

基于IV-RF模型地质灾害易发性及可解释性分析

陈菊芳

云南师范大学地理学部, 云南 昆明

收稿日期: 2026年4月10日; 录用日期: 2026年5月11日; 发布日期: 2026年5月29日

摘要

云南省地质构造复杂, 地质灾害多发易发, 开展地质灾害易发性预测对防灾减灾意义重大。本研究以镇雄县为研究区域, 采用信息量模型(Information Value, IV)量化了各因子属性值对灾害的影响, 并耦合随机森林模型(Random Forest, RF)建立评价模型同时, 采用沙普利可加和解释方法(Shapley Additive Explanations, SHAP)对模型进行可解释性分析, 以明确各因子对地质灾害的影响。研究收集了镇雄县地质灾害点及16个影响因子数据, 经处理后用于模型构建与验证。结果表明, IV-RF复合模型ROC曲线下面积(AUC)达到0.93, 具有较高的预测精度和可靠性, 且地质灾害易发性呈现显著空间差异, 极高和高易发区集中在县域中部、南部及东部部分区域。SHAP分析显示, 高程(DEM)、距道路距离(Distance to Roads, DistRd)、归一化植被指数(NDVI)等因子对地质灾害影响显著, 各因子信息量值增大通常与地质灾害易发性增加相关。本研究为区域地质灾害风险管理提供了科学有效的工具和重要参考。

关键词

地质灾害, 易发性评价, 信息量模型, 随机森林模型, 可解释机器学习

Susceptibility and Interpretability Analysis of Geological Hazards Based on the IV-RF Model

Jufang Chen

Faculty of Geography, Yunnan Normal University, Kunming Yunnan

Received: April 10, 2026; accepted: May 11, 2026; published: May 29, 2026

Abstract

Yunnan Province is characterized by complex geological structures and a high frequency of geological

hazards. Accurate prediction of geological hazard susceptibility is of great significance for disaster prevention and mitigation. Taking Zhenxiong County as the study area, this study uses the Information Value (IV) model to quantify the influence of attribute values of each factor on hazards, and couples it with the Random Forest (RF) model to construct an evaluation model. Meanwhile, the Shapley Additive Explanations (SHAP) method is employed to conduct interpretability analysis of the model, so as to clarify the impact of each factor on geological hazards. Data on geological hazard locations and 16 influencing factors in Zhenxiong County were collected, processed, and then used for model construction and validation. The results show that the area under the ROC curve (AUC) of the IV-RF hybrid model is 0.93, indicating high prediction accuracy and reliability. In addition, geological hazard susceptibility shows significant spatial heterogeneity, with very high and high susceptibility zones concentrated in the central, southern and parts of eastern regions of the county. SHAP analysis reveals that factors such as elevation (DEM), distance to roads (DistRd), and Normalized Difference Vegetation Index (NDVI) impose significant impacts on geological hazards, and the increase in information values of each factor is generally positively correlated with the increase in geological hazard susceptibility. This study provides a scientific and effective tool and an important reference for regional geological hazard risk management.

Keywords

Geological Hazards, Susceptibility Assessment, Information Value Model, Random Forest Model, Interpretable Machine Learning

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

地质灾害是指由自然因素或人为活动引发的危害人民生命和财产安全的山体崩塌、滑坡、泥石流、地面塌陷、地裂缝和地面沉降等与地质作用有关的灾害[1]。云南省复杂的地质构造与地质环境条件致使滑坡、泥石流、崩塌等地质灾害多发易发[2]，因此开展易发性预测不仅对中国防灾减灾具有重要意义，也为区域地质灾害风险管理提供了重要参考。

为有效防范和减轻地质灾害的影响，研究者不断挖掘与分析多源数据，构建不同的模型去评估分析地质灾害易发性。地质灾害评价模型可分为数理统计模型和数据驱动模型，其中数理统计模型有信息量模型(LV) [3]、证据权法[4]、确定系数法等，数据驱动模型主要有逻辑回归法[5]、支持向量机法[6]、随机森林模型(RF) [7]等。LV 模型作为一种基于概率的统计方法，通过计算每个孕灾因子在历史灾害发生与未发生区域的概率分布差异来评估该因子的致灾贡献。然而，该模型对于处理大规模非线性关系、交互效应以及高维数据时具有局限性。RF 模型作为一种强大的机器学习方法，能够处理大量的数据样本，并且在处理坡度、岩性、高程等地质环境因子与滑坡发生间的复杂非线性关系时表现出色。然而，在构造 - 地貌分区显著的区域，该模型对空间异质性适应不足，易出现局部过拟合，且该模型存在黑箱性问题；面对样本量稀少的滑坡类型时，识别性能也明显下降。

