

论文标题: “Threshold Effects and Generalization Bias in AI-based Urban Pluvial Flood Prediction: Insights from a Dataset Design Perspective.” **作者:** Hao Hu 等
该手稿对 AI-based 城市暴雨洪水预测中一个重要但常被忽视的问题——数据集设计——进行了及时且结构严谨的探讨。研究并未仅聚焦于模型架构,而是系统性地评估了数据集长度、特征组成以及降雨强度分布如何影响预测技能、计算成本和泛化能力。总体而言,这是一项扎实且具有相关性的贡献,我认为手稿在经过少量修改后值得发表。

1. 手稿应提高统计学术语的一致性。第 2.6 节提到 MANOVA, 而第 4.3 节将分析描述为多因素 ANOVA。作者应明确说明实际使用的是哪种方法,并确保全文术语一致。

回复: 修订稿 2.5.2 节统一为“多因素方差分析 (Multi-factor ANOVA)”,并在 3.2.4 节敏感性分析中明确“MANOVA 用于多因变量 (NRMSE、 R^2 、训练时间) 联合分析,单指标采用 ANOVA”。

2. 目标变量的描述应更加一致。手稿某些部分将预测目标定义为淹没面积,而其他部分则讨论峰值水位、径流总量和水深。

回复: 修订稿全程统一为“地面积水面积 Y ” (表、公式、评价指标分母均为 \bar{Y})。

3. 结论部分的某些措辞比典型科学论文更强烈且更口语化。例如“respect the 14k-sample ceiling”和“start lean, enrich later”等短语,建议改为更正式的学术风格。这些观点很有价值,但语气可更加精确和中性。

回复: 修订稿讨论、结论 (5、6 节) 和讨论已改为正式学术表述 (如“性能增益趋于平缓”“混合雨量数据集显著提升鲁棒性”),删除“respect the 14k-sample ceiling”等口语。对原版进一步精炼为中性语言。

4. 此外,我强烈建议作者考虑引用以下近期且高度相关的研究,该论文与城市排水网络中的 AI 代理建模以及节点级水力状态预测密切相关: Zafarmomen, N., Samadi, V., and Borgomeo, E. (2026). Spatiotemporal SWMM-LSTM surrogate modeling for efficient node-level water depth and inflow prediction in urban drainage networks. Cambridge University Press, published online 13 January 2026.

总体上,本文选题及时、分析系统,结论具有实际指导意义,仅需针对上述少量问题进行针对性修改即可推荐发表。

专家三

论文审稿意见: **论文标题:** “Threshold Effects and Generalization Bias in AI-based Urban Pluvial Flood Prediction: Insights from a Dataset Design Perspective.” **作者:** Hao Hu 等

该手稿所研究的话题高度相关,聚焦于深度学习模型 (特别是 LSTM) 在城市暴雨洪水预测中数据集设计的优化问题。本研究识别了数据集规模约 14,400 个样本的“阈值效应”,并强调降雨强度分布对模型泛化能力的影响比单纯增加原始数据量更为关键。虽然该工作有趣且动机明确,但仍需在方法论上进行若干关键澄清,并改进结果呈现方式,方可推荐发表。

具体意见:

引言与泛化能力：研究动机清晰且结构严谨。然而，第 87 行指出 LSTM 模型的洞见有望迁移至其他机器学习模型。作者应论证这一说法，解释为何这些发现（尤其是阈值效应）能够适用于其他模型。

回复：修订稿摘要、结论明确：“本文在 LSTM 基准模型框架下得到的规律对其他序列学习模型具有参考意义，但相关推广仍需进一步验证”。

第 2.2 节：文中报告了水力模型的 NSE 值，但需提供更多细节。该 NSE 代表的是哪一个变量（水位、流量还是洪水范围）？在哪些控制点、针对哪些降雨事件进行了验证？

回复：修订稿 2.1 节新增验证表（3 场事件）、明确变量为“管道流量”、控制点为“监测管道”、NSE>0.5。图 1-2 给出过程线对比。

- 第 134 行：所采用的 IDF 公式是否有具体文献引用？

回复：修订稿 2.1 节公式（1）后新增“基于区域设计暴雨强度公式（参考当地水文手册）。

- 第 147 行：分布式模型的方程存在过度简化问题，因为它忽略了单元格之间的侧向流动和地表蓄水。该部分需要更严谨的解释。

回复：修订稿 2.2.2 节补充说明“地表径流量 Q 结合区域平均积水深度与地形参数，经水力转换关系映射为积水面积 Y……基于 InfoWorks ICM 2D 模块浅水方程和 DEM 建立的容积-面积-水深关系曲线”，已说明非纯水量平衡。

