变量驱动条件。这些数据集通过整合多种气候和环境变量, 帮助我们更精确地了解 and 预测潜在蒸散量的变化[5]。在模型方法方面, 深度学习模型以及物理引导型神经网络等技术, 正逐步应用于蒸散发估算和干旱监测的研究中。这些先进的模型不仅提升了估算的精度, 也使得干旱监测更加高效和精准[6]-[10]。这些方法能够有效捕捉传统物理模型难以量化的非线性机制, 从而在提高 PET 预测精度的同时, 也增强了时空一致性。

国内外在 PET 预测、降尺度以及农业水资源管理方面已经开展了大量的综合研究。在传统水文事件预测领域, LSTM 模型曾成功应用于洪水预测, 并为 ET-降水 - 径流耦合研究提供了重要的方法支持[11]。近年来, 机器学习和深度学习逐渐成为 PET 预测中的主流方法。例如, *Sustainability* 上的一项研究展示了多模型 PET 预测框架在流域可持续水资源管理中的应用潜力, 突显了这些技术在实际应用中的巨大价值[12]。此外, Roy 等对蒸散量预测及灌溉调度中的机器学习方法进行了综述, 指出植被指数、

辐射变量与气象因子是决定模型精度的核心特征[13]。研究者把 LSTM 等循环神经网络引入参考作物蒸散量(ET_0)预测, 成功提升了华北及嘉陵江流域等地的月、日尺度 ET_0 预报精度, 显示出在观测资料较为完备条件下深度序列模型的优势[14][15]。综合来看, 深度学习技术正在从单点预测模型逐步走向区域尺度、水文-气候耦合、多变量融合与可解释学习的发展方向, 尤其是注意力机制、GAN、ConvLSTM 等结构的加入, 使得 PET 预测在干旱监测、水资源管理和气候变化研究中具有更高的精度与应用潜力。

从建模角度来看, 全球尺度 PET 预测面临严峻挑战。首先, PET 受到温度、辐射、湿度等多重因素的共同影响, 具有显著的跨月累积效应和非线性响应, 不同气候带的主导因子也有很大差异。这要求模型不仅能够捕捉长期的依赖关系, 还要能够自适应地突出关键月份和关键变量; 其次, 传统的深度预测模型通常以 MSE、MAE 等点对点损失作为目标, 这可能导致预测结果趋向整体均值, 从而降低对偏态分布、尾部概率以及极端高 PET 事件的表达能力, 而这些极端过程恰恰是与高温、辐射异常等气候扰动密切相关的关键因素。

为应对上述问题, 本文依托全球多源再分析与气候格点数据, 构建了一个面向全球统一网格的深度时序预测框架。方法上, 本文以 LSTM 为序列依赖建模核心, 引入时间注意力(Temporal Attention)以识别对预测贡献最大的关键月份, 并引入特征注意力(Feature Attention)以动态调整不同输入变量在不同区域的贡献, 实现“序列依赖-关键时间识别-特征重要性强调”的多层级表征; 进一步地, 本文将生成对抗网络(GAN)引入 PET 预测, 通过生成器-判别器对抗训练实现分布层面对齐, 从而提升预测分布与观测分布在整体形态上的一致性, 并在一定程度上缓解传统点对点回归可能产生的均值收缩现象。

2. 数据与方法

2.1. 数据来源与时空尺度

本文构建全球统一网格化数据集用于月尺度 PET 预测。数据覆盖 1981~2020 年, 空间分辨率统一为 $1^\circ \times 1^\circ$, 所有变量均按月尺度组织并在时间维度逐月对齐, 以满足深度时序模型对长序列输入的需求。为避免时间信息泄露, 模型训练与测试采用按时间顺序划分, 前期用于训练、后期用于测试, 并仅在陆地区网格点上构建样本。

预测目标(PET)采用 CRU TS v4.08 提供的月度 PET 产品(原始 0.5°), 数据基于 FAO-56 Penman-Monteith 方法计算[16]。为与输入特征统一分辨率, 本文将其重采样至 $1^\circ \times 1^\circ$ 网格, 采用双线性插值法, 并保持 PET 量纲一致为 mm/月。

驱动因子以 LandBench 再分析资料为核心来源。LandBench 作为全球 LSV 预测基准数据集, 提供多源的大气、陆面与地理信息, 便于不同数据驱动模型在统一框架下进行公平比较[17]。

由于蒸散驱动因子具有明显月际变化与累积特征, 因此本文采用月尺度进行研究。月平均可以降低天气噪声并增强气候信号。同时, 月尺度更容易保证全球网格数据的时空完整性, 有利于端到端统一建模。

2.2. 输入变量体系与物理机制划分

本文共使用 12 个动态输入变量, 包括: 2 m 温度(T_{2m})、比湿(Q)、月最小、最大气温(T_{min} 、 T_{max})、降水(tp)、水量收支(pb)、净短波辐射(ssr)、下行长波辐射(strd)、潜热通量(slhf)、地表温度(skt)、PET (pet) 以及 SPEI1 (spei)。其中, 辐射与能量项(ssr、strd、skt、slhf)用于刻画“入射能量-地表吸收-蒸散能量消耗”的闭环过程; 降水与水量收支等用于体现水分可得性约束。上述特征均来自预测时点之前的历史

窗口, 从而避免信息泄露。

2.3. 时间序列样本构建与数据预处理

对于全球所有陆地网格点, 都构建了完整的 480 个月(1981~2020)时间序列。采用固定的历史窗口长度 $T = 24$ 个月作为输入, 并通过窗口预测未来 1 个月的 PET。最终, 所有网格点的输入张量采用统一的数据格式, 从而确保模型能够在相同结构下学习到共性规律与区域差异。

为保证训练的稳定性, 本文采用统一的数据处理流程, 包括:

- 1) 缺测处理: 对少量缺测值采用仅基于邻近时空信息的插值、填补。
- 2) 标准化: 对每个变量进行 Z-score 标准化。
- 3) 空间重采样: 将各变量统一至 1° 网格并完成逐月对齐。

2.4. 深度预测模型

1) 基线序列模型(LSTM)

