

DOI: 10.11766/trxb202207020361

陈琦涛, 王 添, 李占斌, 张 皎, 李 鹏, 李斌斌. 基于机器学习的紫色土坡面水沙预测模型研究[J]. 土壤学报, 2024, 61 (2): 424–433.
CHEN Qitao, WANG Tian, LI Zhanbin, ZHANG Jiao, LI Peng, Li Binbin. Research on Water and Sand Prediction Model of Purple Soil Slope
Based on Machine Learning[J]. Acta Pedologica Sinica, 2024, 61 (2): 424–433.

基于机器学习的紫色土坡面水沙预测模型研究*

陈琦涛¹, 王 添^{1, 2†}, 李占斌^{1, 2}, 张 皎¹, 李 鹏^{1, 2}, 李斌斌³

(1. 西安理工大学省部共建西北旱区生态水利国家重点实验室, 西安 710048; 2. 西安理工大学旱区生态水文与灾害防治国家林业局重点实验室, 西安 710048; 3. 水利部水土保持监测中心, 北京 100053)

摘要: 为了精准预测土壤侵蚀变化、科学合理地预防水土流失的发生, 为紫土地水土流失防治提供参考, 基于典型紫土径流小区 43 场降雨资料, 降水量、历时和 I_{30} 为特征指标, 采用 K-means 聚类算法对降雨分类; 使用随机森林 (Randomforest, RF) 算法评估各影响因子对径流深 (H)、土壤流失量 (S) 的重要性, 结合通径分析法进行重要因子筛选; 将筛选所得关键因子作为模型输入变量, H、S 为输出, 使用 BP 神经网络构建预测模型。结果表明: (1) 划分出 3 种雨型, B 雨型 (短历时、大雨强、小雨量) 是主要降雨类型, A 雨型 (长历时、中雨强、大雨量) 最剧烈。(2) 各雨型下特征因子对 H、S 的重要程度明显不同。A 雨型下 T 对 S 重要程度最高 (31%), 各因素对 H 的重要程度相对均匀; B 雨型下各因素对 S 的影响差异小, F 对 H 的重要程度最高 (29%)。在 C 雨型下, Pr 对 H、S 的重要程度最高 (33%、36%)。(3) 三种雨型下的 H、S 均受到 Pr 的显著影响, B、C 雨型下的 H、S 分别同时受 F 和 V_s 、Ph 的影响显著。(4) 利用 BP 神经网络对 H、S 的预测精度均较高, Nash-Suttcliffe 效率系数均高于 0.95, 且对 H 的预测模型精确度高于对 S 的预测模型。

关键词: 机器学习; 通径分析; 降雨; 侵蚀; 紫色土

中图分类号: S157 **文献标志码:** A

Research on Water and Sand Prediction Model of Purple Soil Slope Based on Machine Learning

CHEN Qitao¹, WANG Tian^{1, 2†}, LI Zhanbin^{1, 2}, ZHANG Jiao¹, LI Peng^{1, 2}, Li Binbin³

(1. State Key Laboratory Base of Eco-Hydraulic Engineering in Arid Area, Xi'an University of Technology, Xi'an 710048, China; 2. Key Laboratory of National Forestry Administration on Ecological Hydrology and Disaster Prevention in Arid Regions, Xi'an 710048, China; 3. Monitoring Center of Soil and Water Conservation, Ministry of Water Resources, Beijing 100053, China)

Abstract: Purple soil is one of the unique soils in China. Because of its low degree of soil development and poor soil erosion resistance, soil erosion in the purple soil area is very serious, which seriously hinders local agriculture and economic development.

【Objective】 To accurately predict soil erosion changes, we need to scientifically and rationally prevent the occurrence of soil erosion, provide a reference for the prevention and control of soil erosion on purple soil slopes, and promote research on soil and

* 国家自然科学基金项目 (U2040208、52009104) 和陕西省水利科技计划项目 (2022slkj-04) 资助 Supported by the National Natural Science Foundation of China (Nos.U2040208, 52009104); Shaanxi Provincial Water Conservancy Science and Technology Project (No.2022slkj-04)

† 通讯作者 Corresponding author, E-mail: t.wang@xaut.edu.cn

作者简介: 陈琦涛 (1998—), 男, 四川凉山人, 硕士研究生, 主要从事流域泥沙与生态水利研究。E-mail: chen7taoo@163.com

收稿日期: 2022-06-22; 收到修改稿日期: 2023-02-15; 网络首发日期 (www.cnki.net): 2023-07-19

water conservation in purple soil areas. 【Method】Based on the rainfall data of 43 typical purple soil run-off communities, rainfall, duration and I30 are used as characteristic indicators, and the K-means clustering algorithm was used to classify rainfall; the Random forest (RF) algorithm was used to evaluate the importance of each influencing factor on the depth of run-off (H) and soil loss (S), and the important factor screening was carried out in combination with the path analysis method. The key factors obtained from the screening were used as the input variables of the model, with H and S as the outputs. Also, a predictive model was constructed using BP neural network. 【Result】The results showed that: (1) The erosive rainfall in the purple mound area in the southwest can be divided into three categories. Among them, the A rain type had a long duration, medium rain intensity, heavy rain volume, the lowest frequency of occurrence, and the greatest erosion, which was an important rain type that caused slope erosion in the area; the B rain type had a short duration, heavy rain, and light rain volume, and the most frequent occurrence, which was the main rain type that caused slope erosion in the area; and the C rain type, which had a medium duration, medium rain intensity, and medium rainfall, which occurred more frequently and also made a greater contribution to slope erosion in the area. (2) The importance of characteristic factors to H and S under each rain type was significantly different. For rain type A, the rain duration (T) was the most important to S (31%), and the importance of each factor to H was relatively uniform. Under rain type B, the influence of each factor on S was small, and the rainfall erosion force (F) was the most important to H (29%). Also, the degree of importance was the highest (29%). Under the rain type C, rainfall (Pr) was the most important to H and S(33%, 36%). (3) The H and S under the three rain types were significantly affected by Pr and the H and S under the B and C rain types were most significantly affected by F, vegetation cover (Vs), and average plant height (Ph) at the same time. (4) The prediction accuracy of H and S using the BP neural network was high, the efficiency coefficient of Nash-Sutcliffe is higher than 0.95, and the accuracy of the prediction model for H was higher than that of the prediction model for S. 【Conclusion】Research on the purple soil mound area in the southwest needs to focus on preventing soil erosion caused by high-frequency heavy rainfall. In this study, the two models that used rainfall and other factors to calculate H and S had high forecasting capabilities, which provide technical support for soil erosion prediction to achieve accurate forecasting of soil erosion.

