

RF-BP神经网络耦合模型在城市地面塌陷易发性评价中的应用

于博帆, 邢怀学, 周丽玲, 严嘉兴, 张锦瑞, 徐美君

Assessment of urban ground collapse susceptibility based on RF-BP neural network coupling model: A case study of typical areas in Hangzhou City

YU Bofan, XING Huaixue, ZHOU Liling, YAN Jiaying, ZHANG Jinrui, and XU Meijun

在线阅读 View online: <https://doi.org/10.16031/j.cnki.issn.1003-8035.202311017>

您可能感兴趣的其他文章

Articles you may be interested in

机器学习模型在滑坡易发性评价中的应用

Application of machine learning model in landslide susceptibility evaluation

刘福臻, 王灵, 肖东升 中国地质灾害与防治学报. 2021, 32(6): 98-106

基于RBF神经网络信息量耦合模型的滑坡易发性评价

Landslide susceptibility assessment by the coupling method of RBF neural network and information value: A case study in Min Xian, Gansu Province

黄立鑫, 郝君明, 李旺平, 周兆叶, 贾佩钱 中国地质灾害与防治学报. 2021, 32(6): 116-126

基于SMOTE-Tomek和CNN耦合的滑坡易发性评价模型及其应用

Landslide susceptibility mapping model based on a coupled model of SMOTE-Tomek and CNN and its application: A case study in the Zigui-Badong section of the Three Gorges Reservoir area

于宪煜, 汤礼 中国地质灾害与防治学报. 2024, 35(3): 141-151

基于机器学习的伊犁河谷黄土区泥石流易发性评估

Susceptibility assessment of debris flow disaster based on machine learning models in the loess area along Yili Valley

李志, 陈宁生, 侯儒宁, 吴铭洋, 张瀛玉龙, 杜鹏 中国地质灾害与防治学报. 2024, 35(3): 129-140

基于深度神经网络模型的雅安市滑坡易发性评价

Evaluation of landslide susceptibility in Ya'an City based on depth neural network model

牟家琦, 庄建琦, 王世宝, 孔嘉旭, 杜晨辉 中国地质灾害与防治学报. 2023, 34(3): 157-168

基于原型网络的云南怒江州泥石流灾害易发性评价与区划

A case study on the susceptibility assessment of debris flows disasters based on prototype network in Nujiang Prefecture, Yunnan Province

韩俊, 王保云 中国地质灾害与防治学报. 2023, 34(5): 117-129



关注微信公众号, 获得更多资讯信息

DOI: 10.16031/j.cnki.issn.1003-8035.202311017

于博帆, 邢怀学, 周丽玲, 等. RF-BP 神经网络耦合模型在城市地面塌陷易发性评价中的应用——以杭州市典型区为例[J]. 中国地质灾害与防治学报, 2025, 36(3): 160-170.

YU Bofan, XING Huaixue, ZHOU Liling, et al. Assessment of urban ground collapse susceptibility based on RF-BP neural network coupling model: A case study of typical areas in Hangzhou City[J]. The Chinese Journal of Geological Hazard and Control, 2025, 36(3): 160-170.

RF-BP 神经网络耦合模型在城市地面塌陷 易发性评价中的应用 ——以杭州市典型区为例

于博帆^{1,2}, 邢怀学^{1,3}, 周丽玲^{4,5}, 严嘉兴², 张锦瑞², 徐美君⁶

(1. 中国地质调查局南京地质调查中心, 江苏南京 210016; 2. 中国地质大学(武汉)地质调查
研究院, 湖北武汉 430074; 3. 自然资源部城市地下空间探测评价技术创新中心, 江苏南京
210016; 4. 浙江省地质院, 浙江杭州 310007; 5. 自然资源部浙江沿海城市地质安全
野外科学观测研究站, 浙江杭州 310007; 6. 自然资源部滨海城市地下空间地质安全
重点实验室, 山东青岛 266101)

摘要: 为了改变地面塌陷易发性评价主要通过知识驱动模型实现的现状, 文章探讨了将数据驱动模型引入城市地面塌陷评价的可能性, 选取杭州市填土-粉砂土典型区域为研究区, 进行了研究区地面塌陷指标因子的选择以及相关性检验, 筛选出了排水管线密度、社会活动密度、地下承压水位埋深、表层填土层厚度、与暗河暗浜距离、饱和砂土顶板埋深、软土层厚度 7 个评价因子对研究区地面塌陷易发性进行了评价, 通过对比 RF、I-RF 集成模型、RF-BP 神经网络模型, 得到了在该研究区背景下集成模型相对单模型对地面塌陷易发性评价结果精确度更高, 最后选取了效果最好的 RF-BP 神经网络集成模型进行了易发性评价。评价结果显示: 易发性分区与地面塌陷隐患区高度吻合, 预测效果较好, 证明了数据驱动模型在城市地面塌陷易发性评价方面应用的可能性。

关键词: 地面塌陷灾害; 易发性评价; 机器学习模型; 评价因子选取

中图分类号: P694 **文献标志码:** A **文章编号:** 1003-8035(2025)03-0160-11

Assessment of urban ground collapse susceptibility based on RF-BP neural network coupling model: A case study of typical areas in Hangzhou City