LSTM 使用门控机制有效地缓解了传统 RNN 梯度消失的问题, 可以捕捉长序列中的长期依赖关系, 因此特别适用于 PET 这类具有季节积累性与滞后效应的时间序列预测任务。

2) 注意力增强 LSTM

为刻画多变量时序中“关键月份、关键变量”的贡献差异, 本文在 LSTM 框架中引入注意力机制, 构建三类结构:

- ① 特征注意力(FALSTM): 在变量维度对输入特征进行自适应加权, 以强调与 PET 更相关的驱动因子。
- ② 时间注意力(TALSTM): 在时间维度对历史窗口内不同月份的隐藏状态进行加权汇聚, 以突出关键月份信息。
- ③ 联合注意力(FTALSTM): 同时在特征维度与时间维度实施加权, 从而实现对变量异质性与时间异质性的协同建模。

3) 对抗分布学习

传统回归训练大多数以 MSE、MAE 等点估计损失为目标, 预测结果可能向条件均值收敛并造成波动幅度被低估。为缓解该问题, 本文引入 GAN 进行对抗式分布学习: 生成器基于 LSTM、注意力结构生成 PET 预测序列, 判别器用于区分真实序列与生成序列, 并将分布差异通过对抗损失反馈给生成器, 从而在保持数值精度的同时促进预测分布与观测分布的一致性。

2.5. 全球统一式建模策略

在全球尺度任务中, 传统的分区建模、多模型并行方法常常导致模型碎片化、训练成本高、区域间不可比等问题。本文采用全球统一式建模策略: 对全球全部格点使用单一模型训练与预测。

该策略的关键依据包括:

- 1) PET 形成机制具有全球一致的物理基础(辐射 - 能量平衡、热力状态、水汽压差与季节循环等), 统一模型可学习跨区域共享规律。
- 2) 通过参数共享实现跨区域迁移学习, 模型可隐式学习不同气候区差异并降低区域独立训练的过拟合风险。
- 3) 训练效率更高且对数据稀缺区域具有更好的泛化能力。
- 4) 对抗学习中判别器可在全局范围学习真实分布, 从而提升分布逼近的稳定性, 并增强预测分布与观测分布在整体形态上的一致性。

2.6. 评价指标

为了全面评估模型的预测性能, 本文使用了包括 RMSE、BIAS、皮尔逊相关系数 R、决定系数 R^2 、KGE 和 NSE 在内的六个统计指标。其中, RMSE 反映了模型的平均误差水平; BIAS 衡量了系统偏差的绝对大小; R 和 R^2 分别用来评估线性相关性和模型的解释度; NSE 和 KGE 则作为水文一致性指标, 其中 KGE 通过综合考虑相关性、偏差和变异性, 为模型提供了更为全面的评估。

2.7. 空间分布与分布特征分析方法

为进一步分析模型在全球不同区域上的适应性差异, 本文在测试集上逐网格计算 RMSE 和 KGE, 并绘制其空间分布图。RMSE 用于反映不同区域的误差幅值, KGE 用于综合评估相关性、偏差和变异性的一致性。通过比较基线模型与最优模型在全球陆地区域上的空间分布特征, 可以更直观地识别模型性能改善的区域分布及其空间异质性。

为直接验证 GAN 在分布层面的作用, 本文进一步对测试集全部有效样本进行汇总, 比较观测值、TALSTM 预测值与 TALSTM-GAN 预测值的概率密度函数(PDF)和累积分布函数(CDF), 以分析不同模型对总体分布形态的拟合能力。

此外, 为考察模型在分布两端的性能, 本文依据观测值的第 10 百分位(P10)和第 90 百分位(P90)分别定义低值区和高值区, 并绘制相应的预测值 - 观测值散点图, 通过与 1:1 参考线的偏离情况以及 R、RMSE 和 BIAS 等指标, 比较 TALSTM 与 TALSTM-GAN 在低值段和高值段的预测差异。

3. 结果与分析

3.1. 总体性能对比

为了评估所构建框架的整体性能, 本文在独立测试集上对八种模型进行了对比。考虑到不同区域可能存在离群误差, 本文采用各指标在全球网格点上的中位数来汇总总体表现(见表 1), 以减少少数异常网格对整体结论的影响。

Table 1. The median of each indicator on the global grid points

表 1. 各指标在全球网格点上的中位数

模型/评价指标	RMSE	BIAS	R	R^2	KGE	NSE
LSTM	8.7187	6.9987	0.9859	0.9534	0.8445	0.9504
LSTM-GAN	7.9486	6.5749	0.9845	0.9401	0.8612	0.9401
TALSTM	6.6280	4.5034	0.9867	0.9739	0.9575	0.9702
TALSTM-GAN	6.6079	4.5485	0.9870	0.9712	0.9576	0.9712
FALSTM	6.9433	4.5888	0.9867	0.9715	0.9442	0.9682
FALSTM-GAN	6.7583	4.5550	0.9867	0.9700	0.9543	0.9700
FTALSTM	7.4028	5.1853	0.9840	0.9661	0.9331	0.9624
FTALSTM-GAN	6.7230	4.6546	0.9860	0.9697	0.9570	0.9697

整体上, 八种模型在全球网格上均呈现出较强的预测能力: 相关性指标普遍较高(R 普遍大于 0.98), R^2 与 NSE 多数高于 0.94, 说明在所选多源气象 - 水文驱动因子支持下, 月尺度 PET 具有良好的可预测性。

在此基础上, 模型结构的改进带来了性能的稳定提升。与 LSTM 相比, 引入注意力机制后, 模型的

误差和偏差显著下降; 进一步引入对抗学习后, 大多数模型的系统偏差得到了有效减小, 分布一致性相关指标(如 KGE、NSE)也得到了提升, 从而使得整体性能更加均衡。

3.2. 注意力机制的性能提升

表 1 显示, 相比基准 LSTM, 注意力机制能显著提升 PET 预测精度与一致性。

其中 TALSTM 效果最为突出: TALSTM 相较于 LSTM 来看, RMSE 由 8.7187 降至 6.6280 (约下降 24%), BIAS 由 6.9987 降至 4.5034 (降幅超过 35%); 同时 R^2 提升至 0.97 以上, KGE 约提升至 0.96, NSE 约为 0.97, 表明时间注意力有效增强了模型对 PET 月际演变与关键月份贡献的表征能力。