Key words: Machine learning; Path analysis; Rainfall; Erosivity; Purple soil

紫色土是我国特有土壤之一，主要分布于三峡库区和川中丘陵区^[1]，其矿质养分充足，肥力高，是当地重要的农业生产资源。紫色丘陵区坡耕地主要分布在四川盆地，其中坡耕地占耕地的 49.9%^[2]。然而由于紫色土土壤发育程度低，土壤抗蚀性差，导致该地区的水土流失情况十分严重，仅次于黄土地区，严重阻碍了当地的农业及经济发展^[3]。同时，也是影响长江中上游流域水环境的主要污染源，加剧了三峡水库水体的水环境压力^[4]。

产沙产流是评估水土流失的重要指标之一^[5]，紫色土分布区域降水充沛，降雨产生的径流携带大量土壤，加之该地区土地利用以坡耕地为主，大于 10°的坡耕地占全区面积的 70%以上^[6]，大面积的坡耕地为土壤侵蚀提供有利的侵蚀场所，导致该地区水土严重流失。由此可见，降雨和坡度是紫色土区影响土壤侵蚀产流的重要因素。

国内学者针对紫色土坡面产沙产流进行了很多研究。马星等^[7]根据降雨过程划分出的 4 种雨型，

分析其对不同坡度坡耕地产流产沙的作用；孙丽丽等^[8]运用人工模拟降雨装置，研究紫色土坡面产流产沙过程受不同雨强的影响，得出雨强和紫色土坡面产流产沙量均为极显著正相关。王玉宽等^[9]建立了紫色土坡面降雨侵蚀模型；霍云梅等^[10]研究认为，坡面在恒定的基础上，雨强增大，土壤入渗率降低，产流总量和汇流速度增大。Liu 等^[11]设计了一种壤中流观测系统，在不同雨强条件下对比坡面，产流和壤中流的产流过程进行研究，认为在暴雨和大暴雨时，表面径流和壤中流均有发生，且壤中流滞后更明显。

目前研究多集中于基于人工降雨，讨论不同植被措施和耕地措施对产流产沙的影响^[12-14]，而针对不同降雨类型、不同坡度的坡耕地水沙研究较少，且目前少有学者将机器学习与划分降雨类型和评价降雨等因子对水沙影响程度相结合。

鉴于上述，以紫色土坡地为研究对象，径流小区降雨和侵蚀观测资料，基于机器学习算法对西南

紫色土地区的侵蚀性降雨类型进行聚类划分, 计算各影响因子对径流深和土壤流失量的重要性评分。探明不同降雨特征以及坡地坡度在多种雨型下对坡地侵蚀的影响。旨在进一步揭示紫色土坡地降雨土壤侵蚀机制, 为西南紫色土区坡地水土保持工作提供科学依据。

1 材料与方法

1.1 研究区域概况

万安小流域位于四川省盐亭县林山乡境内 (105°24'E、31°15'N), 属于四川中部典型丘陵区, 地貌以丘陵和低山为主, 为嘉陵江水系四级支流区, 面积约为 12.36 km², 本区属于中亚热带湿润季风气候, 年均气温 17.3℃, 年蒸发量 604 mm, 日照时长为 1 104 h。多年平均降雨 826 mm, 夏季占 65.5%, 降雨集中并多暴雨。土壤类型主要为白垩系和侏罗系紫色砂页岩发育而成的紫色土和水稻土。缓坡台地和陡坡地面积各占流域丘陵的 1/3 和 2/3。农业生产以种植业为主, 为典型的紫色土丘陵农业区, 农作物主要包括水稻、玉米、小麦、油菜等。

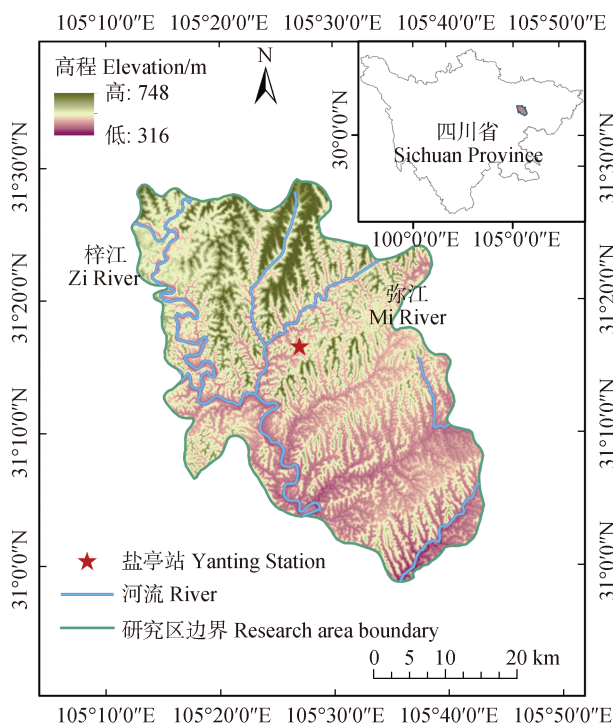


图 1 研究区地理位置图

Fig. 1 Map of the geographical location of the study area

1.2 试验设计

本研究的水文泥沙数据为四川省盐亭县万安小流域综合观测站 2015—2019 年降雨气象、水文实测资料。盐亭站设有 5 种不同坡度的径流小区 (6.5°、10°、15°、20°、25°), 每个径流小区均设置 3 个面积为 20 m×5 m 的样区, 采用冬小麦—夏玉米轮作 (表 1)。各个样区的播种、施肥方式均采用当地常规耕作方式进行, 作物种植期间不再灌溉与追肥。为确保各个径流样区水系的独立性, 四周用混凝土围埂, 以避免水流侧漏, 同时依照紫色土土层的结构特征在土层下端 60 cm 左右浇筑混凝土, 以模拟土壤, 岩石不透水界面, 回填土壤均按自然土壤结构分层填入。每个样区坡底地表处均外接有集流槽, 并通过导水管将槽中的径流导入连接着的径流池, 用以收集径流和泥沙, 各个出水口和集水池均相互独立。