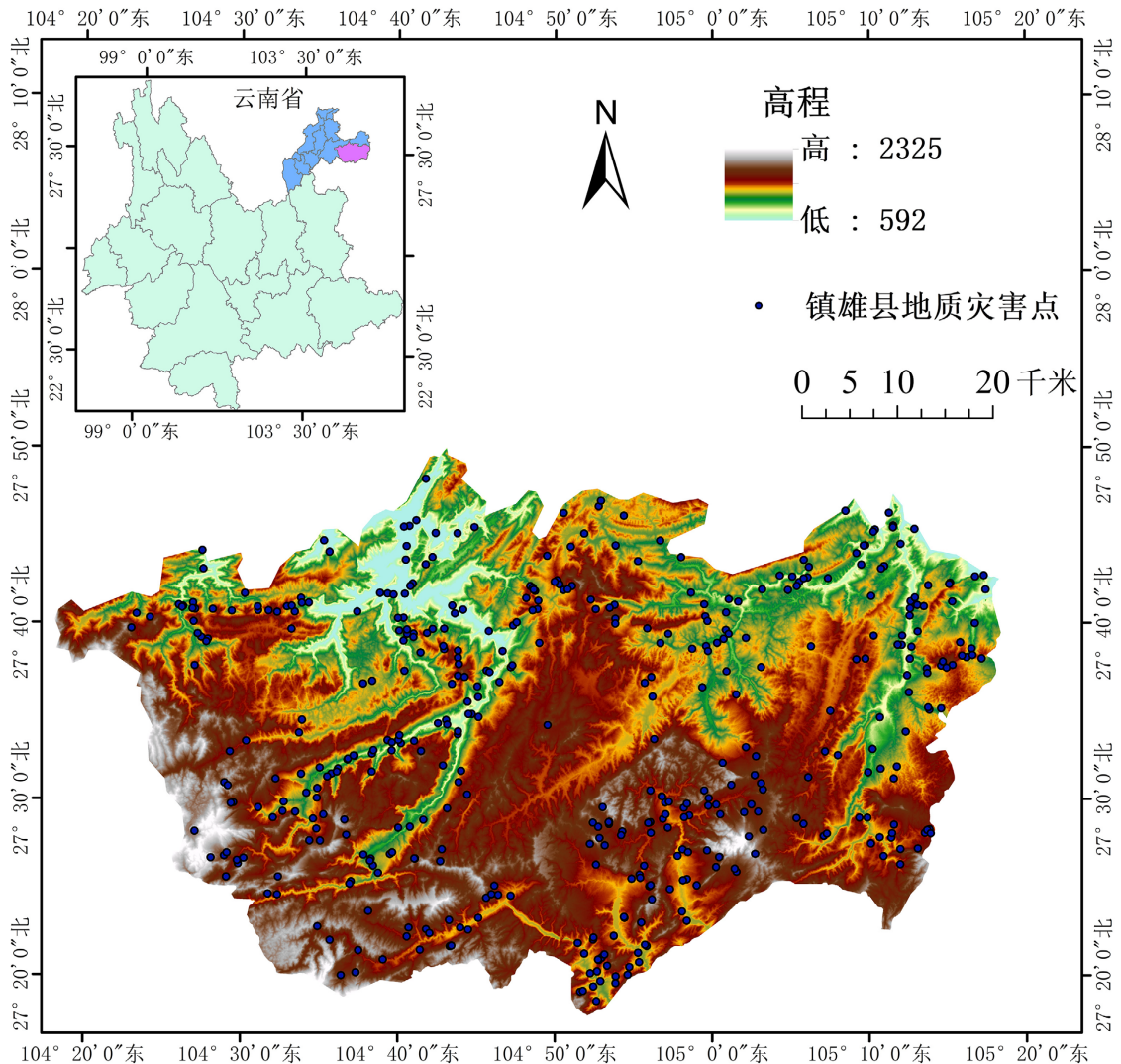
基于此，本文以镇雄县为研究区，建立信息量 - 随机森林(Information Value-Random Forest, IV-RF)复合模型，该模型将 LV 模型与 RF 模型有机结合，充分发挥了二者优势。一方面，它利用 LV 模型对各孕灾因子的致灾贡献进行初步量化，为 RF 模型提供更具物理意义和可解释性的输入特征，有效解决了 RF 模型在构造 - 地貌分区显著区域对空间异质性适应不足、易局部过拟合的问题，同时在一定程度上缓解

了面对样本量少滑坡类型时识别性能下降的状况；另一方面，借助 RF 模型强大的数据处理能力和对复杂非线性关系的捕捉能力，弥补了 LV 模型在处理大规模非线性关系、交互效应以及高维数据时的局限性，从而显著提升了地质灾害易发性评价的精度和可靠性。为解决 RF 模型的黑箱性问题，采用沙普利可加和解释方法(Shapley Additive Explanations, SHAP)对模型进行可解释性分析。

2. 研究区概况与数据来源

2.1. 研究区概况

镇雄县位于云南省西部，隶属于昭通市，地处滇东北山区，东邻曲靖市，西界四川省，南接贵州省。如图 1，境内山势险峻、地表破碎、水系发育、降雨集中，地势总体由西南向东北逐渐递减。地质构造复杂，属于云贵高原的一部分，地质构造上主要受扬子地块和康滇地块的交汇影响[8]。地形复杂，山地、丘陵和高原交错，是典型的山区县[9] [10]。



注：该图基于自然资源部标准底图服务网站下载的审图号为 GS(2024)0650 号的标准地图制作，底图无修改。

Figure 1. Overview map of the research area

图 1. 研究区概况图

2.2. 数据来源

本实验所用到的数据包括镇雄县的地质灾害点, 以及 16 个影响因子, 包括: 高程(DEM)、坡度(Slop)、坡向(Aspect)、地形起伏度(Terrain Ruggedness Index, TRI)、高程变异系数(Elevation Variation Coefficient, EVC)、地形粗糙度(Surface Roughness, SR)、地表切割深度(Surface Incision, SI)、水流强度指数(Stream power index, SPI)、地形湿度指数(Topographic wetness index, TWI)、岩性(Lithology, LIT)、土壤类型(Soil Type, ST)、距断层距离(Distance to Faults, DistF)、距离河流的距离(Distance to Rivers, DistR)、距道路距离(Distance to Roads, DistRd)、土地利用(LULC)、归一化植被指数(NDVI)。本文将各类影响因子统一处理为栅格格式, 并重采样至 30 m × 30 m 的空间分辨率, 在统一坐标系下进行配准、裁剪与分析。数据来源见表 1。

Table 1. Research data and sources

表 1. 研究数据及来源

数据类型	地质灾害点及影响因子	数据源
地质灾害点	历年地质灾害点	地理遥感生态网 https://gisrs.cn
地貌数据	DEM、Slop、Aspect、TRI、EVC、SR、SI	地理空间数据云 https://www.gscloud.cn/
地质数据	LIT	国际土壤研究与信息中心
	DistF	地质科学数据出版社系统 1:20 万地质图
	ST	中国科学院土壤科学研究所
水文数据	DistR	OpenStreetMap https://github.com/openstreetmap
	SPI、TWI	地理空间数据云 https://www.gscloud.cn/
人类活动数据	DistRd	OpenStreetMap https://github.com/openstreetmap
	LULC、NDVI	地球资源数据源 https://gis5g.com/home