FALSTM 同样带来显著提升: FALSTM 相较 LSTM 在 RMSE 与 BIAS 上分别降低约 20% 和 34%, 并显著提高 R^2 、KGE 与 NSE, 说明从特征维度进行动态加权有助于突出与 PET 更相关的输入变量、减弱冗余信息干扰。

FTALSTM 整体提升但并非单项最优: FTALSTM 融合两类注意力后, RMSE 与 BIAS 较 LSTM 分别降低约 15% 和 26%, 相关性与一致性指标亦整体提高, 体现出联合注意力在“稳健综合表现”上的优势, 但在部分指标上可能略低于单一注意力结构的最优值。

综合来看, 注意力机制的主要贡献在于: 在多变量、跨月窗口的输入条件下, 模型能够更有针对性地强调“对 PET 更关键的月份、变量”, 从而在全球尺度上提升误差控制与一致性指标。

3.3. GAN 的误差与一致性优化

除注意力机制外, 本文进一步引入对抗式分布学习, 以缓解传统点对点回归易产生的“均值收缩”与系统偏差问题, 并提升预测分布与观测分布在整体形态上的一致性。对比结果表明, GAN 在不同结构中带来的收益存在差异, 但总体趋势为“减弱系统偏差, 提升综合一致性”。

基线层面: LSTM-GAN 相较 LSTM 在 RMSE 与 BIAS 上有所降低(RMSE 从 8.7187 降至 7.9486, BIAS 从 6.9987 降至 6.5749), KGE 提升(0.8445~0.8612), 体现出对抗训练对误差幅值与一致性具有一定改善作用。

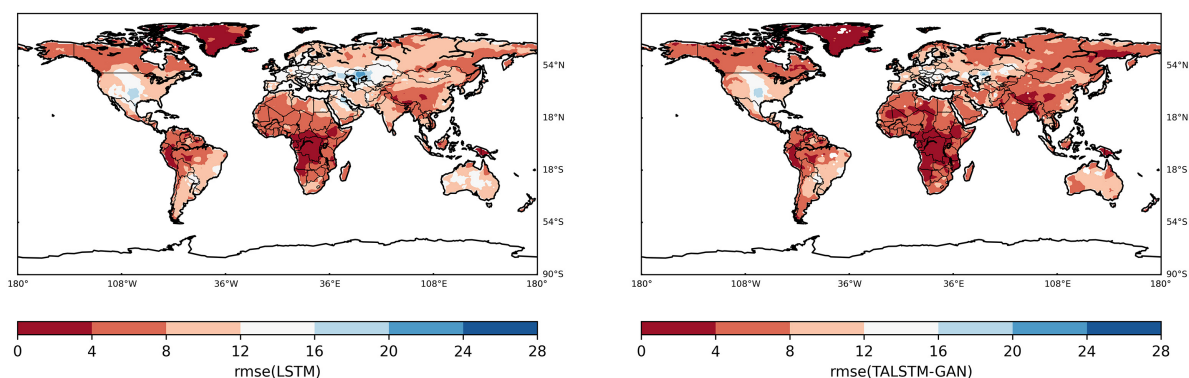
时间注意力框架: TALSTM-GAN 相较 TALSTM 在 RMSE 上进一步略微降低, 并将 R 提升至 0.987、KGE 提升至约 0.958、NSE 提升至约 0.971, 在相关性与水文一致性方面达到所有模型中的最高水平, 说明在时间注意力结构中加入 GAN 有助于改善预测分布与观测分布的一致性, 并提升时序形态拟合的稳健性。

特征注意力框架: FALSTM-GAN 相较 FALSTM 在 RMSE 与 BIAS 上均有小幅降低, KGE 与 NSE 略有提升, 整体性能更均衡, 表明对抗学习能够在特征注意力框架下进一步优化预测分布。

联合注意力框架下增益尤为明显: FTALSTM-GAN 相较 FTALSTM, 其 RMSE 由 7.4028 降至 6.7230, BIAS 由 5.1853 降至 4.6546, 同时 R 与 R^2 提升至约 0.986 和 0.970, KGE 与 NSE 分别达到约 0.957 和 0.970。结果表明, GAN 显著抑制了系统偏差并提升误差稳定性, 使其在“误差类指标更优、相关性指标接近最优”的前提下获得更均衡的综合表现。

3.4. 空间分布特征与区域适应性分析

为进一步分析模型在全球不同区域上的适应性差异, 本文补充给出了基线模型 LSTM 以及最优模型 TALSTM-GAN 在测试集上的 RMSE 和 KGE 空间分布图, 如图 1 和图 2 所示。相较于表 1 中基于全球网格点中位数的总体统计结果, 空间分布图能够更直观地揭示模型在不同气候背景和地表条件下的性能差异。



注：文中图片全部由作者基于实验结果使用 python 自行绘制。空间分布图的地图边界来源于 python 的 Basemap 与 Matplotlib 绘图工具。