1.3 研究方法

(1) 聚类算法。K-means 聚类算法属于非监督学习的硬聚类算法^[15], 是一种迭代重定位的算法^[16]。基于 R 语言采用机器学习的无监督学习聚类对 43 场降雨事件进行分类。

(2) 随机森林算法。随机森林是一种集成学习方法, 组合多个决策树算法对相同现象产生重复的预测结果^[17]。对每棵树进行自助采样, 使误差估计的计算能够基于袋外样本数据^[18]。各因子的决策树数量根据各自训练结果而定, 分割节点的数量设置为 3。采用 R 语言进行编程计算。

(3) BP 神经网络。BP 神经网络是一种典型的监督式神经网络分类器, 实现从输入到输出的非线性映射功能, 以通过学习自动提取合理的求解规则, 从而具有一定的推广和概括能力^[19]。使用 Matlab 进行 BP 神经网络模型的训练和计算。

1.4 评价指标

对实测值和模拟值进行评价, 平均绝对误差 (Mean Absolute Error, MAE)、均方误差 (Mean Squared Error, MSE)、均方根误差 (Root Mean Square Error, RMSE)。Nash-Suttcliffe 效率系数 (Ens)^[20], 通常 MAE、MSE 和 RMSE 越小、预报准确率越大, Ens 越接近 1, 表明实际值与模拟值越接近, 模型的预测能力越强。

表 1 2015—2019 年径流小区概况

Table 1 Overview of runoff plot 2015-2019

小区号 Plot number	坡度 Slope ^o	坡长 Slope length/m	坡宽 Slope width/m	土壤类型 Soil type	水土保持措施 Soil conservation measure	植株密度 Plant density/hm ²	
						小麦 Wheat	玉米 Corn
2、3	6.5	20	5	紫色土	轮作	22 770	500
4、5、6	10	20	5	紫色土	轮作	22 770	500
7、8、9	15	20	5	紫色土	轮作	22 770	500
10、11、12	20	20	5	紫色土	轮作	22 770	500
13、14、15	25	20	5	紫色土	轮作	22 770	500

2 结果与讨论

2.1 侵蚀性降雨的确定与分类

(1) 侵蚀性降雨的确定。基于通用土壤流失方程 (USLE) 及根据雨量大小拟定的侵蚀性降雨标准来划分降雨事件, 将降雨间隔时间在 6 h 以上、或连续 6 h 降雨量不足 1.2 mm 的看作二次降雨事件, 否则看作一次降雨事件^[21] (以一次降雨量 ≥ 12.7 mm 划分侵蚀性降雨)。

2015—2019 年观测侵蚀性降雨 43 次。侵蚀性降雨发生时间为 5—9 月, 月均侵蚀性降雨次数分别为 1.5 次、2.5 次、2.75 次、2 次、1.6 次。年均侵蚀性降雨为 493 mm, 占年均降雨量的 57.57% (图 2), 研究区 5—9 月降雨量及侵蚀性降雨场次比较集中, 7 月达到最大。

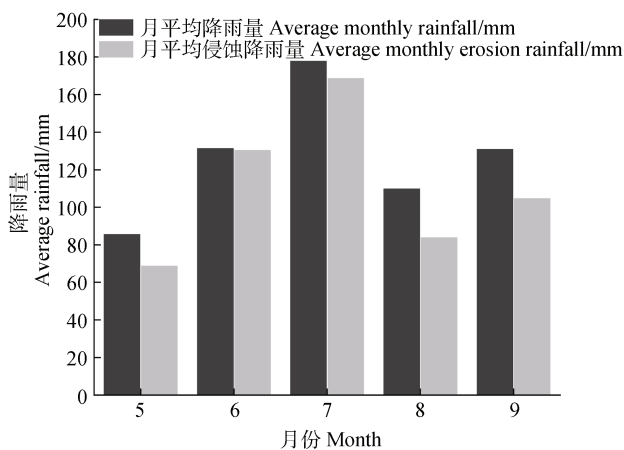


图 2 2015—2019 年降雨频发月侵蚀性降雨统计图

Fig. 2 Statistical map of erosive rainfall in 2015-2019 with frequent rainfall

(2) 侵蚀性降雨类型分类及特征。为进一步研究降雨对坡面径流泥沙的影响, 使用 K-means 算法聚类, 选取降雨量 (P)、降雨历时 (D)、最大 30 分钟降雨强度 (I_{30}) 等 3 个降雨特征指标对 43 场降雨事件进行分类, 共分为 3 类。根据 Silhouette 指标对最佳分类结果进行判断, Silhouette 指标的在 $[-1, 1]$ 范围内变动, 所有样本的平均 Silhouette 指标值越大表示聚类质量越好^[22], 其最大值对应的类数为最佳聚类数, 将样本分为 3 类的平均 Silhouette 指标值在 0.8 以上, 要明显大于其他分类。取分位数 25% 和 75% 为变化范围, 统计各种雨型降雨特征指标 (表 2)。