YU Bofan^{1,2}, XING Huaixue^{1,3}, ZHOU Liling^{4,5}, YAN Jiaying², ZHANG Jinrui², XU Meijun⁶

(1. Nanjing Center, China Geological Survey, Nanjing, Jiangsu 210016, China; 2. The Institute of Geological Survey of
China University of Geosciences, Wuhan, Hubei 430074, China; 3. Ministry of Natural Resources Engineering Innovation
Center for Urban Underground Space Exploration and Evaluation, Nanjing, Jiangsu 210016, China; 4. Zhejiang Institute of

收稿日期: 2023-11-20; 修订日期: 2024-04-02 投稿网址: <https://www.zgdzzyfzxb.com/>

基金项目: 自然资源部滨海城市地下空间地质安全重点实验室开放基金项目(BHKF2022Z02)

第一作者: 于博帆(2000—), 男, 湖北武汉人, 地质工程专业, 硕士研究生, 主要研究安全生产管理与城市地质评价。

E-mail: 1378747279@qq.com

通讯作者: 邢怀学(1981—), 男, 山东青岛人, 城市地质专业, 硕士, 正高级工程师, 主要从事城市地质、环境地质调查研究工作。

E-mail: 57670204@qq.com

Geosciences, Hangzhou, Zhejiang 310007, China; 5. Observation and Research Station of Zhejiang Coastal Urban Geological Security, Ministry of Natural Resources, Hangzhou, Zhejiang 310007, China; 6. Laboratory of Geological Safety of Underground Space in Coastal Cities, Ministry of Natural Resources, Qingdao, Shandong 266101, China)

Abstract: To improve the current situation where ground subsidence susceptibility assessment mainly relies on knowledge-driven models, this study explores the feasibility of incorporating data-driven models into the evaluation of urban ground subsidence. The study focused on a typical area in Hangzhou characterized by fill and silty soil. The selection of ground collapse indicators was conducted, followed by a correlation test. 7 evaluation factors, including drainage pipeline density, social activity density, depth of underground confined water level, thickness of surface fill layer, distance from hidden rivers and beaches, depth of the saturated sand top plate, and thickness of the soft soil layer, were selected for assessing the susceptibility to ground subsidence in the study area. By comparing the random forest (RF) model, RF-I integrated model, and RF-BP neural network integrated model, it was found that the integrated model had higher accuracy in assessing the susceptibility of ground collapses subsidence in this study area compared to single models. Ultimately, the RF-BP neural network integrated model, which showed the best performance, was chosen for susceptibility assessment. The assessment results indicated a high correlation between the susceptibility zones and areas prone to ground subsidence, indicating good prediction performance and proving the potential application of data-driven models in evaluating the susceptibility of urban ground collapses.

Keywords: ground collapse disaster; susceptibility assessment; machine learning model; selection of evaluation factors

0 引言

地面塌陷是我国城市中常见的地质灾害^[1]。随着城市地下建设的不断推进,人类活动对地层的扰动与对地下结构的破坏日益增多,地面塌陷灾害的发生也日趋频繁,对人民的生命及财产安全造成较大威胁^[2-3]。因此,评估与预测地面塌陷尤其是对城市等人口密集区域,并进行预防显得十分重要。

在地面塌陷的评估与预测工作中,地面塌陷的易发性评价是十分重要的一环,地面塌陷的易发性评价是指地面塌陷以地质环境条件为基础,结合孕灾因子和诱发因子预测其在一定区域内发生的可能性大小。目前对地面塌陷易发性的评价方式主要是以知识驱动模型为主,其中层次分析法运用最多,通过对已经筛选好的评价指标进行人为主观打分,得到不同评价指标的权重,最后进行易发性分析。于翠翠等^[4]利用传统 AHP 法从岩溶覆盖、水文地质条件等方面对鲁中泰安东地区某拟建高速铁路的岩溶塌陷可能性进行了分析,结合 GIS 软件对铁路建设提供了建议;Xie 等^[5]为解决泉州某高速岩溶塌陷问题,通过现场识别结合 AHP 法,最后利用 GIS 总结出该地区岩溶塌陷易发性最终结果;庄树裕^[6]针对层次分析法的判断矩阵存在一致性检验问题的现状,利用模糊层次分析法(FAHP)对福建西南部永定区的岩溶地面塌陷易发区进行了评价,使计算过程更

加简单且计算结果更加准确;江思义等^[7]利用专家-层次分析法,即先让专家根据研究区实际情况确定评价因子,后再利用 AHP 法确定评价因子权重,比起 AHP 法,通过专家筛选的方式可以更加贴合研究区的需求与特征;王桂林等^[8]利用地理探测器,对重庆中梁山地区的潜在因子进行了因子探测,结合研究区实际情况,对评价因子进行定量化分析,在得到适当的因子后,结合 AHP 法进行完整评价。