Figure 1. RMSE spatial distribution of baseline model LSTM and TALSTM-GAN

图 1. 基线模型 LSTM 与 TALSTM-GAN 的 RMSE 空间分布图

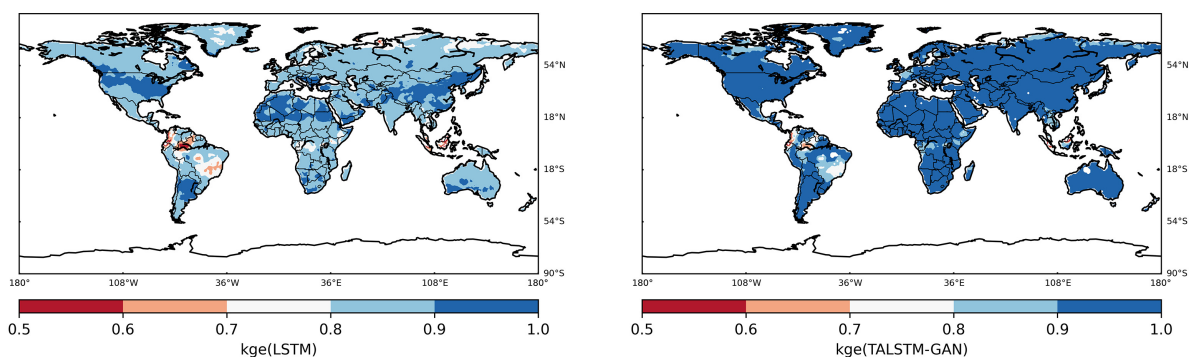


Figure 2. KGE spatial distribution of baseline model LSTM and TALSTM-GAN

图 2. 基线模型 LSTM 与 TALSTM-GAN 的 KGE 空间分布图

从 RMSE 的空间分布来看，TALSTM-GAN 在多数陆地区域相较基线模型 LSTM 表现出更低的误差水平，说明引入时间注意力与对抗学习后，模型在全球范围内的误差控制能力得到了整体改善。与此同时，不同区域之间仍存在明显差异：在气候季节性较强、驱动关系相对稳定的区域，模型误差普遍较低；而在部分复杂地形区、气候过渡带以及高纬地区，RMSE 仍相对较高，表明统一模型对局地异质性过程的刻画仍存在一定不足。

从 KGE 的空间分布来看，TALSTM-GAN 在大多数区域保持了较高的一致性水平，且相较于 LSTM 呈现出更广泛的高值分布范围。这说明改进模型不仅在误差幅值上有所优化，而且在相关性、偏差和变异性等综合水文一致性方面也更接近观测。整体而言，空间分布结果与表 1 中的统计指标结论相一致，即改进模型的优势并非仅体现在少数网格点，而是在全球多数区域具有较好的普遍性。

需要指出的是，不同区域之间的性能差异可能与气候分区和地表环境密切相关。在湿润区、季风区或季节循环明显区域，PET 的月际变化通常具有较强规律性，模型更容易从历史窗口中学习稳定关系；而在高海拔地区、干旱半干旱过渡区以及受地形起伏、地表异质性影响较强的区域，PET 的形成机制更复杂，统一模型的表现相对受限。

3.5. GAN 对分布形态的改善验证

为直接验证 GAN 在分布层面的作用，本文进一步比较了观测值、TALSTM 预测值和 TALSTM-GAN

预测值的概率密度函数(PDF)与累积分布函数(CDF), 结果如图 3 所示。总体来看, 两种模型的预测分布均与观测分布较为接近, 说明在多源驱动因子支持下, 月尺度 PET 的总体分布可以被较好拟合。在此基础上, 引入 GAN 后, 预测分布在整体形态上与观测分布保持了较高一致性, 表明对抗学习有助于提升分布层面的相似性。但从总体 PDF/CDF 曲线来看, 这种改善更主要体现为分布形态的稳健贴合, 而非所有统计量上的显著变化。

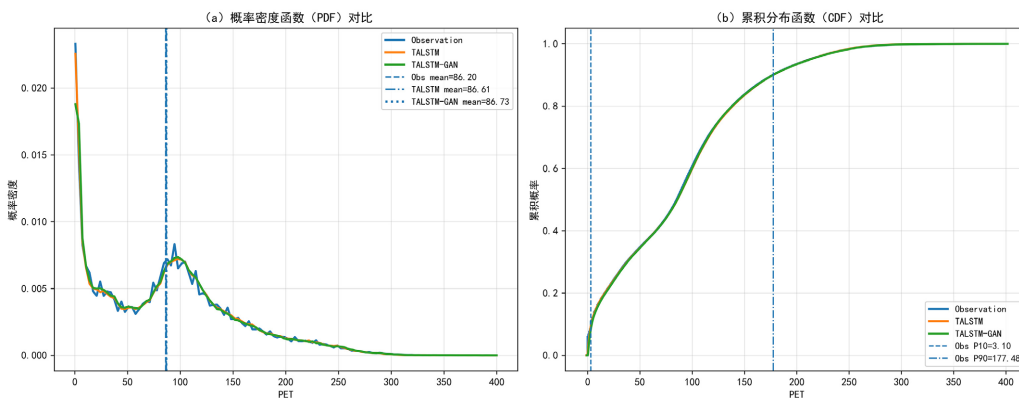


Figure 3. Comparison of probability density function and cumulative distribution function of observed values, TALSTM and TALSTM-GAN predicted values