由表 2 可知, A 型降雨的发生频次最少, 仅发生 6 次, 占比 8.84%; B 型降雨发生频次最高, 发生 21 次, 占比 68.7%, 是研究区的主要雨型; C 型降雨发生 16 次, 占比 37.2%。A 型降雨平均降雨历时最长 (46.29 h), 平均降雨量最多 (94.38 mm), 平均最大 30 min 降雨强度最低 ($19.48 \text{ mm}\cdot\text{h}^{-1}$), 但是单场侵蚀结果为 S ($0.092 \sim 0.882 \text{ t}\cdot\text{hm}^{-2}$), H ($1.988 \sim 21.61 \text{ mm}$), 是造成土壤侵蚀最强的降雨类型; B 型降雨平均降雨历时和平均降雨量均最低, 分别为 0.65 h 和 37.94 mm, 但平均降雨强度最大 ($39.01 \text{ mm}\cdot\text{h}^{-1}$), 单场侵蚀强度 S 介于 $0.0095 \sim 0.361 \text{ t}\cdot\text{hm}^{-2}$ 之间, H 介于 $0.782 \sim 4.846 \text{ mm}$ 之间, 说明是造成土壤侵蚀的次要雨型; C 型降雨的降雨历时、降雨量均高于 B 型降雨, 其降雨强度与 A 型降雨差异较小, 但是单场侵蚀强度远低于 A 且亦低于 B, 对土壤侵蚀贡献较小。因此, A 型降雨具有发生频率低、降雨历时长、降雨量大、降雨强度低,

侵蚀结果高的特点；B 型降雨发生频率高、降雨历时短、降雨量小、降雨强度大、侵蚀结果中等；C 型降雨是发生频率较低、降雨历时较长、降雨量中等、降雨强度较低、侵蚀结果较低的降雨类

型。3 类降雨的单场侵蚀能力呈 A 型雨 > B 型雨 ≥ C 型雨。其中高频次、短历时、大雨强的 B 雨型是造成紫色土坡地土壤侵蚀的主要降雨类型，应重点防范。

表 2 不同雨型降雨特征

Table 2 Rainfall characteristics of different types of rain

降雨类型 Type of rainfall	样本个数 Samples number	特征指标 Characteristic indexes						
			$I_{30}/(\text{mm}\cdot\text{h}^{-1})$	Pr/mm	T/h	$I/(\text{mm}\cdot\text{h}^{-1})$	$S/(\text{t}\cdot\text{hm}^{-2})$	H/mm
A	6	平均值	19.48	94.38	46.30	3.10	0.55	17.65
		下限	5.60	42.50	39.42	1.20	0.09	1.98
		上限	38.4	185.8	52.12	4.70	0.88	21.61
B	21	平均值	39.01	37.94	0.65	6.90	0.24	6.51
		Q ₂₅	7.20	14.80	1.00	4.60	0.01	0.78
		Q ₇₅	89.50	102.00	11.00	8.00	0.36	4.85
C	16	平均值	24.78	63.08	26.28	1.90	0.27	2.69
		Q ₂₅	1.20	17.40	16.83	1.40	0.01	0.61
		Q ₇₅	64.9	185.8	35.67	2.80	0.08	2.59

注：Q₂₅、Q₇₅、I₃₀、Pr、T、I、S 和 H 分别表示 25%分位数、75%分位数、最大降雨 30 min 雨强、降雨量、降雨历时、平均雨强、土壤流失量和径流深。Note: Q₂₅, Q₇₅; I₃₀, Pr, T, I, S and H represent 25% quantile value, 75% quantile value, the maximum 30 min rainfall intensity, rainfall, duration, average rainfall intensity, soil loss and runoff depth.

2.2 主要影响因子筛选

(1) 降雨因子对径流深、土壤流失量影响的重要性。径流和侵蚀发生的关键驱动力为降雨特征^[23]，选取降雨量 (Pr)、坡度 (θ)、降雨历时 (T)、平均雨强 (I)、最大 30 min 降雨强度 (I₃₀)、降雨侵蚀力 (F)、雨前土壤含水量 (Sw)、植被盖度 (Vs)、植株平均高度 (Ph) 9 个影响因子，使用随机森林 (Randomforest, RF) 机器学习算法计算各影响因子对径流深 (H) 和土壤流失量 (S) 的重要性评分^[24]，其中决策树的数量根据各个坡度的训练结果来确定，各雨型下影响因子对径流深影响对应的决策树数量依次为 380、200、400，对土壤流失量对应的决策树数量依次为 500、800、400、460、500，分割节点的数量设置为 3。

在 A 雨型下， θ 对 H 的影响程度最大 (19%)，因为坡度可以通过改变径流沿坡面方向的重力分力、坡面承雨面积来影响径流深度。I₃₀、Pr (15%) 以及 Vs、F (13%) 也对 H 有较大影响，但是总体而言，各因素对 H 的影响程度均不高，Sw 对 H 的

影响程度最低 (1%)，这是因为 A 型降雨雨量大，土壤层含水量一定会达到饱和。T、I、Ph 对 S 的影响程度最大，分别为 31%、23%、19%，原因是玉米植株冠层对降雨起到缓冲作用，消减降雨动能，使其转化为能量较小的茎秆流和穿透雨。玉米为高秆作物，其较高的株型使叶缘、叶尖处下落的大雨滴重新获得了能量，会对冠下溅蚀、增加坡面径流紊动性产生积极影响^[25]。

在 B 雨型下，F、I₃₀、Pr 对 H 的影响最大 (29%、28%、20%)， θ 对 H 的影响最低 (1%)，对 S 而言，9 个因素对 S 的影响程度差异不大，原因是 B 雨型雨强大，历时短，雨量小，但是玉米植株有较强的截流能力，降雨穿透雨量低，所以雨强相关元素对 S 的重要程度降低， θ 、Pr 对其影响相对较高。在 C 雨型下，I、Pr 对 H、S 的重要程度最高，分别为 23%、33%和 22%、36%。这与张嘉懿等^[26]、向宇国等^[27]、Forster 等^[28] 对降雨特征因子与土壤侵蚀量关系的统计分析结果相似。

由此也表明不同雨型下 H、S 差异明显，不同

雨型下对 H、S 的重要影响特征因素亦有不同。

(2) 径流深、土壤流失量的关键影响因子。利用通径分析法不仅可以筛选各特征因子对径流深的关键影响因子,还能依据间接通径系数对上节得到的特征因子重要性进行补充说明。

计算各因子对 H 的通径系数时,将不显著因子剔除。A 雨型下, Pr 为主要影响因子; B 雨型下 F 为主要影响因子; C 雨型下, Ph 对 H 的直接作用最强, I_{30} 和 F 通过 Pr 以及 Ph 通过 Vs 对 H 的间接作用明显; 综合来看,对 H 的影响能力排序为 $Pr > Vs > F > Ph > I_{30}$, 其中 Vs、Ph 的决定系数为负,表明二者对 H 起到了抑制的作用。三种雨型下的 H 均受到 Pr 的显著影响, B、C 雨型同时也受到 F 的强烈影响。