3. 研究方法

通过耦合 LV 模型和 RF 模型研究镇雄县地质灾害易发性的空间分布以及模型的可解释性分析, 主要内容及方法包括:

- 1) 计算影响因子的信息量值。检查 16 个影响因子的多重共线性, 使用 LV 模型计算筛选后的因子二级分类的信息量值。
- 2) 构建 LV_RF 耦合模型。将 IV 模型计算获得的各因子信息量值作为输入特征值, 输入 RF 模型中计算地质灾害易发性。
- 3) 基于可解释模型的驱动因子分析。运用 SHAP 解释模型, 以明确各因子对地质灾害的影响。技术路线图如图 2。

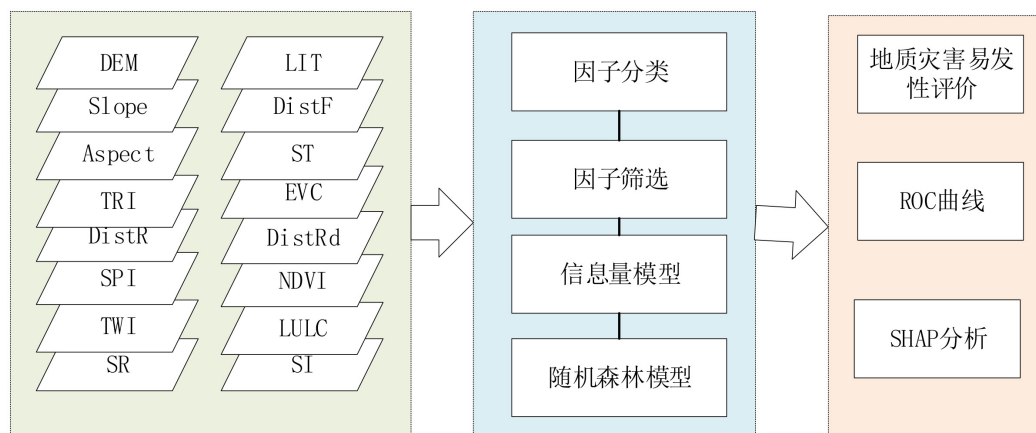


Figure 2. Technical process

图 2. 技术流程

3.1. 因子筛选

为了确保所选取因子的相互独立性，需要对这些因子进行多重共线性检验和独立性检验。研究使用方差膨胀因子(Variance Inflation Factor, VIF) [11]和容忍度(Tolerance, TOL)量化因子之间的共线性程度，当 $VIF > 5$ ， $TOL < 0.2$ 时，因子之间的共线性强，需要剔除该因子；当 $VIF < 5$ ， $TOL > 0.2$ 时，因子之间没有共线性。

使用皮尔逊相关系数(Pearson Correlation Coefficient, PCCs) [12]计算地质灾害易发性影响因子之间的相关系数，其数值大小反映因子之间的线性相关程度，当两个因子之间的相关系数小于-0.7 时，两个因子之间存在较强的负相关性；当两个因子之间的相关系数大于 0.7 时，两个因子之间存在较强的正相关性；相关系数接近 0，则两个因子之间几乎没有相关性。

3.2. LV 模型

LV 模型将每个灾害影响因子下各个级别或类别对灾害的影响量化，通过比较每个因子条件下的灾害发生率与地区内平均发生率，计算信息量值，反映各因子对灾害易发性的贡献。计算公式如(1)所示

$$IV(X_i, K) = \ln \frac{N_i/N}{S_i/S} \quad (1)$$

式中： $IV(X_i, K)$ 表示影响因子 X_i 对地质灾害事件 K 提供的信息量值， N_i 表示分布在评价因子 X_i 区域内地质灾害个数， N 表示研究区内总的地质灾害个数， S_i 表示评价因子 X_i 的面积， S 表示研究区总面积。

3.3. LV_RF 模型

为了综合 LV 模型和 RF 模型的优点，构建 LV_RF 耦合模型对镇雄县进行地质灾害易发性评价。

RF 模型是集成学习算法，通过构建多决策树基分类器实现预测集成，经随机特征选择与采样训练后，依据训练样本结果对数据进行预测投票。在 RF 模型构建评价模型的过程中，将 IV 模型计算获得的各因子信息量值作为输入特征值，在研究区距离地质灾害点 800 m 外随机选取和历史地质灾害点等量的非地质灾害点，与历史地质灾害点一起构成地质灾害易发性区划的样本。其中，将地质灾害点设定为“1”，非地质灾害点设定为“0”。分别随机选取地质灾害点及非地质灾害点样本中的 70% 作为训练样本，剩余的作为验证样本。运用有放回随机抽样方式从训练样本中抽取数据生成新的训练样本，同时随机选取部分特征参与模型训练，那些在抽样过程中未被抽到的数据则被定义为袋外数据，可利用这部分数据对

模型拟合性能进行初步验证。而验证样本则用于对模型最终精度进行验证，从而确保评价结果的准确性与可靠性。核心超参数设置为：决策树数量 100、随机种子 42、最大特征数为平方根、最大深度无限制、最小分裂样本 2、最小叶节点样本数 1、启用自助采样；通过多次重复迭代运行模型，直至精度、变量重要性与评价结果保持稳定，完成超参数优化与模型稳定性验证。

3.4. 精度评价

采用受试者工作特征曲线(Receiver Operating Characteristics, ROC)对模型的预测性能进行检验，以曲线下的面积(Area Under Curve, AUC)对模型进行评价。其横坐标表示假阳性率(False Positive Rate, FPR)，表示实际没有地质灾害的地方被模型错误预测为有灾害的概率；纵坐标为真阳性率(True Positive Rate, TPR)，表示实际发生地质灾害的地方被模型正确预测出来的概率，曲线越靠近左上角，曲线下面积越接近 1，说明模型的预测能力越强。计算公式

$$AUC = \sum_{i=1}^{n-1} \frac{(FPR_{i+1} - FPR_i) \times (TPR_i + TPR_{i+1})}{2} \quad (2)$$