图 3. 观测值、TALSTM 与 TALSTM-GAN 预测值的概率密度函数和累积分布函数对比

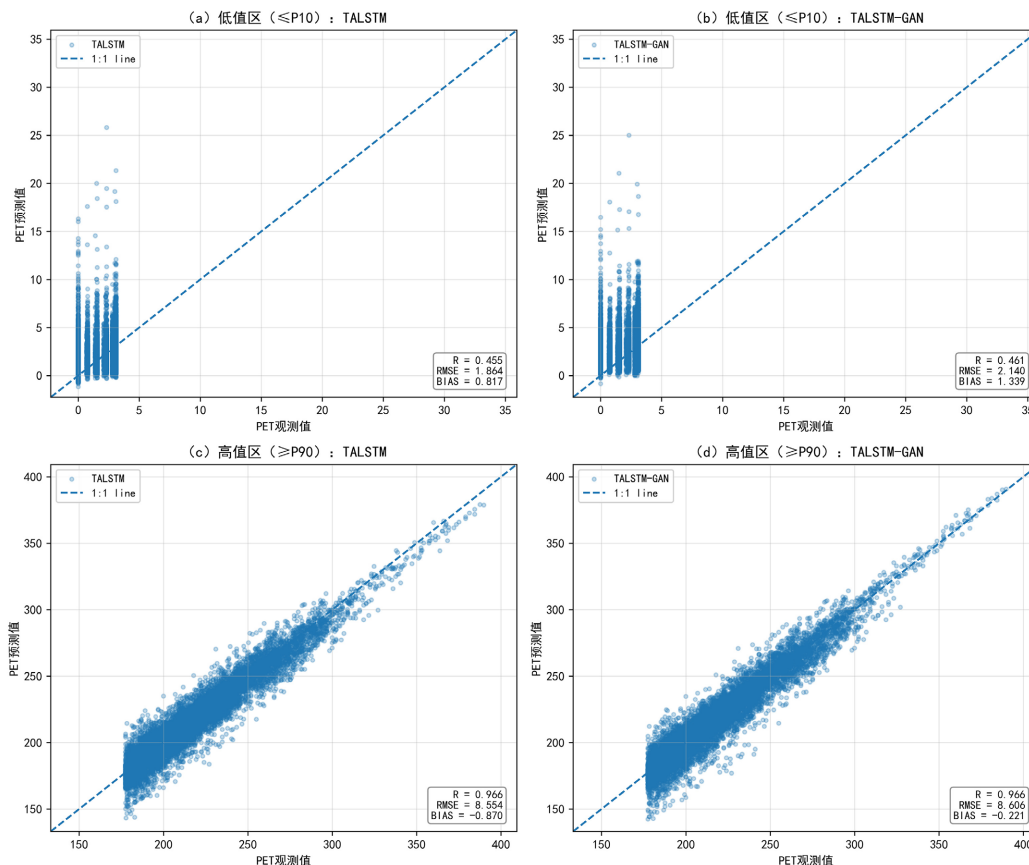


Figure 4. Comparison of predicted-observed scatters of TALSTM and TALSTM-GAN in low-value and high-value areas

图 4. TALSTM 与 TALSTM-GAN 在低值区和高值区的预测值 - 观测值散点对比

进一步地, 本文依据观测值的第 10 百分位和第 90 百分位分别构建低值区和高值区样本, 并绘制 TALSTM 与 TALSTM-GAN 的预测值 - 观测值散点图, 如图 4 所示。结果显示, 在高值区, TALSTM-GAN 的散点分布更接近 1:1 参考线, 且系统偏差有所减弱, 说明对抗学习对高值区预测压缩现象具有一定缓解作用; 而在低值区, 两种模型差异相对较小, GAN 带来的改善并不显著, 部分误差指标甚至未表现出一致优势。因此, GAN 的增益主要体现在分布上端样本的偏差缓解, 而在低值段尚未表现出一致优势。

综合来看, GAN 的主要作用体现在提升预测分布与观测分布在整体形态上的相似性, 以及改善高值区的偏差表现。对于其是否能够显著增强极端事件的预测能力, 仍需结合高分位误差、极端事件命中率等更专门的指标进一步验证。