A 雨型下,对 S 的直接作用由大到小为 T、F、Pr,间接作用由大到小为 F、T、Pr。 I_{30} 、Sw、Vs、Ph 对 B 雨型下的 S 影响较大,除 I_{30} 的直接通径系数大于 0 以外,其他均小于 0,说明 I_{30} 是导致坡面土壤流失的关键因素。C 雨型下,各显著影响因子对 S 的直接作用由大到小为 Vs、Ph、Pr、 I_{30} ,综合决定能力由大到小为 Ph、Pr、 I_{30} 、Vs。除开土壤流失量最大的 A 型降雨, B、C 型降雨下的 S 均受到 I_{30} 的正面影响以及 Vs、Ph 的负面影响,在 B、C 雨型下, Vs、Ph 对 S 有显著的抑制作用,其中 Vs 通过 Ph 对 S 的负作用最为明显。

这表明降雨特征因子是促进影响紫色土坡面径

流深和土壤流失的主要影响因子集,降雨是影响 H 的最关键因子,雨强是影响 S 的关键因子。

在侵蚀性最强的 A 型降雨条件下,通径分析显示玉米作物在坡地上的固土能力不足。但是在侵蚀强度较低的 B、C 雨型下,玉米作物可以有效减少坡地土壤流失,可以依据植株生长高度与种植密度的情况来判定。

2.3 BP 神经网络预测坡面水沙量

(1) 基于 BP 神经网络的坡面水沙量预测模型。使用随机森林 (Randomforest, RF) 机器学习算法和通径分析筛选所得关键因子作为模型输入变量,选择重要程度大于 10% 的因素作为关键因子, H、S 分别为输出变量,利用 BP 神经网络分别建立不同雨型下对应的 H、S 模型,试验期间共 534 组数据,其中 A 雨型共 46 组数据, B 雨型共 275 组数据, C 雨型共 213 组数据。随机选取 70% 作为训练集,验证集和检验集各为 15%。传递函数为 Tansig,迭代次数采用早停法 (Early Stopping) 通过式 (1) 来确定节点个数,对径流深的神经网络得出隐藏层个数为 4~11 个,在多次预测评估后得出,隐藏层神经元为 6 个时,模型泛化性最优。对土壤流失量的神经网络得出隐藏层个数为 2~6 个,在多次预测评估后得出,隐藏层神经元为 3 个时最适宜。

基于 BP 神经网络模型的 H、S 拟合结果评价如表 3、表 4 所示。

表 3 不同雨型下使用 BP 神经网络对坡面径流深预测模型的评价

Table 3 Evaluation of slope runoff depth prediction model using BP neural network under different rain patterns

降雨类型 Rainfall pattern	样本数 Sample number	用于检验的特征 Features for inspection	平均绝对误差 MAE	均方误差 MSE	均方根误差 RMSE	纳什效率系数 Ens
A	46	Pr、 θ 、 I_{30} F、Vs、Ph	3.69	27.7	5.27	0.58
B	275	Pr、 I_{30} 、F	0.7	1.75	1.32	0.94
C	213	Pr、I、 I_{30} F、Sw	1.21	4.25	2.06	0.99

在对 H 的预测模型下, A 雨型采用 Pr、 θ 、 I_{30} 、F、Vs、Ph 6 个因子来构建模型,但其 MAE、MSE、RMSE 为最大 (3.69、27.7、5.27),纳什效率系数最低 (0.58),说明 A 雨型下对 H 的预测模型效果不佳,原因是采用的因素特征效果并不明显,从而影响模型准确度。B 雨型使用 3 个特征因子 Pr、 I_{30} 、

F 构建模型,其 MAE、MSE、RMSE 为三者最小 (0.7、1.75、1.32),纳什效率系数为 0.94,模型效果很好; C 雨型使用 Pr、I、 I_{30} 、F、Sw 5 个因素,其 MAE、MSE、RMSE 略高于 B 雨型,远低于 A 雨型 (1.21、4.25、2.06),但 C 雨型 Ens 最高 (0.99),模型效果佳。

表 4 不同雨型下使用 BP 神经网络对坡面土壤流失量预测模型的评价

Table 4 Evaluation of BP neural network on slope soil loss prediction model under different rain patterns

降雨类型 Rainfall pattern	样本数 Sample number	用于检验的特征 Features for inspection	平均绝对误差 MAE	均方误差 MSE	均方根误差 RMSE	纳什效率系数 Ens
A	46	T、I、Ph	0.11	0.05	0.22	0.76
B	275	Pr、 θ 、 I_{30} Sw、Vs	0.15	0.06	0.25	0.92
C	213	Pr、I、 I_{30} 、F	0.11	0.08	0.29	0.87

在对 S 的预测模型下，A 雨型采用 T、I、Ph 3 个因素，B 雨型使用 Pr、 θ 、 I_{30} 、Sw、Vs 5 个因素，C 雨型使用 Pr、I、 I_{30} 、F 4 个因素，三者的 MAE、MSE、RMSE 差异不大，但 Ens 排名为 B (0.92) > C (0.87) > A (0.76)，三者的模型效果均良好，B 雨型的最佳。

总体来看，对于 3 种雨型，使用 BP 神经网络

建立的对 H、S 的预测模型拟合效果良好。

(2) BP 神经网络预测模型精度检验。对 H、S，A 雨型从 7、8、9 月每月随机抽取 3 场数据，B、C 雨型从 6、7、8、9 月随机抽取 5 场数据，以保证预测范围涵盖整个玉米植株生长期。利用 BP 神经网络训练得出的合格的预测模型来预测 H、S。通过随机选择的降雨特征数据可以观察到 (表 5，表 6)，

表 5 不同雨型下使用 BP 神经网络对坡面径流深预测模型的评价

Table 5 Evaluation of BP neural network on slope runoff depth prediction model under different rain patterns