式中， n 为地质灾害点的数量；FPR 表示被错误地判定为正例的比例，即假阳性率；TPR 表示被正确地判定为正例的比例，即真阳性率。

3.5. SHAP 可解释分析

机器学习算法虽然能快速准确地完成预测任务，但大多数都是黑盒模型，需要对其结果进行解释。SHAP 是一种解释机器学习模型预测结果的方法，通过计算特征对预测结果的贡献度，基于合作博弈理论中的 Shapley 值，来为每个特征分配一个贡献值，从而解释模型的决策过程[13]。计算公式(3)如下

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} [f(S \cup \{i\}) - f(S)] \quad (3)$$

式中， ϕ_i 表示第 i 个地质灾害易发性评价因子的贡献， N 表示所有评价因子的集合， S 表示给定预测的评价因子的子集， $f(S \cup \{i\})$ 和 $f(S)$ 分别表示有或无第 i 个评价因子的模型结果[14]。SHAP 通过加法特征归因的方法生成可解释模型，即输出模型定义为输入变量的线性和，如公式(4)所示。

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i \quad (4)$$

式中， $z' \in \{0,1\}^M$ ，当样本中含有评价因子 i 时等于 1，否则等于 0； M 是输入评价因子的数量； ϕ_0 是平均预测值； ϕ_i 是第 i 个评价因子的贡献值，即 SHAP 值，SHAP 值为正，则表示该因子是正向贡献，为负则为负向贡献。

4. 结果与分析

4.1. 地质灾害易发性影响因子筛选

在地质灾害易发性评价体系的构建中，从地质、地貌、植被以及人类活动等方面初步选取高程、断层距离、岩性、土壤类型等 16 个因子对地质灾害易发性进行评价。

在进行评价前利用 VIF 对影响因子进行多重共线性分析，结果如表 2 所示，其中，EVC、SI、Slop、SR、TRI 这 5 个因子的 $VIF > 5$ ， $TOL < 0.2$ ，存在多重共线性问题，且这 5 个因子均反映地形陡峭程度与地表破碎程度，在镇雄县高山峡谷地貌中，坡度大的区域必然伴随起伏度大、切割深、粗糙度高，物理意义高度趋同，不具备独立贡献，故剔除相应因子。

Table 2. Calculation results of factor multicollinearity
表 2. 因子多重共线性计算结果

因子	TOL	VIF
Aspect	0.968	1.033
DEM	0.512	1.952
DistF	0.939	1.065
FistR	0.923	1.084
DistRd	0.881	1.135
EVC	0.144	6.951
LIT	0.942	1.061
LULC	0.767	1.304
NDVI	0.973	1.028
SI	0.116	8.633
Slop	0.09	11.151
SPI	0.668	1.497
SR	0.127	7.902
ST	0.897	1.115
TRI	0.081	12.279
TWI	0.525	1.905

在剔除多重共线性因子后，进一步采用 PCCs 计算影响因子之间的相关性，因子之间的相关性系数都小于 0.7，表明剩余因子之间不存在强相关性。如图 3 所示。

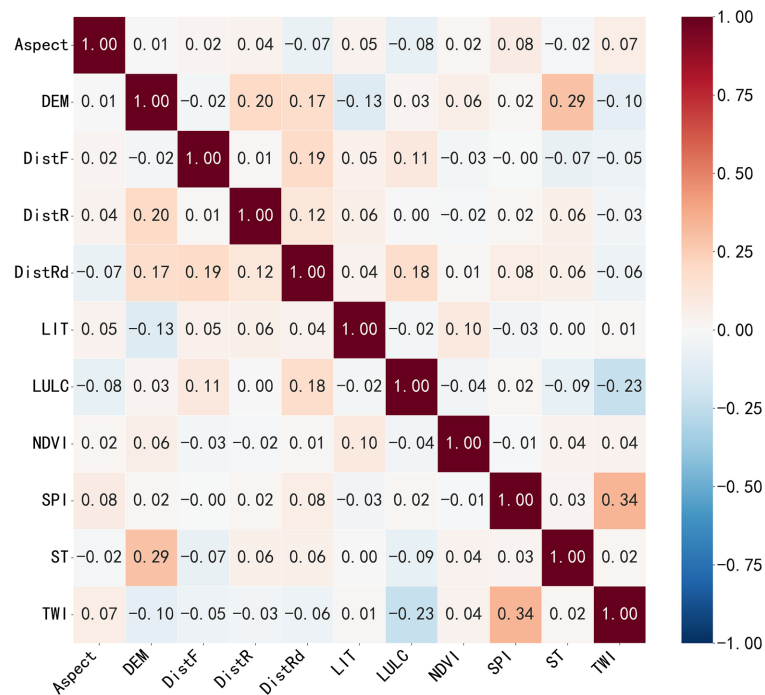
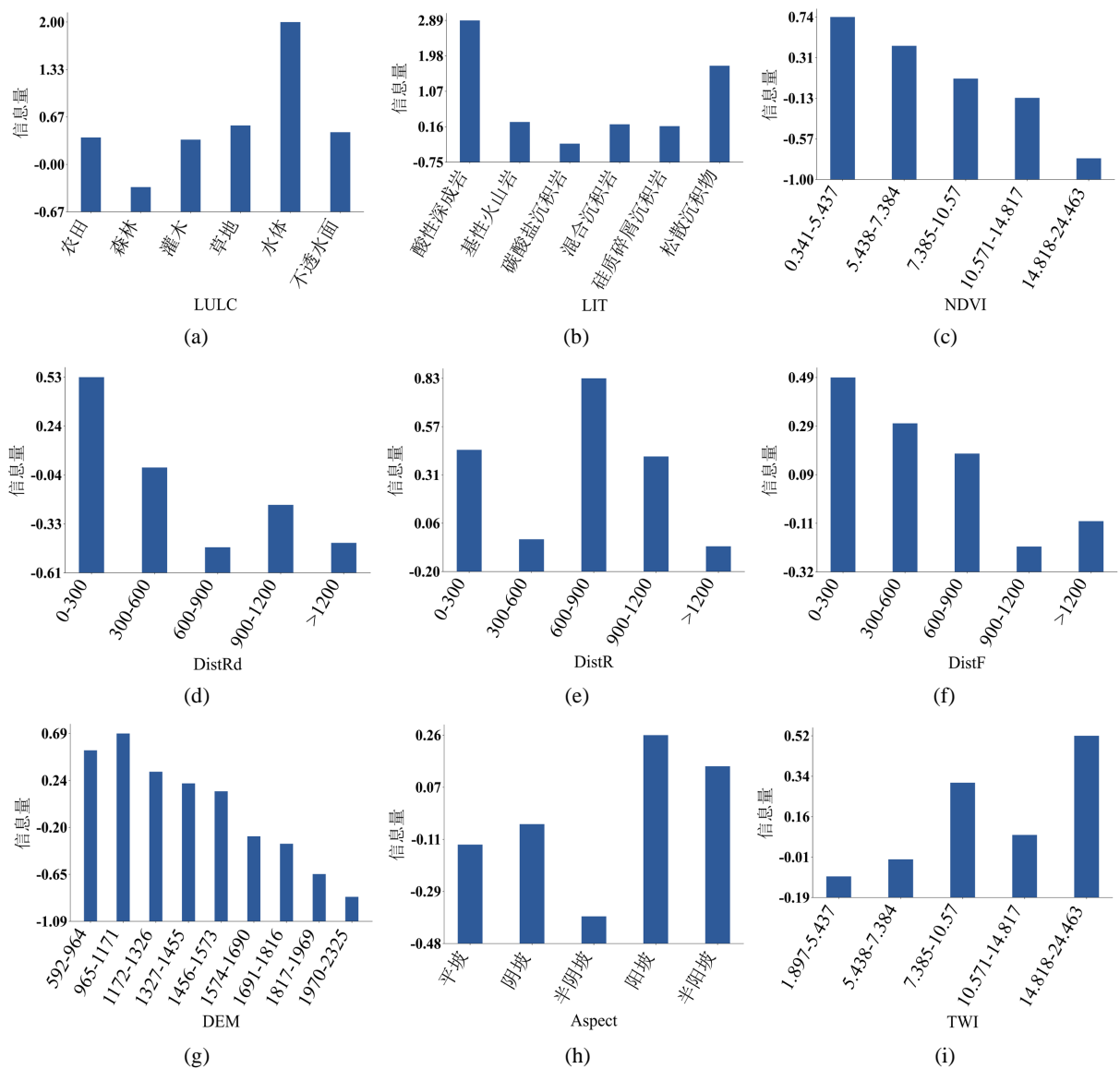