然而,随着科技的进步和地质灾害条件的复杂性增加,层次分析法的局限性日渐显现^[9]。其主要问题在于过度依赖专家对因素权重的主观判断,当涉及评价因子过多时,可能影响评价的准确性。为此,研究者开始探索基于数据的驱动模型,这在滑坡等地质灾害易发性评价中已得到广泛应用^[10-13]。不同于传统模型,数据驱动模型能够处理更复杂的数据集,从而在一定程度上提高预测的准确性。近年来,一些学者发现,通过耦合多种机器学习模型,可以进一步优化预测结果^[14-16],如耦合 BP 神经网络与 SVM 模型便可以实现 SVM 模型高泛化能力与 BP 神经网络模型深度学习的特点,这类耦合模型的方法利用了各个模型的优势,减少了单一模型可能存在的缺陷,从而在多个案例研究中显示出更高的准确度。

虽然数据驱动模型在滑坡、崩塌等灾害中运用广泛,但是受制于灾害点范围过小、评价指标与人类活动

而非自然因子相关性较大等特点,数据驱动模型在地面塌陷尤其是城市地面塌陷易发性评价中应用案例较少;冉星星^[14]为了弥补地面塌陷点塌陷面积与深圳市总面积差距太小可能导致的机器学习训练效果不佳的问题,扩大了考虑的年份,选取了 2013—2021 年的所有塌陷事故共 2 139 起,在样本数量足够的基础上进行机器学习,通过运用随机森林、XGBoost 和 SVM 三种经典单体机器学习模型对深圳市地面塌陷灾害风险进行了评价;王子童^[15]则在考虑矿区岩溶塌陷时考虑更大的影响范围并将研究区周边区域也确定为风险区并在此基础上进行评价,最后通过 RF 模型评估计算,最后利用三维可视化软件将评价结果直观地展示出来。

本文针对地面塌陷尤其是城市地面塌陷中易发性评价主要是通过知识驱动模型进行操作的现状,通过精确化研究区以及塌陷点的实体勾绘,本次样本点包括研究区发生过的地面塌陷灾害点以及经雷达探测演化为地面塌陷可能性较高的地下空洞点。在确认研究样本后,根据《T/CAGHP 018—2016 地质灾害分类分级标准》以及研究区现场勘察资料,确定了地面塌陷点的塌陷深度、地下空洞点的大小以及样本点影响范围半径为塌陷深度的 10 倍,并在此基础上划分了地面塌陷隐患区。保证了研究区内灾害点的空间占比以及研究的准确性,同时也存在较高的合理性。进一步验证了滑坡等其他地质灾害易发性评价中常见的数据驱动模型(机器学习)应用于城市地面塌陷评价中的可能性,对比了随机森林模型(random forest, RF)、RF-I(information, I)模型、RF-BP 神经网络模型的评价效果并得出了最终评价结果。

1 研究区概况

杭州市粉砂土主要分布于平原区,并且由于地形地貌特质特征,平原区堆积冲积海积平原为最典型区域,综合受灾情况以及地面塌陷隐患点,本文选取了钱塘江南岸滨江区与萧山区填土-粉砂土典型区域为研究区,面积 36.88 km²,其范围内包含地铁 4 号、5 号、6 号线等城市重要地铁线以及风情快速路、时代高架路等城市主干道,是杭州市人类活动改造较为明显的一片区域。

研究区位于萧绍平原西沿,北靠杭州复向斜丘陵,南临浦阳江平原。境内以平原为主,有少量低丘,地貌类型单调,界线清晰,地势低平,河网密布,主要土体成分为砂质粉土、粉砂,砂砾层埋深近 35~50 m。人类活动影响深刻,高层建筑物采用砂砾石层或基岩作为基础持力层。沉积物方面主要为灰色-深灰色含粉砂淤泥质黏

土以及淤泥质粉质黏土,具有明显的水平层理,含水量高。

研究区第四纪地层主要为全新统下段地层,形成于富阳海侵早期,以洪冲积为主,常表现为河谷平原、河流阶地等地貌。地层由砂、砂砾石组成,分选性和磨圆度较好,结构相对松散,厚度 2~12 m。经选取研究区内的钻孔并根据实际采样结果绘制研究区东西向与南北向剖面各一条(图 1)。