3.6. 跨网格稳定性与分布特征分析

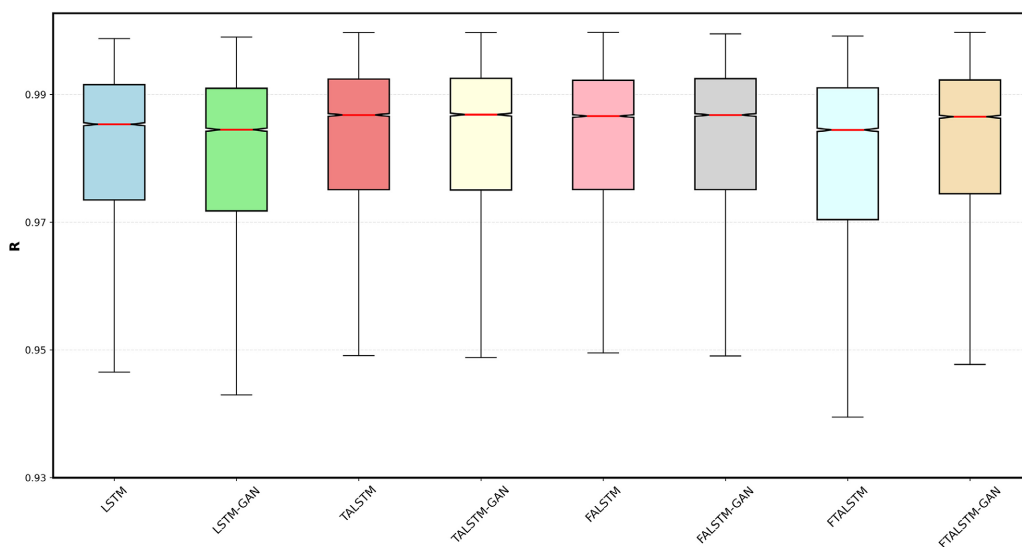


Figure 5. R index box diagram of different models

图 5. 不同模型的 R 指标箱型图

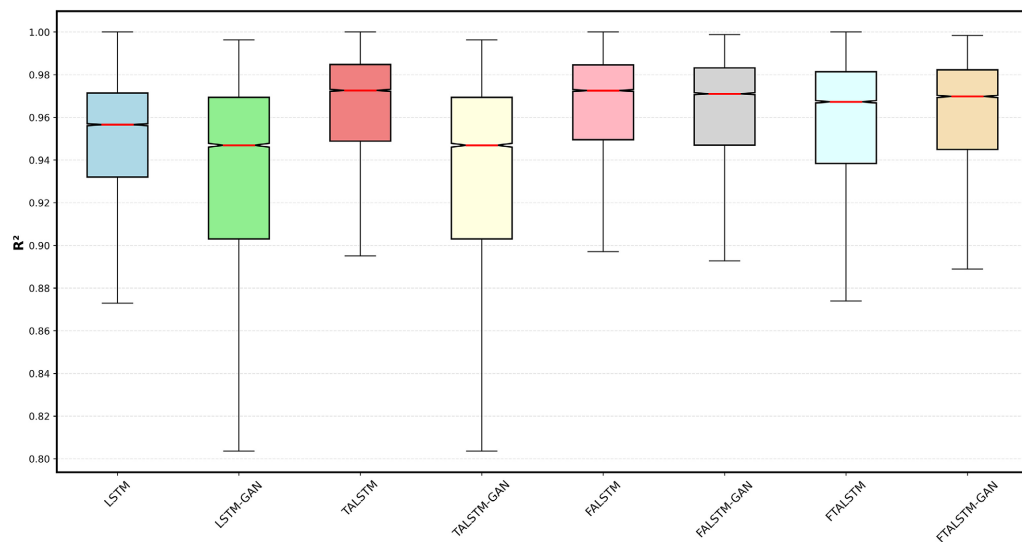


Figure 6. R² index box plots of different models

图 6. 不同模型的 R² 指标箱型图

除中值对比外, 本文还从网格点分布角度给出了 R 、 R^2 、KGE 与 NSE 的箱型图(图 5~8), 用于观察模型在全球不同区域上的稳定性差异。

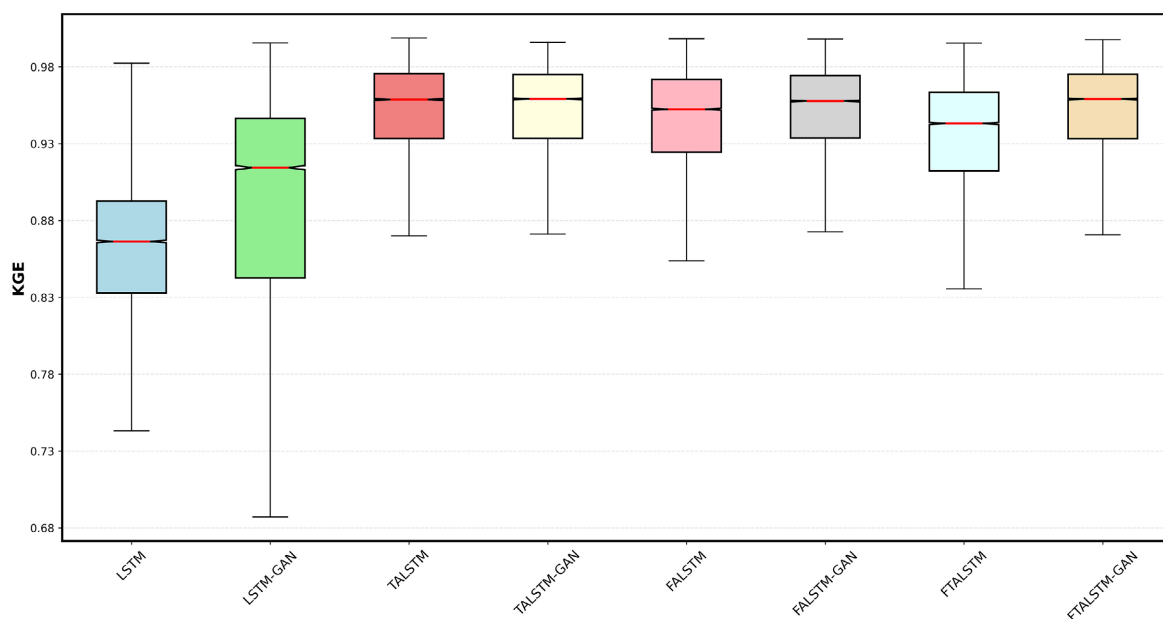


Figure 7. KGE index box plots of different models

图 7. 不同模型的 KGE 指标箱型图

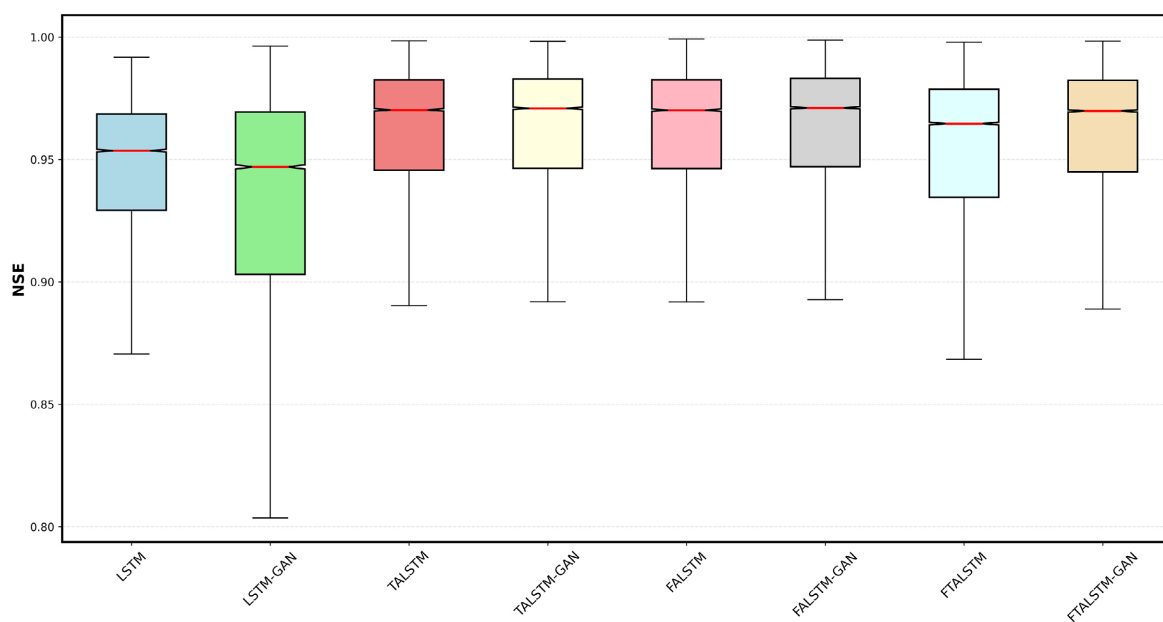


Figure 8. NSE index box diagram of different models

图 8. 不同模型的 NSE 指标箱型图

从箱线图可见, 多数模型在 R 、 R^2 、KGE 与 NSE 上均维持较高水平, 说明框架具备较强的全球一致性。相较基线 LSTM, 注意力模型与其 GAN 扩展模型的中位数普遍更高, 且部分指标的箱体更窄、离群点更少, 表明在更多网格点上性能更稳定。

3.7. 最优模型选择

为保证模型推荐具有可复现性, 本文优先考虑综合一致性指标 KGE 与 NSE, 并结合 RMSE 与 BIAS 共同判断模型性能。由表 1 可知, TALSTM-GAN 在 KGE、NSE 及相关性指标上整体表现最优, 同时保持了较低的 RMSE 和 BIAS; 此外, 从图 1~4 可以看出, 该模型不仅在全球多数区域具有较好的空间适应性, 而且在分布形态上与观测更为一致。因此, 本文将 TALSTM-GAN 作为最终推荐模型。与此同时, FTALSTM-GAN 在误差类指标上表现较为均衡, 可作为对误差幅值更敏感场景下的备选方案。