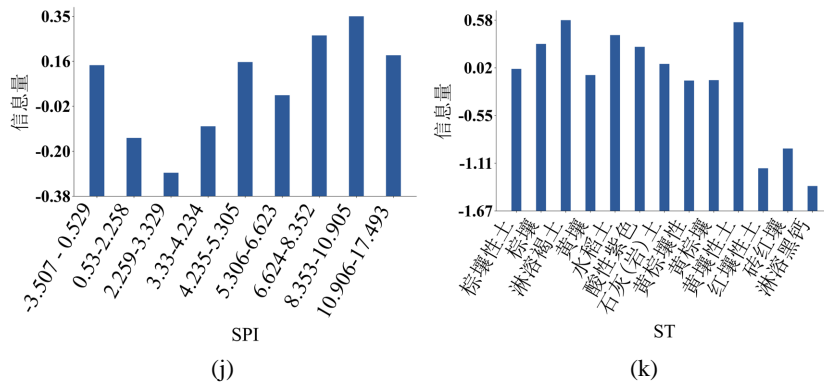
Figure 3. Results of factor correlation analysis
图 3. 因子相关性分析结果

综上，剔除 EVC、SI、Slop、SR、TRI 这 5 个因子，剩余 11 个因子之间不存在多重共线性，也不存在强相关性，构成最终评价因子体系。

4.2. 影响因子信息量计算

依据式(1)算得各评价因子二级分类信息量值(见图 4)，值越大对地质灾害影响越显著。其中，坡向因子阳坡、半阳坡信息量值为正，促进地质灾害发生，程度阳坡更甚，平坡等为负起抑制作用；DEM 因子中 965~1171 高程区间信息量值正且居首，较高及高海拔区域为负起抑制作用；DistF、DistR、DistRd 因子距断层等 0~300 范围内信息量值正且突出，随距离增加呈波动变化，体现影响复杂；LIT 因子中酸性深成岩等为正起促进作用，碳酸盐沉积岩为负起抑制作用；LULC 因子中水体等为正起促进作用，森林为负起抑制作用；NDVI 信息量随植被增加先增后减，特定区间起促进作用；SPI 特定区间起促进作用，部分区间起抑制作用；土壤类型中淋溶褐土等为正起促进作用，红壤性土等为负起抑制作用；TWI 在 14.818~24.463 之间为正且较高起促进作用，其他区间有正有负。





(a) LULC; (b) LIT; (c) NDVI; (d) DistRd; (e) DistR; (f) DistF; (g) DEM; (h) Aspect; (i) TWI; (j) SPI; (k) ST.