降雨场次 Rainfall event	雨型 Rain pattern	用于检验的特征 Features for inspection							预测关系检验 Predictive Relationship Test					
		Pr /mm	θ / $^{\circ}$	I /($\text{mm}\cdot\text{h}^{-1}$)	I_{30} /($\text{mm}\cdot\text{h}^{-1}$)	F /($\text{MJ}\cdot\text{mm}\cdot\text{hm}^{-2}\cdot\text{h}^{-1}$)	Sw /%	Vs /%	Ph /m	实测 Measured /mm	预测 Predict /mm	平均绝对 误差 MAE	均方误 差 MSE	纳什效率 系数 Ens
2017/6/20	A	42.5	10		14.3	36		40	0.75	2.9	2.92	0.13	0.03	0.99
2017/6/27		85.7	20		25.4	186.58		65	0.5	11.13	10.85			
2017/9/8		131.2	6.5		26.64	643.62		90	1.6	15.37	15.28			
2017/6/15	B	54.2			35.3	407				0.7	0.78	0.06	0.01	0.99
2017/7/13		62			47.89	826.17				3.36	3.34			
2017/7/28		46			20.95	57.97				0.25	0.31			
2017/8/6		54.2			35.3	407				8	8.09			
2017/8/25		24.1			35.81	228.97				0.89	0.94			
2017/6/11	C	43.5		2.3	10.8	75.5	27.7			2	2.08	0.54	0.9	0.99
2017/6/24		33.2		1.08	18.53	53.24	18.5			0.86	0.86			
2017/7/18		185.8		9.5	64.96	3498.83	0.24			66.25	66.16			
2017/8/7		52.7		2.3	51.4	595.74	18.7			3.18	3.33			
2017/9/15		17.4		0.6	1.2	3.5	31.9			0.3	0.36			

表 6 不同雨型下使用 BP 神经网络对坡面土壤流失预测模型的评价

Table 6 Evaluation of BP neural network on slope soil loss prediction model under different rain patterns

降雨场次 Rainfall event	用于检验的特征 Features for inspection										预测关系检验 Predictive Relationship Test				
	雨型 Rain pattern	T /min	Pr /mm	θ / $^{\circ}$	I / $(\text{mm}\cdot\text{h}^{-1})$	I_{30} / $(\text{mm}\cdot\text{h}^{-1})$	F / $(\text{MJ}\cdot\text{mm}\cdot\text{hm}^{-2}\cdot\text{h}^{-1})$	Sw /%	Vs /%	Ph /m	实测 Measured	预测 Predict	平均绝对误差 MAE	均方误差 MSE	纳什效率系数 Ens
											/mm	/mm	/mm	/mm	MAE
2017/6/4	A	315			11.5					0.05	0.97	1.02	0.06	0.01	0.92
2017/6/20		3050			0.8					0.75	0.55	0.53			
2017/6/27		2500			3.2					0.5	0.43	0.53			
2017/7/31	B		30.4	25		41.64		0.24	68		0.85	0.92	0.07	0.01	0.95
2017/7/13			62	15		47.89		25.2	25		0.61	0.57			
2017/7/22			14.8	15		28.73		0.27	50		0.94	0.89			
2017/8/25			24.1	20		35.81		16.5	91		0.12	0.04			
2017/7/22			14.8	20		28.73		0.27	50		1.11	0.99			
2017/7/18	C		185.8		9.5	64.96	3498.83				1.85	1.98	0.1	0.01	0.98
2017/9/9			185.8		4.7	38.4	595.6				1.61	1.71			
2017/7/11			121.1		3.4	56.3	683.5				1.63	1.73			
2017/6/23			106.6		3.59	59.6	1411.52				0.09	0.17			
2017/8/7			52.7		2.3	51.4	595.74				0.06	0.14			

除 A 雨型中的 6 月 4 日、7 月 18 日以外, 其余降雨的特征指标均处于对应范围。因为 6 月 4 日的降雨特征属于短历时降雨, 可以解释降雨平均强度为何高于 A 型; 6 月 24 日、7 月 18 日、9 月 9 日雨强均溢出本雨型范围。总体上, 随机取出的降雨均符合对应雨型基本特征, 具有用于检验的典型性。

模型模拟效果较好, 三种雨型下对 H 的预测结果较好 (纳什效率系数均为 0.99), 三种雨型平均绝对误差 MAE 分别为 0.13、0.06、0.54, 均方误差 MSE 为 0.03、0.01、0.9。由表 6, 模型对 S 的预测模拟精度亦高, 与对 H 的预测水平相似, 纳什系数均大于 0.9, 平均绝对误差 MAE 分别为 0.06、0.07、0.1, 均方误差 MSE 均为 0.01。可以得出, 对 H、S 的预测模型效果佳, 且对 H 的预测模型精确度高于对 S 的预测模型。

土壤侵蚀受到多元影响, 相关预报模型误差通常较高。如利用 WEPP 模型对川中紫色区小流域径流深预报的 Nash-Suttcliffe 效率系数为 0.723^[29], 利

用 WEPP 模型对川东紫色土裸露坡地次降雨侵蚀强度预报的相对误差为 24%^[30], 使用 ANN-CA 模型预测黄土丘陵区土壤侵蚀, 总体精度为 87.9%^[31], 采用以径流侵蚀功率和径流深为自变量、次降雨侵蚀产沙为因变量的统计模型^[32], 其预报结果的平均相对误差均在 15%左右, 而采用分布式水蚀物理模型预报的平均相对误差则为 24%^[33]。本研究较高的 Nash-Suttcliffe 效率系数和对径流深小于 20%的预测相对误差, 对土壤流失量小于 30%的预测相对误差, 表明该 BP 模型具有较好的应用精度。需要说明的是, 该研究是基于 20 m 的定长小区实验, 最后获得的神经网络模型无法反映坡长对侵蚀的贡献。