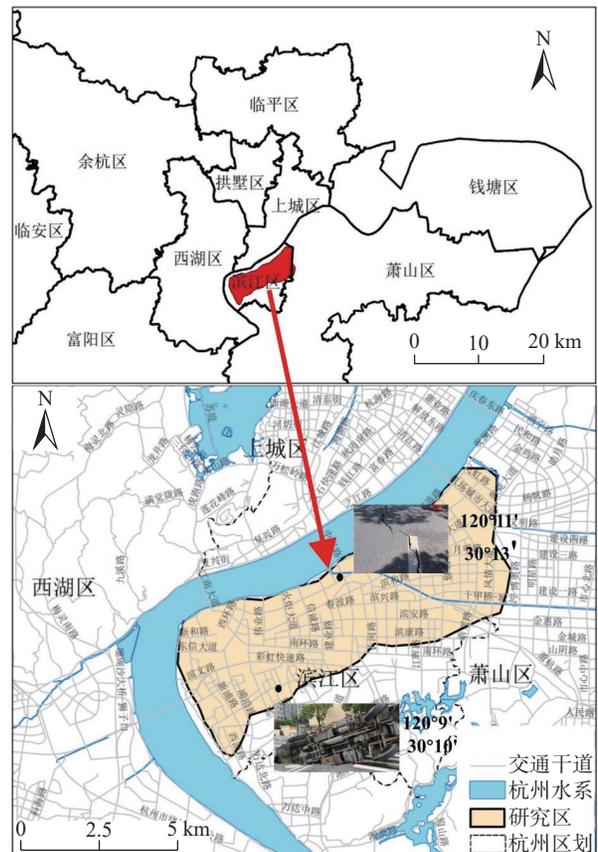


图 1 研究区地理位置图

Fig. 1 Geographical location map of the study area

研究区含水层近地表分布,一般水位埋藏于 3 m 以浅,厚度 1.5~28.8 m,底板埋深一般在 10.8~38.5 m,主要为全新统冲海积、冲湖积含水层。含水层组底板为全新统下组海积软土层,岩性主要为第四系全新统黏质粉土、砂质粉土、粉质黏土、粉砂潜水含水层,透水性相对较差;深部孔隙承压水,在天然条件下,因其水力坡度极其平缓,地下径流极其缓慢,处于相对“静止”状态,水循环交替作用几乎停止。地下水补给、径流、排泄条件较差。仅隔水顶、底板缺失时,才与上部潜水和下部承压水发生一定的水力联系。研究区地层富水性较好,潜水单井涌水量普遍在 1 000 m³/d 以上,承压水单井涌水量普遍在 3 000 m³/d 以上(图 2)。

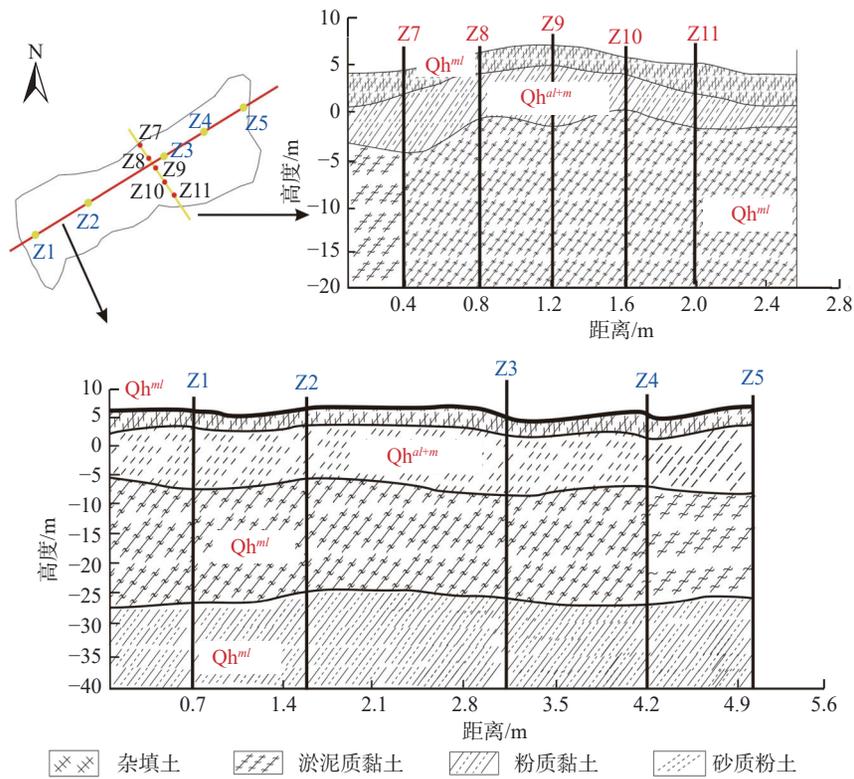


图 2 研究区典型第四纪地质剖面图

Fig. 2 Typical quaternary geological cross-sectional profile map of the study area

杭州平原原始地面标高较低, 河网密布、地表水系发育, 在城市建设过程中, 大量地表水体被回填形成了众多的暗河暗滨, 城区地平面标高的不断提升形成了大面积分布的填土层, 其物理力学性质差, 且极不均匀, 暗河暗滨地段及填土层中的大块体间, 极易形成渗流管道, 造成土体流失, 由于水土流失发生在路面以下, 从隐患形成到塌陷发生的过程一般目视难以发现, 一旦出现路面开裂或显著变形等地面迹象, 可预警时间已极短, 并且路面刚度越大, 塌陷发生的规模会越大。塌陷发生还具有滞后性, 除因高压供水管道破损引发的塌陷发展速度

快以外, 多见的是因排水管道发生破损漏水、漏土或地下水渗流等引发的水土流失, 从形成路基疏松、脱空, 到发生路面塌陷的整个过程需经历较长的时段, 具有明显的滞后性, 因管线破损引起的地面塌陷示意图见图 3。研究区地面塌陷深度基本在填土范围内或填土层与下覆地层交界 2 m 范围内, 填土厚度一般 1~5 m, 塌陷深度一般 1~2 m。发生塌陷的土体性质为上部填土下部粉土或上部填土下部粉质黏土。经实地勘察, 大部分塌陷病害点分布与粉砂、粉土层分布相关性较大且明显分布于人工填土层区域(表 1)。