- 第 148 行：“III”存在笔误，应为“I”（表示 Infiltration，入渗）。

回复：修订稿统一为“I”（入渗量）。

第 2.3 节：

- 第 161 行：为保持一致性，“Configuration 4”应标记为：“降雨（P）+ 土壤入渗（I）+ 管道排水（D）”。

- 第 172 行：请明确序列长度的单位，是时间步长还是小时？

回复：修订稿 2.2.2 节表 2 明确“组合四：降雨量 P+土壤入渗量 I+管道排水量 D”；2.3 节明确“45 分钟时间序列片段”。

第 3.3 节：

- 第 332 行：关于数据遗漏预防？手稿提到使用了重叠滑动窗口。必须澄清数据划分（训练/验证/测试）是在生成这些窗口之前还是之后进行的。若在之后进行，则存在较高的信息泄漏风险，这将使报告的泛化性能失效。

回复：感谢审稿人提出的重要问题。我们已在修订稿 2.3 节（数据集抽取规则与验证策略）中对滑动窗口生成与数据划分的时序关系进行了明确澄清，并进一步强化了防遗漏措施。同时，修订稿 2.5 节（随机交叉验证）明确采用 5 折交叉验证，且“划分针对每种数据集配置（不同长度、特征组合、雨量分布）单独进行”，基本单元为已打乱的 45 分钟独立时间序列片段，从而严格避免了未来信息遗漏。

第 4.1 节：

- 第 402 行：关于性能在 Level 4 之后衰减的说法，在当前图表中对于低强度和混合强度降雨的情况难以辨识。

回复：修订稿 3.2.1 节补充“当样本量达到约 14,400 个样本后，性能提升趋于平缓”，并在图中用虚线标注阈值。修改思路：可进一步在 3.2.1 节增加分雨量场景的子图或表格数据。

- 第 423 行：模型是否考虑了检查井溢流？这在城市暴雨洪水中是一个关键因素。

回复：修订稿 2.2 节说明“InfoWorks ICM 1D 管网模块基于圣维南方程……2D 模块基于浅水方程”，已隐含溢流处理。

总体意见：

- MANOVA 与 ANOVA 的使用存在混淆。描述内容显示是对单个指标进行因子 ANOVA 分析，但又提到了 MANOVA。请明确说明所采用的精确统计框架，以及如何处理多个因变量之间的协方差。

回复：修订稿 2.5.2 节统一 MANOVA 框架，明确“MANOVA 用于多指标联合分析”。

- 手稿存在一个重大遗漏，即生成数据集所需的时间。由于 InfoWorks ICM 模拟计算量大，作者必须提供总模拟时间、所用硬件配置，并对比“数据生成投入时间”与 AI 实时预测的优势。

修改思路（数据生成时间）：可在 2.1 节技术路线或讨论中新增“InfoWorks ICM 单场模拟耗时约 XX 分钟，总生成时间约 X 小时（Intel i9+RTX 3090）”，对比 LSTM 训练时间，突出“数据成本 vs 预测收益”。

图表与可读性：

- 图 7：“Y0333333.1”代表什么？图例标签应具有描述性且自解释。

回复：已删除冗余图例标签

- 图 8、图 9 和图 11：图 8（Training time 1、2、3）难以阅读且似乎存在冗余。建议将核心信息合并至图 9，并若图 11 未提供独特洞见则予以删除。

回复：对图 8、图 9 和图 11：图 8（Training time 1、2、3）进行修改，考虑换成清晰的趋势变化对比图。

- 图 17：请澄清“Value”纵轴的含义。目前该图不足以支撑其下方得出的结论。

回复：图 17 拟删除。

讨论部分：讨论缺乏足够的文献引用来支撑其观点，特别是第 534 行。研究结果应与计算水文学领域中现有关于数据集设计的文献进行对比，以更好地进行情境化分析。

结论： 本文将 AI 在水文学中的应用焦点从“模型架构”转向“数据集设计”，是一项及时且有价值的贡献。然而，仍存在若干关键问题需要解决。首先，关于数据生成计算成本的透明度不足，使得难以评估所提出框架的效率，因为“数据成本”与模型精度同等重要。其次，滑动窗口方法中潜在的数据泄漏问题必须解决，以确保结果的有效性。最后，水文表示存在过度简化，质量守恒考虑和校准流程描述过于简略。加强合成数据的物理基础至关重要。为推荐该工作，作者需针对上述要点进行修改，并解决计算可行性、数据处理以及模拟水文基础方面的相关顾虑。