3 结 论

(1) 西南紫色土丘陵区侵蚀性降雨可以划分为 3 类。A 雨型长历时、中雨强、大雨量、发生频次最低、侵蚀性最大, 是造成该区坡面侵蚀的重要雨

型; B 雨型短历时、大雨强、小雨量, 发生频次最多, 是造成该区坡面侵蚀的主要雨型; C 雨型中历时、中雨强、中雨量, 发生频次较多, 对该区坡面侵蚀亦有较大贡献。因此, 应重点防范高频次的强降雨导致的土壤侵蚀。(2) 各雨型下特征因子对 H、S 的重要程度明显不同。A 雨型下 T 对 S 重要程度最高(31%), 各因素对 H 的重要程度相对均匀; B 雨型下各因素对 S 的影响差异小, F 对 H 的重要程度最高(29%)。在 C 雨型下, Pr 对 H、S 的重要程度最高(33%、36%)。(3) 利用通径分析法得出: 三种雨型下的 H、S 均受到 Pr 的显著影响, B、C 雨型下的 H、S 分别同时受 F 和 Vs、Ph 的影响显著。这从另一方面证明了雨量和雨强是影响 H、S 的关键因子, 并且 Vs、Ph 对 H、S 至关重要。同时, 雨强较大时, S 还受到 Sw 的较大影响。(4) 将经过随机森林(Randomforest, RF) 机器学习算法和通径分析筛选所得关键因子作为模型输入变量, H、S 为输出, 结合 BP 神经网络构建预测模型。结果表明, 对 H、S 的预测精度均较高, 且对 H 的预测模型精确度高于对 S 的预测模型。

本研究利用降雨和其他因素计算 H、S 的两种模型均具有较高的预报能力, 为实现土壤侵蚀量的准确预报, 为水土流失预测工作提供技术支持。

参考文献 (References)

- [1] He X B, Bao Y H, Nan H W, et al. Tillage pedogenesis of purple soils in southwestern China[J]. *Journal of Mountain Science*, 2009, 6 (2): 205—210.
- [2] Zhao X J, Liu D H. Soil management and soil and water conservation of dry farming in the purple hilly area of Sichuan; 2002 Academic Annual Conference of China Association for Science and Technology, Chengdu, Sichuan, China, F, 2002 [C]. 赵燮京, 刘定辉. 四川紫色丘陵区旱作农业的土壤管理与水土保持[C]//加入 WTO 和中国科技与可持续发展——挑战与机遇、责任和对策(下册). 成都, 2002: 83—84.
- [3] Liu L, Zheng F L. Summary of the research on purple soil erosion[J]. *Journal of Anhui Agricultural Sciences*, 2006, 34 (12): 2804—2805. [刘力, 郑粉莉. 紫色土土壤侵蚀研究进展[J]. *安徽农业科学*, 2006, 34 (12): 2804—2805.]
- [4] Zhu B, Wang T, Kuang F H, et al. Measurements of nitrate leaching from a hillslope cropland in the central Sichuan Basin, China[J]. *Soil Science Society of America Journal*, 2009, 73 (4): 1419—1426.
- [5] Wang Z. Study on influencing factors of erosion and sediment yield[J]. *China Science and Technology Information*, 2008 (6): 27—29, 31. [王政. 侵蚀产沙的影响因素研究[J]. *中国科技信息*, 2008 (6): 27—29, 31.]
- [6] Gao M R, Liu G C, Zhu B. Study on process of runoff in different three tillage systems in hilly area, Sichuan Province[J]. *Journal of Soil Water Conservation*, 2000, 14 (S1): 118—121. [高美荣, 刘刚才, 朱波. 四川紫色丘陵区不同耕作制的产流过程初步分析[J]. *水土保持学报*, 2000, 14 (S1): 118—121.]
- [7] Ma X, Zheng J K, Wang W W, et al. Characteristics of the runoff and sediment yield of sloping farmland in the purple soil area under different rainfall patterns[J]. *Journal of Soil and Water Conservation*, 2017, 31 (2): 17—21. [马星, 郑江坤, 王文武, 等. 不同雨型下紫色土区坡耕地产流产沙特征[J]. *水土保持学报*, 2017, 31 (2): 17—21.]
- [8] Sun L L, Cha X, Huang S Y, et al. Effects of different rainfall intensity on the slope erosion process in purple soil[J]. *Journal of Soil and Water Conservation*, 2018, 32 (5): 18—23. [孙丽丽, 查轩, 黄少燕, 等. 不同降雨强度对紫色土坡面侵蚀过程的影响[J]. *水土保持学报*, 2018, 32 (5): 18—23.]
- [9] Wang Y K, Wang Y Q, Fu B, et al. A study on the process simulation of soil erosion by rainfall on the purple soil slopes[J]. *Journal of Mountain Science*, 2006, 24 (5): 597—600. [王玉宽, 王勇强, 傅斌, 等. 紫色土坡面降雨侵蚀试验研究[J]. *山地学报*, 2006, 24 (5): 597—600.]
- [10] Huo Y M, Bi H X, Zhu Y J, et al. Erosion process and its affecting factors of southern typical clay slope under simulated rainfall condition[J]. *Journal of Soil and Water Conservation*, 2015, 29 (4): 23—26, 84. [霍云梅, 毕华兴, 朱永杰, 等. 模拟降雨条件下南方典型粘土坡面土壤侵蚀过程及其影响因素[J]. *水土保持学报*, 2015, 29 (4): 23—26, 84.]
- [11] Liu D, She D, Yu S E, et al. Rainfall intensity and slope gradient effects on sediment losses and splash from a saline-sodic soil under coastal reclamation[J]. *Catena*, 2015, 128: 54—62.
- [12] Wang X T, Wang Y K, Fu B, et al. Experimental study on runoff characteristics of slope land in Sichuan Basin[J]. *Journal of Soil and Water Conservation*, 2006, 20 (5): 9—11, 19. [王先拓, 王玉宽, 傅斌, 等. 川中丘陵区紫色土坡耕地产流特征试验研究[J]. *水土保持学报*, 2006, 20 (5): 9—11, 19.]
- [13] Zhang B, Jiang G Y, Chen Z F, et al. Factors of soil erodibility of purplish soil in hilly region[J]. *Acta Pedologica Sinica*, 2010, 47 (2): 354—358. [张兵, 蒋光毅, 陈正发, 等. 紫色丘陵区土壤可蚀性因子研究[J]. *土壤学报*, 2010, 47 (2): 354—358.]
- [14] Liu J G, Zhang X C, Li L, et al. Research of effect of vegetation coverage on soil and water loss in purple soil