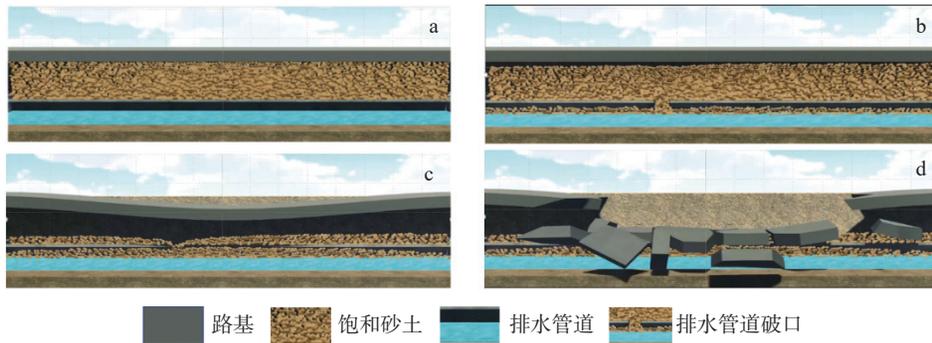


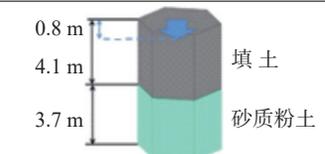
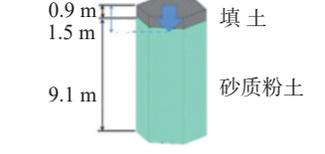
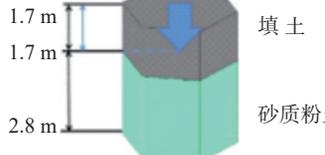
图 3 饱和砂土区管道破损引发地面塌陷过程

Fig. 3 The collapse process caused by pipeline damage in saturated sand area

注: a 为原始地层及管道结构; b 为管道破损吸入饱和砂土层; c 为地下形成空洞地面开始变形; d 为地面塌陷形成。

表 1 研究区部分典型塌陷点垂向分布特征

Table 1 Table of vertical distribution characteristics of some typical collapse points in the study area

位置	塌陷深度/m	填土厚度/m	5~10 m内岩性组合	塌陷示意图
滨江区闻涛路江陵路路口附近塌陷	0.8	4.1	填土-砂质粉土	
滨江区西兴路和丹枫路交叉口塌陷	1.5	0.9	填土-砂质粉土	
滨江区通和路空洞隐患	1.7	1.7	填土-砂质粉土	

2 研究模型

2.1 I-RF

RF 是一种基于决策树的集成学习技术,其通过融合多个决策树来形成一个鲁棒性更为良好的模型,其主要特点包括随机采样和随机特征选择。该模型利用随机采样策略和特征选择,在维持高精确度的同时,提高了模型的泛化能力。

I 法为评估地质灾害易发性提供了一种可靠的框架,该模型通过分析各影响因素与地质灾害之间的相关性,并利用信息量值计算来量化每个因素的重要性,从而评估地区的灾害易发性,以下公式为信息量值计算公式。

$$I_v = \ln(P_i/P_n) \quad (1)$$

式中: I_v ——信息量值,每个因素信息量值的核心是信息量模型的核心;

P_i ——在发生地质灾害的条件下特定因素的概率;

P_n ——在未发生地质灾害的条件下特定因素的概率。

总信息量值越高,表示地质灾害易发性越高,因为各因素的贡献更大。反之,总信息量值越低,则地质灾害易发性越低。

I-RF 是信息量与随机森林模型的耦合,模型同时接受了 RF 和 I 模型的预测结果,该耦合模型继承了两个独立模型的优点,其核心在于投票系统,通过一个统一的投票机制对两者的预测结果进行整合,该系统基于 RF 和 I 模型的输出来制定最终决策(图 4)。通过比较

各自模型的预测优势,投票系统选择最可能的结果,从而确保了一个更可靠和稳健的分类决策。模型综合了来自信息量法的关键影响因素分析和随机森林的高精度预测能力。此耦合模型不仅系统地减少了过拟合的风险,还增强了数据的可解释性和结果的多样性。

2.2 RF-BP 模型

BP 神经网络(backpropagation neural network, BP)在机器学习和深度学习领域占有重要地位,特别是在处理复杂的非线性问题方面表现出色。该网络是一种多层前馈神经网络,其独特之处在于使用了反向传播算法,允许系统通过后向调整各层的权重和偏差,不断优化网络性能,以实现预测误差的最小化。这一特性使 BP 神经网络在需要精细调整和优化的场景中特别有效。

RF-BP 模型是 RF 模型与 BP 神经网络模型的耦合,在该框架中,不同的模型被分层处理:首先,随机森林在第一层中生成初步预测,这些预测输出随后被用作第二层神经网络的输入。这种方法的优势在于,随机森林的强大能力在处理大量输入变量并捕获其非线性关系时,可以作为一个有效的特征转换器和降维工具。此外, BP 神经网络则可以在细粒度层面进一步优化和精细调整预测。此集成方法的实施采用了 KFold 交叉验证策略,极大地提高了模型的准确性和鲁棒性。通过这种方式,模型在各个不同的数据子集上进行了训练和验证,确保了评估的全面性和避免了过拟合,从而实现了测试集的高精度预测。这不仅提高了模型的泛化能力,还增强了其在面对未知数据时的预测可靠性(图 5)。