4. 讨论

本文结果表明, 相较基线 LSTM, 引入注意力机制后, RMSE 与 BIAS 明显下降, 且 KGE、NSE 等一致性指标整体提升, 说明在全球尺度、长序列、多变量驱动的 PET 预测任务中, “选择性关注”是提升预测性能的关键能力。其原因在于 PET 受季节循环和跨月累积影响显著, 历史窗口中不同月份对下一月 PET 的贡献存在差异。时间注意力(TALSTM)能够突出关键月份的信息, 更有效地捕捉季节跃迁以及异常热力和辐射背景引起的月际波动。同时, 不同气候区的主导驱动因子差异较大, 特征注意力(FALSTM)通过动态调整变量权重, 帮助模型在统一结构下兼顾区域差异, 并抑制冗余特征带来的噪声传播。虽然联合注意力(FTALSTM)在单项指标上可能并非最优, 但在多指标权衡下, 它展现了更为稳健的表现, 突显了其在处理时间和变量双重异质性时的协同刻画能力。

本研究引入 GAN 的动机是为了缓解传统回归训练中可能出现的“均值收缩”问题, 即在 MSE 和 MAE 目标的驱动下, 模型更倾向于输出接近均值的预测, 从而影响了对分布形态的表达。实验结果表明, 在多种结构中加入 GAN 后, BIAS 普遍降低, KGE、NSE 等一致性指标也有所提升。尤其是 FTALSTM-GAN, 在 RMSE 和 BIAS 上的增益更为明显, 展示了更均衡的误差控制能力。总体来看, 对抗训练有助于提升预测结果与观测序列在分布层面的整体一致性, 并对高值区偏差表现出一定改善; 至于其对尾部尤其是极端事件预测能力的提升, 仍需通过高分位误差、超阈值统计或极端事件命中率等更专门的指标进一步验证。

本文采用单一模型覆盖全球网格点, 其优势在于: PET 形成机制具有共性基础(能量供给、热力状态与水汽条件的耦合), 统一模型通过参数共享学习可迁移规律; 同时, 在样本相对稀缺或不均衡区域, 统一训练具有一定“跨区域迁移”效应, 可降低局地过拟合并提升泛化能力。相较于区域分割模型, 统一建模亦更利于工程部署与公平对比。

结合图 1 和图 2 所示的空间分布结果可以进一步看出, 统一模型的性能提升并非在所有区域均匀发生, 而是呈现出明显的空间异质性。总体上, 改进模型在多数区域相较基线模型表现出更低的 RMSE 和更高的 KGE, 说明统一建模能够在全局范围内学习到较强的共性规律; 但在部分高海拔地区、干旱半干旱过渡区、高纬寒冷区以及可能受季风转换影响较强的区域, 误差仍相对较大、一致性相对较低。这表明即使在参数共享的统一框架下, 不同区域的局地机制差异仍可能导致模型难以达到一致的最优表现。

需要指出的是, 统一模型也存在一定的局限性。空间误差图显示, 复杂地形区、气候过渡带以及高纬寒冷区仍是误差相对较高或一致性相对较低的区域, 这说明单一参数集合难以在所有区域同时实现局地最优。造成这种区域差异的潜在原因可能包括: 第一, 不同气候带中 PET 的主导控制因子及其相互作用方式并不完全一致, 统一模型虽然能够学习跨区域共性规律, 但对局地特有驱动关系的刻画仍有限; 第二, 复杂地形区域的热力条件、辐射收支和水汽输送过程更为复杂, 而本文当前输入变量以动态气象-水文因子为主, 尚未显式引入地形、植被、土壤等静态环境信息, 这可能削弱模型对局地差异的表达能力; 第三, 目标 PET 数据来自 CRU 产品并被重采样至 $1^\circ \times 1^\circ$ 网格, 空间插值和平滑处理可能在一定程度上削弱小尺度地表异质性和局地极端, 从而影响复杂区域的误差表现。尽管注意力机制在一定程

度上能够缓解区域异质性问题, 但统一模型在表达局地特例方面仍有进一步的改进空间, 未来可考虑引入区域嵌入、静态因子或分层建模策略以增强模型的区域适应性。

此外, 本文结论仍受以下不确定性影响:

- 1) 目标 PET 来自 CRU 产品并重采样至 1° 网格, 插值、重采样可能平滑局地极端。
- 2) 标准化与缺测处理虽有利于训练稳定, 但可能影响尾部分布区域的相对尺度, 从而影响极值相关评估。
- 3) 对抗学习的效果受模型容量、判别器强度和训练稳定性的影响。可能需要更严格的训练策略来检验模型的鲁棒性。

综合来看, 注意力机制可以作为全球 PET、ET 等多因子耦合变量预测的核心组件。在时间和变量的异质性得到有效建模后, 引入分布层面对齐更容易发挥增益作用。如果优先考虑一致性指标, 建议模型更加注重 KGE 和 NSE 的表现; 而在对误差敏感的场景中, 比如 FTALSTM-GAN 这样更为均衡的结构则可以作为备选方案。面向实际应用, 后续仍有必要进一步开展极端事件导向的专门评估与区域敏感性分析, 并结合更多静态因子或区域嵌入, 以提升模型对局地特殊过程的解释能力和区域适应性。尤其是针对空间误差图中表现较弱的复杂地形区、过渡气候区和高纬区域, 后续需要开展更有针对性的区域敏感性分析与结构改进, 以提升统一模型在全球不同环境背景下的稳健性。

5. 结论

本文基于 1981~2020 年、 $1^\circ \times 1^\circ$ 全球统一网格多源再分析气候数据, 构建并评估了“LSTM + 注意力机制 + 对抗分布学习(GAN)”的 PET 预测统一框架。主要结论如下:

1) 在全球月尺度任务上, 模型整体具有较强可预测性; 注意力机制可有效降低 RMSE 与 BIAS, 并提升 KGE、NSE 等一致性指标, 其中时间注意力在误差控制方面更具优势; 进一步引入 GAN 后, 模型在分布形态相似性和综合一致性方面表现出一定增益, 尤其在高值区偏差控制上有所改善, 但其对极端事件预测能力的提升仍需进一步验证。