- slope land[J]. Research of Soil and Water Conservation, 2015, 22 (3): 16—20. [刘纪根, 张昕川, 李力, 等. 紫色土坡面植被覆盖度对水土流失影响研究[J]. 水土保持研究, 2015, 22 (3): 16—20, 27.]
- [15] Mondal A, Murthy K R, Ghosh A, et al. Robust image segmentation using global and local fuzzy energy based active contour[C]//2016 IEEE International Conference on Fuzzy Systems(FUZZ-IEEE). Vancouver, BC, Canada. IEEE, 2016: 1341—1348.
- [16] Qian S Y, Liu H H, Li D Y. Research and application of improved K-means algorithm in text clustering[J]. DEStech Transactions on Computer Science and Engineering, 2018, 47 (11): 161—166.
- [17] Rodriguez-Galiano V, Mendes M P, Garcia-Soldado M J, et al. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain) [J]. Science of the Total Environment, 2014, 476/477: 189—206.
- [18] Anne, Puissant, . Object-oriented mapping of urban trees using Random Forest classifiers[J]. International Journal of Applied Earth Observation and Geoinformation, 2014, 26: 235—245.
- [19] Marakhimov , Avazjon. Convergence analysis of feedforward neural networks with backpropagation[J]. Bulletin of National University of Uzbekistan (Mathematics and Natural Sciences), 2019, 2 (2): 1—1.
- [20] Nash J E, Sutcliffe J V. River flow forecasting through conceptual models part I—A discussion of principles[J]. Journal of Hydrology, 1970, 10 (3): 282—290.
- [21] Wischmeier W H, Smith D D. Rainfall energy and its relationship to soil loss[J]. Transactions , American Geophysical Union, 1958, 39 (2): 285.
- [22] Dudoit S, Fridlyand J. A prediction-based resampling method for estimating the number of clusters in a dataset[J]. Genome Biology, 2002, 3 (7): 1—21.
- [23] dos Santos J C N, de Andrade E M, Medeiros P H A, et al. Effect of rainfall characteristics on runoff and water erosion for different land uses in a tropical semiarid region[J]. Water Resources Management, 2017, 31 (1): 173—185.
- [24] Zhu Q, Guo J X, Guo X, et al. Research on influencing factors of soil erosion based on random forest algorithm—A case study in upper reaches of Ganjiang River Basin[J]. Bulletin of Soil and Water Conservation, 2020, 40 (2): 59—68. [朱青, 国佳欣, 郭熙, 等. 基于随机森林算法的土壤侵蚀影响因子研究——以赣江上游流域为例[J]. 水土保持通报, 2020, 40 (2): 59—68.]
- [25] Ma B, Li Z B, Ma F, et al. Effects of maize plants on the redistribution of water under simulated rainfall conditions[J]. Acta Ecologica Sinica, 2015, 35 (2): 497—507. [马波, 李占斌, 马璠, 等. 模拟降雨条件下玉米植株对降雨再分配过程的影响[J]. 生态学报, 2015, 35 (2): 497—507.]
- [26] Zhang J Y, Wang J H, Cheng X F, et al. Effects of rainfall on slope runoff and sediment yield under different crop measures in the Jiangzi River minor watershed[J]. Journal of Soil and Water Conservation, 2021, 35 (2): 8—14. [张嘉懿, 汪军红, 程先富, 等. 降雨对江子河小流域不同农作物措施坡面产流产沙的影响[J]. 水土保持学报, 2021, 35 (2): 8—14.]
- [27] Xiang Y G, Zhang D, Chen F, et al. Effect of rainfall and slope on runoff and sediment production of tobacco slope farmland[J]. Southwest China Journal of Agricultural Sciences, 2021, 34 (5): 1121—1127. [向宇国, 张丹, 陈凡, 等. 降雨和坡度对植烟坡耕地产流产沙的影响[J]. 西南农业学报, 2021, 34 (5): 1121—1127.]
- [28] Foster G R, Lombardi F, Moldenhauer W C. Evaluation of rainfall-runoff erosivity factors for individual storms[J]. Transactions of the ASAE, 1982, 25 (1): 124—129.
- [29] Ma H. Application of WEPP in soil and water loss in a small purple soil watershed in central SiChuan[D]. Beijing: Beijing Forestry University, 2010. [马浩. WEPP 模型在川中紫色土区小流域水土流失中的应用[D]. 北京: 北京林业大学, 2010.]
- [30] Yan D C, Wen A B, Zhang Z Q, et al. Using study of sloping WEPP model in Sichuan hilly basin[J]. Journal of Soil and Water Conservation, 2007, 21 (5): 42—45, 63. [严冬春, 文安邦, 张忠启, 等. 坡面版 WEPP 模型在川中丘陵区的应用研究[J]. 水土保持学报, 2007, 21 (5): 42—45, 63.]
- [31] Zhao J T, Ma Y X, Shi Y, et al. Prediction of soil erosion evolution in counties in the loess hilly region based on ANN-CA model[J]. Science of Soil and Water Conservation, 2021, 19 (6): 60—68.
- [32] Yu G Q, Li Z B, Lu K X, et al. Piecewise prediction model for erosion and sediment yielding of individual rainfall events in small watershed on loess plateau[J]. Acta Pedologica Sinica, 2010, 47 (4): 604—610. [于国强, 李占斌, 鲁克新, 等. 黄土高原小流域次降雨侵蚀产沙分段预报模型研究[J]. 土壤学报, 2010, 47 (4): 604—610.]
- [33] Jia Y Y, Zheng F L, Yang Q K. Distributed water erosion prediction model for small watershed in Loess Plateau[J]. Journal of Hydraulic Engineering, 2005, 36 (3): 328—332. [贾媛媛, 郑粉莉, 杨勤科. 黄土高原小流域分布式水蚀预报模型[J]. 水利学报, 2005, 36 (3): 328—332.]

(责任编辑: 檀满枝)