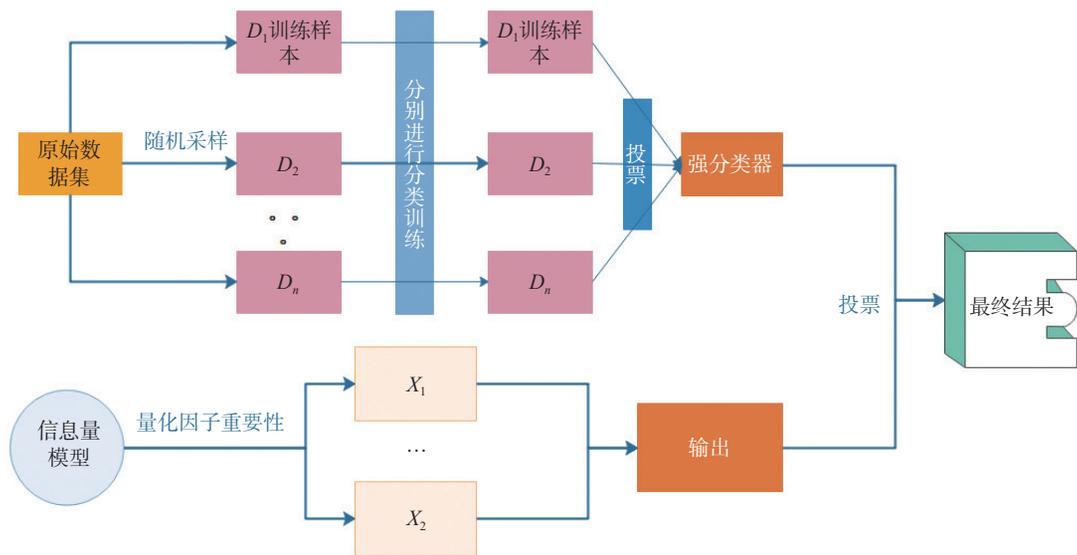


图 4 信息量-随机森林集成模型结构图

Fig. 4 Structure diagram of RF-BP neural network coupled model

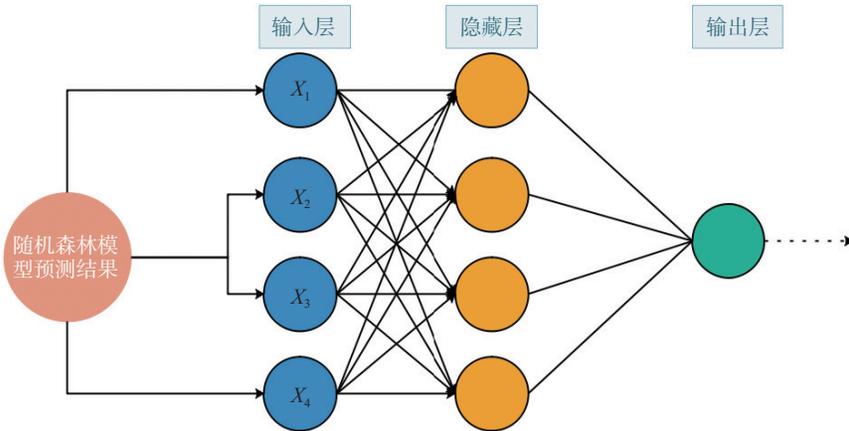


图 5 随机森林-BP 神经网络耦合模型结构图

Fig. 5 Structure diagram of random forest-BP neural network coupled model

3 数据来源与评价因子确定

3.1 数据来源

本次研究的基础数据来自项目调查与研究成果、杭州市经济、人口数据等, 具体包括①研究区地质、水文、人类活动数据, 主要用于选定评价因子与后续易发性评价工作。②近年间研究区地面塌陷的点位、范围与致灾原因③野外调查资料, 如钻孔数据, 塌陷点实际勘察数据。

3.2 评价因子的选取

城市地面塌陷的形成机制与滑坡、崩塌等其他地质灾害相似, 通常是在多种不利地质因素共同或相互作用下触发。其中, 孕灾环境因素涵盖了表层土壤和地下土层的物理特性、地下水动态等自然条件, 而致灾因素

则更多与人类活动及其他特殊触发条件相关。鉴于城市地质环境的复杂性和地面塌陷灾害的特殊性, 选择评价因子时必须综合考虑研究区的人类活动强度, 并结合地质综合条件和人文特征进行深入分析。引发研究区路面塌陷一般需有 3 个要素: 一是路基下有粉土、砂土、暗浜土和填土等易流失的土体分布; 二是有管道渗漏和地下水渗流等承载土体流失的介质-水; 三是有水土流失的通道-破损或废弃的排水管道、地下工程、河道等。在充分考察和综合现有研究资料的基础上, 本文确定了导致研究区地面塌陷的主要因素为: 填土、粉土(砂)、暗浜等不良地质广泛分布、地下排水管道结构存在缺陷、暗浜等地基回填不密实、水文地质条件的影响、人类工程活动扰动。综合以上主要因素, 最终确定本次评价研究的地质指标为本文选取以下 7 个关键因

子作为城市地面塌陷的初步评价指标。

(1)社会活动密度。社会活动密度是一个综合指标,该指标结合了人口密度和社会经济活动的密集程度,更全面地反映了人类活动的强度。通过分析社会活动密度,能够更深入地了解人类活动对区域环境的潜在