2) 从空间分布结果来看, 最优模型 TALSTM-GAN 在全球多数区域相较基线模型 LSTM 表现出更低的误差和更高的一致性, 说明改进模型的性能提升具有较好的空间普遍性。然而, 在复杂地形区、干旱半干旱过渡区及高纬寒冷区, 模型误差仍相对较高, 表明区域异质性仍是全球统一建模需要进一步解决的重要问题。

3) 从观测值与预测值的 PDF/CDF 对比以及高低值区散点图来看, 引入 GAN 后, 模型预测分布与观测分布在整体形态上保持了较高一致性, 并在高值区偏差控制方面表现出一定改善。这说明分布层面对齐有助于提升模型的整体水文一致性和分布形态相似性。

4) 需要指出的是, 本文现有证据主要支持 GAN 对预测分布整体形态和综合一致性的改善作用。对于其是否能够显著增强极端高 PET 事件的预测能力, 仍需结合高分位误差、极端事件命中率或超阈值统计等更专门的指标进一步验证。

总体而言, “注意力增强时序建模 + 分布对齐训练 + 全球统一训练”能够以较高效率获得稳定的全球预测性能。若以一致性指标(KGE、NSE)及分布形态相似性为优先进行模型推荐, 本文框架可为干旱监测与水资源管理等应用中的 PET 评估提供具有推广潜力的数据驱动方案。

参考文献

- [1] Gebrechorkos, S.H., Peng, J., Dyer, E., Miralles, D.G., Vicente-Serrano, S.M., Funk, C., *et al.* (2023) Global High-Resolution Drought Indices for 1981-2022. *Earth System Science Data*, **15**, 5449-5466. <https://doi.org/10.5194/essd-15-5449-2023>

-
- [2] Noguera, I., Vicente-Serrano, S.M. and Domínguez-Castro, F. (2022) The Rise of Atmospheric Evaporative Demand Is Increasing Flash Droughts in Spain during the Warm Season. *Geophysical Research Letters*, **49**, e2021GL097703. <https://doi.org/10.1029/2021gl097703>
- [3] He, Q.L., Xiao, J.L. and Shi, W.Y. (2022) Responses of Terrestrial Evapotranspiration to Extreme Drought: A Review. *Water*, **14**, Article 3847. <https://doi.org/10.3390/w14233847>
- [4] Malamos, N. and Tegos, A. (2022) Advances in Evaporation and Evaporative Demand. *Hydrology*, **9**, Article 78. <https://doi.org/10.3390/hydrology9050078>
- [5] Singer, M.B., Asfaw, D.T., Rosolem, R., Cuthbert, M.O., Miralles, D.G., MacLeod, D., *et al.* (2021) Hourly Potential Evapotranspiration at 0.1° Resolution for the Global Land Surface from 1981-Present. *Scientific Data*, **8**, Article No. 224. <https://doi.org/10.1038/s41597-021-01003-9>
- [6] Li, S., Wang, G., Zhu, C., Lu, J., Ullah, W., Hagan, D.F.T., *et al.* (2022) Attribution of Global Evapotranspiration Trends Based on the Budyko Framework. *Hydrology and Earth System Sciences*, **26**, 3691-3707. <https://doi.org/10.5194/hess-26-3691-2022>
- [7] Christian, J.I., Basara, J.B., Hunt, E.D., Otkin, J.A., Furtado, J.C., Mishra, V., *et al.* (2021) Global Distribution, Trends, and Drivers of Flash Drought Occurrence. *Nature Communications*, **12**, Article No. 6330. <https://doi.org/10.1038/s41467-021-26692-z>
- [8] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., *et al.* (2019) Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature*, **566**, 195-204. <https://doi.org/10.1038/s41586-019-0912-1>
- [9] Ma, Z., Chen, D., Sun, P., Liu, Y. and Ye, L. (2025) Actual Evapotranspiration Dominated Net Primary Productivity Loss Caused by Concurrent Droughts. *Ecological Indicators*, **176**, Article 113645. <https://doi.org/10.1016/j.ecolind.2025.113645>
- [10] Sungmin, O. and Park, S.K. (2024) Global Ecosystem Responses to Flash Droughts Are Modulated by Background Climate and Vegetation Conditions. *Communications Earth & Environment*, **5**, Article No. 88. <https://doi.org/10.1038/s43247-024-01247-4>
- [11] Le, X., Ho, H.V., Lee, G. and Jung, S. (2019) Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water*, **11**, Article 1387. <https://doi.org/10.3390/w11071387>
- [12] Hasan, I.A. and Yuce, M.I. (2024) Prediction of Potential Evapotranspiration via Machine Learning and Deep Learning for Sustainable Water Management in the Murat River Basin. *Sustainability*, **16**, Article 11077. <https://doi.org/10.3390/su162411077>
- [13] Roy, D.K., Sarkar, T.K., Kamar, S.S.A., Goswami, T., Mukhtadir, M.A., Al-Ghobari, H.M., *et al.* (2022) Daily Prediction and Multi-Step Forward Forecasting of Reference Evapotranspiration Using LSTM and Bi-LSTM Models. *Agronomy*, **12**, Article 594. <https://doi.org/10.3390/agronomy12030594>
- [14] 邢立文, 崔宁博, 董娟. 基于 LSTM 深度学习模型的华北地区参考作物蒸散量预测研究[J]. 水利水电技术, 2019, 50(4): 64-72.
- [15] 葛杰, 康秀, 白钰, 等. LSTM 模型在嘉陵江流域参考作物蒸散量计算中的适用性研究[J]. 中国农村水利水电, 2024(11): 62-70.
- [16] Harris, I., Osborn, T.J., Jones, P. and Lister, D. (2020) Version 4 of the CRU TS Monthly High-Resolution Gridded Multivariate Climate Dataset. *Scientific Data*, **7**, Article No. 109. <https://doi.org/10.1038/s41597-020-0453-3>
- [17] Li, Q., Zhang, C., Shangguan, W., Wei, Z., Yuan, H., Zhu, J., *et al.* (2024) Landbench 1.0: A Benchmark Dataset and Evaluation Metrics for Data-Driven Land Surface Variables Prediction. *Expert Systems with Applications*, **243**, Article 122917. <https://doi.org/10.1016/j.eswa.2023.122917>