Figure 4. Information values of evaluation factors

图 4. 评价因子信息量值

4.3. 基于 LV_RF 模型的滑坡灾害易发性评价

将 LV 模型计算的信息量值作为输入特征值，将数据集划分为 70% 的训练数据集与 30% 的测试数据集，构建 LV_RF 模型。得到的结果使用自然断点法将结果划分为 5 个等级：极低易发区(0~0.227)、低易发区(0.227~0.435)、中易发区(0.435~0.639)、高易发区(0.639~0.843)和极高易发区(0.843~1)，结果如图 5 所示，图 6 展示了 ROC 曲线，ROC 曲线下覆盖 AUC 为 0.93，表示模型分类结果可靠，能够有效地区分滑坡区域和非滑坡区域。采用 10 折交叉验证对数据集分类准确率与评估精度 AUC 值进行统计，结果如图 7，交叉验证的 AUC 值为 0.95。极高易发区主要集中分布于县域中部、南部及东部部分区域，呈不规则片状或条带状延伸，这些区域地质灾害发生概率极高；高易发区环绕在极高易发区周边，分布范围较广，同样呈片状或条带状，与极高易发区共同构成地质灾害高敏感地带；中易发区则更为广泛地散布于全县，在极高、高易发区外围以及部分相对低缓区域均有分布，其地质灾害发生可能性处于中等水平；而极低易发区和低易发区多呈零散块状或点状，镶嵌于中易发区及高易发区之间，主要分布在地势相对平坦、地质条件相对稳定的局部区域。

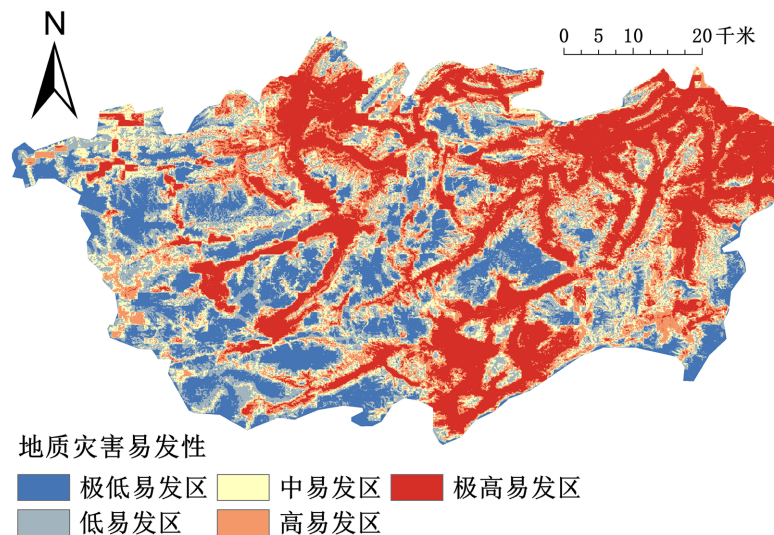


Figure 5. Geological hazard susceptibility in Zhenxiang county

图 5. 基于 LV_RF 模型的地质灾害易发性图

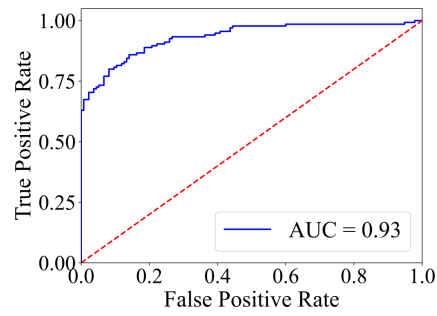


Figure 6. ROC curve of geological hazard susceptibility assessment in Zhenxiang County
图 6. 基于 LV_RF 模型的地质灾害易发性评价 ROC 曲线

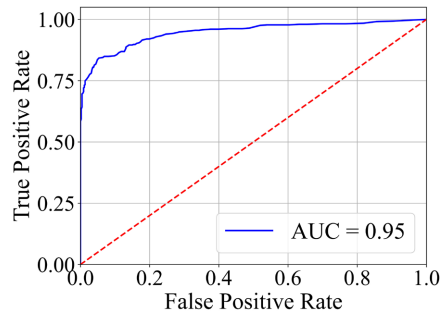


Figure 7. Cross-validation ROC curve of geological hazard susceptibility evaluation based on the LV-RF model
图 7. 基于 LV_RF 模型的地质灾害易发性评价交叉验证 ROC 曲线

4.4. 基于 RF 模型的滑坡灾害易发性评价

用同样的方法选取非滑坡点与滑坡点构成数据集，分别随机选取滑坡点及非滑坡点样本中的 70% 作为训练样本，剩余的作为验证样本。构建 RF 模型，易发性结果如图 8 所示，RF 模型得到的易发性结果与前期 LV_RF 模型的结果在整体空间格局上基本一致，极高易发区主要集中分布于县域中部、南部及东部部分区域，高易发区环绕在极高易发区周边，中易发区则更为广泛地分布，而极低易发区和低易发区主要分布在地势相对平坦、地质条件相对稳定的局部区域。采用 10 折交叉验证对数据集分类准确率与评估精度 AUC 值进行统计，结果如图 9 所示。RF 模型的 AUC 值为 0.87。通过数值可得到初步的判断，LV_RF 模型预测精度显著优于单一的 RF 模型，且 LV_RF 模型的预测效果最佳。

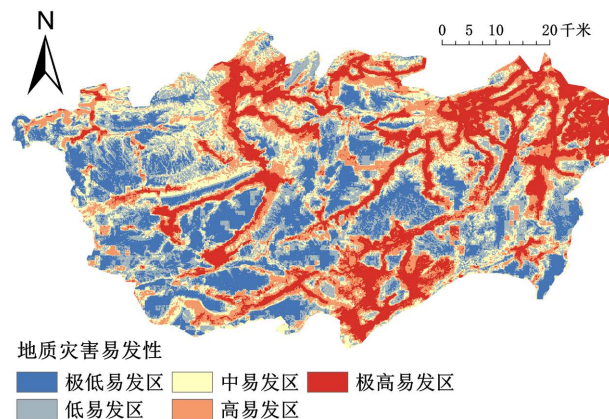


Figure 8. Geological hazard susceptibility map based on the RF model
图 8. 基于 RF 模型的地质灾害易发性图

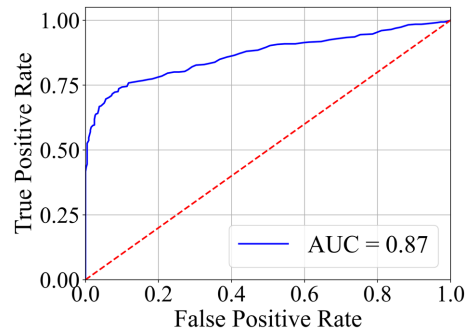


Figure 9. Cross-validation ROC curve of geological hazard susceptibility assessment based on the RF model
图 9. 基于 RF 模型的地质灾害易发性评价交叉验证 ROC 曲线