影响。通常情况下,社会活动密度较高表示人类活动更加频繁,其往往与环境压力增大,地面塌陷等自然灾害的风险增高成正相关。本研究收集并分析了研究区内各乡镇及街道的社会活动数据,合理地根据其强度将社会活动密度划分为 4 个等级,具体情况如图 6(a)所示。

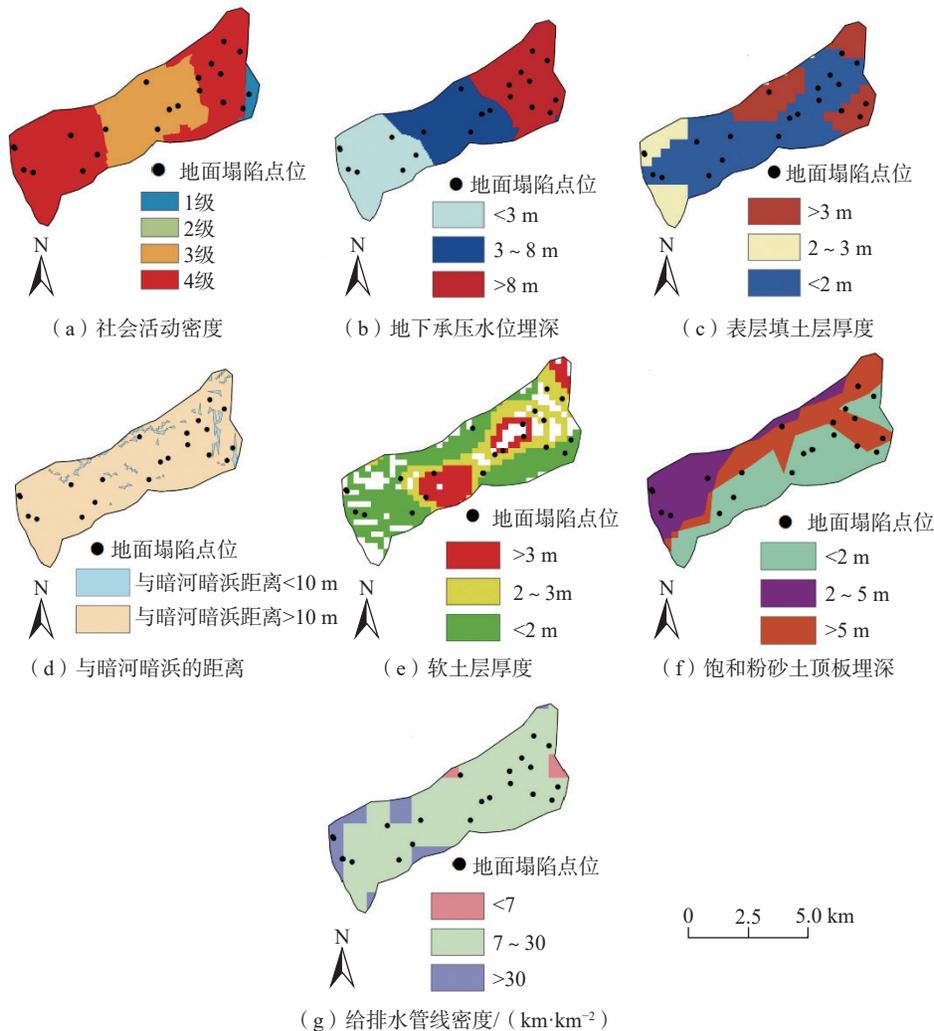


图 6 研究区地面塌陷评价因子分级图

Fig. 6 Grading map of ground collapse evaluation factors in the study area

(2)承压水位是指埋藏在地下的、赋存于两个隔水层之间的地下水。当地下水被过度开采时,承压水位会下降,导致地下的压力减小。这种压力的减小会使得地面上的重力超过地下的支撑力,从而引发地面塌陷。承压水位埋深直接影响潜水位的绝对高程位置,潜水位的最高或最低涨落位次受承压水位控制。通常情况下,承压水位越浅,上覆土层越薄,土体自重越小,潜水位涨落的变幅也越大。当承压水位过浅或过深时,都会导致土层含水量波动增大,潜水位变幅加大。与此同时,两者对于地面塌陷的易发生性及危险性也有很大的影响。

一般情况下,地下承压水埋深越浅,对软土地基的渗流扰动和溶蚀作用越强,减弱土体的结构强度,增大地面塌陷的可能性。将地下承压水位埋深划分为 3 个等级,如图 6(b)所示。

(3)表层填土层厚度。其是指从地表到原始土层之间的人工填充物的垂直深度,表层填土的厚度可能会影响其压实和稳定性,相同的土层结构,填土的厚度越大,其对地下土壤施加的重量压力就越大。如果地下土壤的承载能力不足以支撑这种压力,可能会导致地面塌陷的发生。同时,填土的厚度可能影响地下水的流动和分

布。例如,较厚的填土可能阻止地下水的上升,导致地下土壤的干燥和紧实,从而引发地面塌陷。根据以往调查经验及杭州市地区的实际地质情况,地面塌陷多发生于覆盖层填土厚度大于 3 m 的地区。盖层厚度小于 2 m,发生塌陷的频率明显降低。将表层填土层厚度划分为 3 个等级,如图 6(c)所示。