4.5. SHAP 可解释性分析

运用 SHAP 模型对计算结果展开解释，通过所绘制的特征重要性图(图 10)与 SHAP 汇总图(图 11)深入探究各影响因子与地质灾害之间的内在联系。

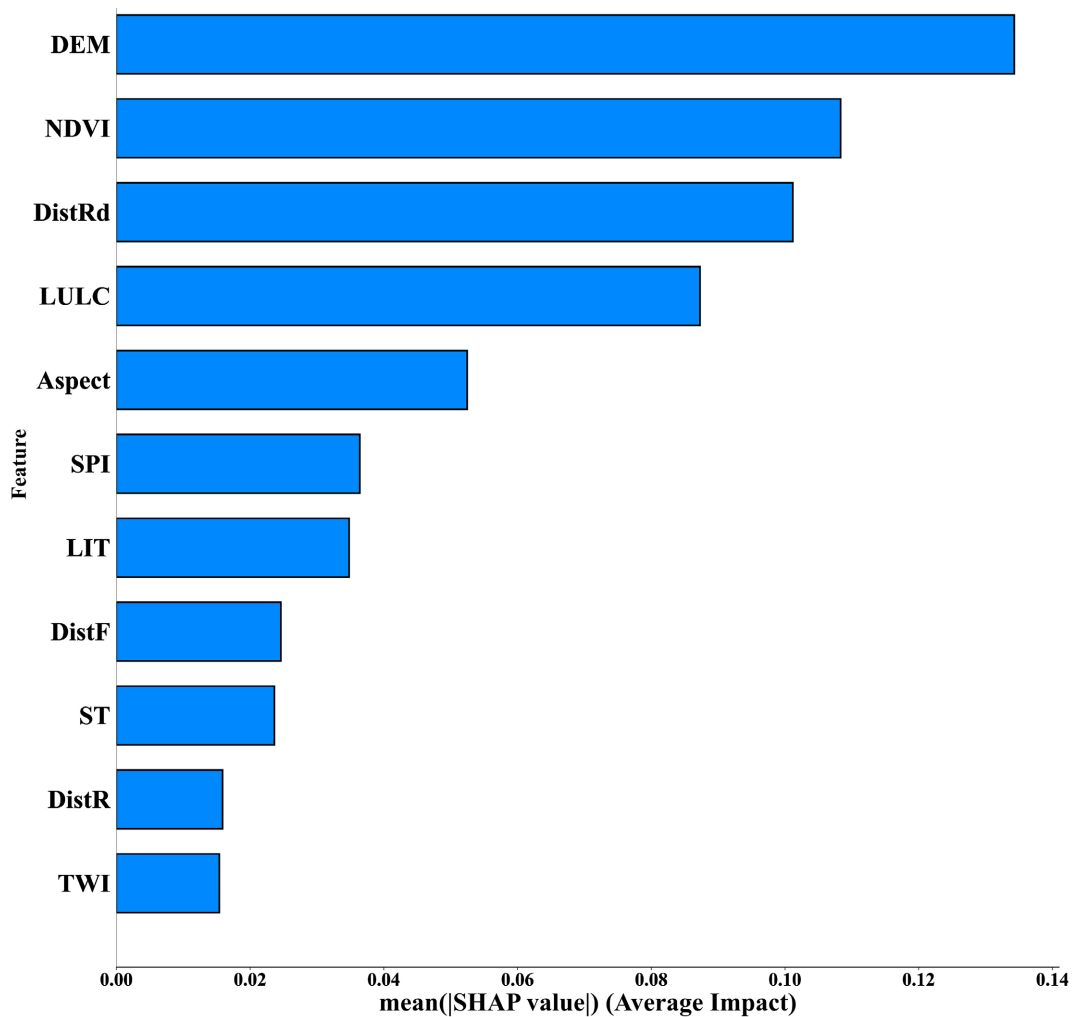


Figure 10. Ranking of feature importance
图 10. 特征重要性排序

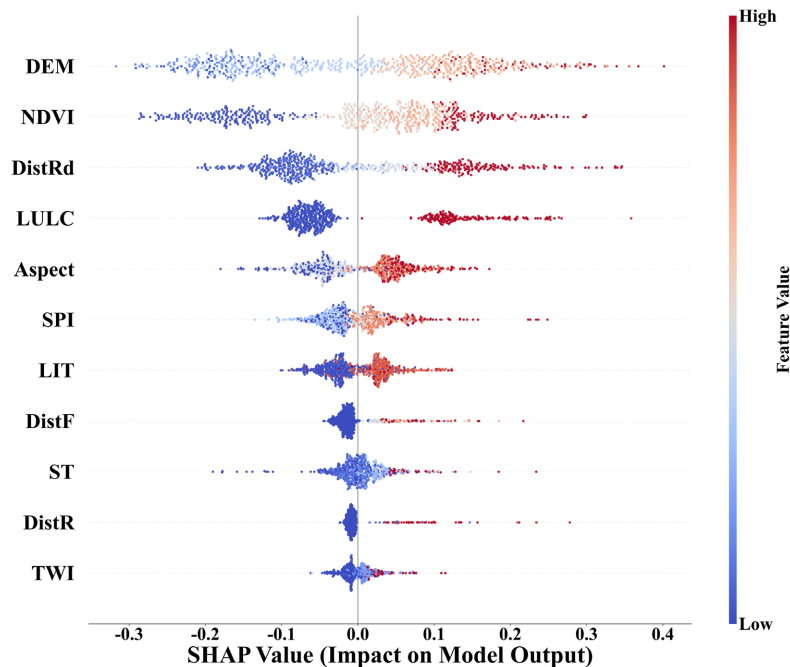


Figure 11. SHAP summary plot
图 11. SHAP 汇总图

从特征重要性图可以明确，DEM、DistRd、NDVI 等因子的信息量对地质灾害影响显著，整体表现出信息量值增大与地质灾害易发性增加相关，点云的横向延展和颜色分布反映了信息量值变化对模型输出的非线性效应和阈值特性。研究区 965~1171 m 中高程带因人类活动集中、坡体应力集中且裂隙发育，在降雨条件下极易发生失稳破坏，与区域中高程陡坡多发滑坡的地质背景一致；DistRd 则反映道路开挖形成高陡边坡、破坏岩土体结构并改变汇流条件，距离道路越近灾害风险越高，体现人类工程活动的强烈扰动效应；NDVI 呈现非线性影响特征，植被虽具固坡与水文调节作用，但高植被覆盖区多对应土层薄、风化强的坡体，中等覆盖区更易在降雨与人类活动共同作用下触发灾害。

研究结果与镇雄县地质构造复杂、地形破碎、降雨集中及人类活动强烈的区域特点高度吻合，与已有文献揭示的滇东北山区地质灾害“沿构造带、沿道路、沿河谷、沿中高程坡体”集中发育规律基本一致。同时，本研究通过 IV-RF 耦合与 SHAP 可解释分析，进一步量化了高程与道路扰动的主导贡献，明确了各因子的非线性驱动阈值，在模型精度与机理解释上更具可靠性，可为山区县域地质灾害防治提供更精细的科学依据。

5. 结论

1) IV-RF 复合模型在镇雄县地质灾害易发性评价中展现出较高的可用性和预测精度。实验结果表明，该模型的 ROC 曲线下面积(AUC)达到 0.92，显示出极强的分类能力和预测可靠性，为地质灾害易发性评价提供了科学有效的工具。

2) 基于 IV-RF 模型的地质灾害易发性评价结果显示，镇雄县地质灾害易发性呈现显著的空间差异性。极高易发区和高易发区主要集中分布在县域中部、南部及东部部分区域，这些区域地质构造复杂、地形陡峭、降水集中，加之人类工程活动强烈，导致地质环境脆弱，极易诱发地质灾害。相反，极低易发区和低易发区则多分布在地势相对平坦、地质条件相对稳定的局部区域。中易发区则广泛散布于全县，作为极高、高易发区与低易发区之间的过渡地带，其地质灾害发生可能性处于中等水平。这种空间分布

特征为区域地质灾害的针对性防治提供了重要参考。

3) SHAP 可解释性分析表明, DEM、DistRd、NDVI 对模型结果影响最重要的三个因素。因子整体表现出相似的影响特性, 即各因子的信息量值增大通常与地质灾害易发性增加相关, 点云的横向延展和颜色分布反映了信息量值变化对模型输出的非线性效应和阈值特性

参考文献

- [1] 马玲玲, 高海东, 于国强, 等. 基于耦合信息量与机器学习的陕西省地质灾害易发性评价[J]. 自然灾害学报, 2026, 35(1): 92-103.
- [2] 王宇. 云南省地质灾害防治与研究历史评述[J]. 灾害学, 2019, 34(3): 134-139.
- [3] 谭玉敏, 郭栋, 白冰心, 等. 基于信息量模型的涪陵区地质灾害易发性评价[J]. 地球信息科学学报, 2015, 17(12): 1554-1562.
- [4] 赵艳南, 牛瑞卿. 基于证据权法的滑坡危险性区划探索[J]. 地理与地理信息科学, 2010, 26(6): 19-23.
- [5] 毕结昂, 徐佩华, 宋盛渊, 等. 基于信息量-逻辑回归耦合模型的玛纳斯河流域地质灾害易发性评价[J]. 工程地质学报, 2022, 30(5): 1549-1560.
- [6] 武雪玲, 任福, 牛瑞卿, 等. 斜坡单元支持下的滑坡易发性评价支持向量机模型[J]. 武汉大学学报(信息科学版), 2013, 38(12): 1499-1503.
- [7] 米国良, 朱兴华, 陶雪文, 等. 基于随机森林模型的城市路面塌陷易发性评价——以杭州市平原沟谷区为例[J/OL]. 中国地质灾害与防治学报, 1-15. <https://link.cnki.net/urlid/11.2852.p.20250825.1534.002>, 2026-03-30.
- [8] 余飞燕, 吕蒙, 杨军. 云南省镇雄县滑坡地质灾害主要致灾因子筛选及危险性分区[J]. 地质与资源, 2015, 24(2): 152-156.
- [9] 胡波, 侯岳林, 李文, 等. 融合 InSAR 与改进 CNN-BiLSTM 的镇雄县滑坡形变分析与预测方法研究[J/OL]. 安全与环境学报, 1-12. <https://link.cnki.net/doi/10.13637/j.issn.1009-6094.2025.1290>, 2026-03-30.
- [10] 薛智文, 许冲, 付登文, 等. 2024 年 1 月 22 日云南镇雄滑坡灾害成因分析[J]. 地震研究, 2025, 48(1): 80-88.
- [11] 张雷雨, 杨毅, 梁霄. 地理加权回归模型的多重共线性诊断方法[J]. 测绘与空间地理信息, 2017, 40(10): 28-31.
- [12] Ma, J., Wang, X. and Yuan, G. (2023) Evaluation of Geological Hazard Susceptibility Based on the Regional Division Information Value Method. *ISPRS International Journal of Geo-Information*, **12**, Article 17. <https://doi.org/10.3390/ijgi12010017>
- [13] Zhang, J., Ma, X., Zhang, J., Sun, D., Zhou, X., Mi, C., *et al.* (2023) Insights into Geospatial Heterogeneity of Landslide Susceptibility Based on the SHAP-XGBoost Model. *Journal of Environmental Management*, **332**, Article 117357. <https://doi.org/10.1016/j.jenvman.2023.117357>
- [14] Lundberg, S.M., Erion, G.G. and Lee, S.I. (2018) Consistent Individualized Feature Attribution for Tree Ensembles. arXiv:1802.03888.