(4)距暗河暗浜的距离。与暗河暗浜的距离是指与暗河或者暗浜这种不良地质情况的距离,暗河和暗浜附近的土层含水量通常较高,土体强度较低,是地面塌陷的易发区。同时,暗河暗浜水位波动较频繁,会导致周边土层的聚集沉降和不均匀沉降,加大局部区域发生地面塌陷的概率。同时,距离暗河和暗浜越近,受到的水流侵蚀和涌浸作用就越强,地面塌陷的可能性也越大。将与暗河暗浜的距离划分为 2 个等级,如图 6(d)所示。

(5)软土层厚度。与填土类似,研究区软土结构强度也较低,在管道渗漏或较强的地下水长期渗流作用下,易发生水土流失,形成地下空洞导致道路塌陷。在一定深度范围内存在更多的软土层可能会增加地面塌陷的风险。将软土层厚度与评价深度划分为 3 个等级,如图 6(e)所示。

(6)饱和粉砂土顶板埋深。饱和粉砂土顶板埋深是指饱和粉砂土层位于地下的深度,饱和粉(砂)土的顶板埋深对地面塌陷的易发性有重要影响。通常,饱和粉(砂)土顶板埋深越大,地面塌陷的可能性越小。因为较深的顶板埋深可以提供更强的地下支撑,抵抗地面塌陷。顶板埋深越浅,受到外部荷载的影响越大,地面塌陷的可能性就越大。浅层土层承受的压力更大,更容易发生破坏。然而,当顶板埋深超过一定限度时,其增加对地面塌陷的抑制作用会越来越小。因为土体自重产生的压力随埋深的增加呈对数衰减。通过对杭州市地区以往地面塌陷点位的地质统计结果显示,地面塌陷多发生于顶板埋深小于 2 m 的地区。如图 6(f)所示。

(7)给排水管线密度。给排水管线密度是指单位面积内每平方公里的给水和排水管线长度的总和,给排水管道的存在和密度可以影响地面的稳定性。排水管道结构缺陷主要包括错位、脱节、破裂、渗漏等,当管道产生结构缺陷时,排水管线为水土流失提供通道,排水管道标高与区域广泛分布的粉土粉砂层基本一致,该层结构松散、透水性强易于流失,排水管线缺陷附近逐渐形成水土流失带、松散发育带、掏空垮落带、变形沉降带等。往往给排水管线密度越大,地面塌陷发生可能性越大。将给排水管线密度划分为 3 个等级,如图 6(g)所示。

3.3 评价因子的相关性分级

在机器模型评价过程中,评价因子对模型的准确性以及最终评价结果的影响是最大的,当因子间的相关性较高时,机器学习过程中可能会出现过拟合或频繁判别失误等情况,模型的精确程度^[17]也会降低,因此对评价因子进行相关性检验是十分必要的。本研究利用 Python 代码实现,通过 matrix 模块,实现因子相关性热力图的创建。

由图 7 可见,评价因子间不存在相关性过大的情况,因此不需要进行因子调整。

4 研究区地面塌陷易发性评价与分析

4.1 RF 模型、I-RF 模型、RF-BP 神经网络模型预测精度对比

在 ArcGIS 软件中将整个研究区进行 20 m×20 m 像素点化,最终得到 340 个隐患区以及 25 682 个非隐患区。为了确保训练模型的准确性,采取隐患区与非隐患区 1:1 的方式划分训练集(其中隐患区在训练过程中用 1 表示,非隐患区用 0 表示),利用 python 软件随机选用 170 个隐患区以及 170 个非隐患区要素作为训练集,利用训练完成后的预训练模型对剩下的点进行预测,最终得到每个样本点发生地面塌陷的概率(0~1 范围内)。

在这个过程中分别对 RF 模型、RF-I 模型、RF-BP 神经网络模型进行相同的训练集训练以及相同的预测集判断,最后通过模型的 ROC(receiver operating characteristic curve, ROC)曲线来展示模型的准确率^[18]。首先,模型的输出分数通常是连续的值,需要将这些分数转化为二分类的预测结果。此时通常通过设定一个阈值,比如 0.5,来判断模型输出是否为正类或负类,根据不同的阈值,可以计算出一系列的真阳率(TPR)和假阳率(FPR)。其中真阳率(TPR)表示正样本被正确分类为正样本的比例,即模型成功识别出的正例在所有正例中的比例,假阳率(FPR)表示负样本被错误分类为正样本的比例,将不同阈值下得到的 TPR 和 FPR 组成一组坐标点,然后连接这些点,就得到了 ROC 曲线。此外,曲线下的面积(area under the ROC curve, AUC)也是一个性能度量。AUC 值越高,表示模型的性能越好。

3 种不同模型的 ROC 曲线见图 8,当采用随机森林单模型时,AUC 值为 0.73,当将随机森林单模型与信息量模型进行耦合后,集成模型的 AUC 值上升至 0.76,准确度提高了 4%。但由于 2 种模型的 AUC 值都低于 0.8,说明这 2 种模型的预测精度都不够好。在将 RF 与 BP 神经网络模型耦合后,模型 AUC 值上升至 0.86,较随

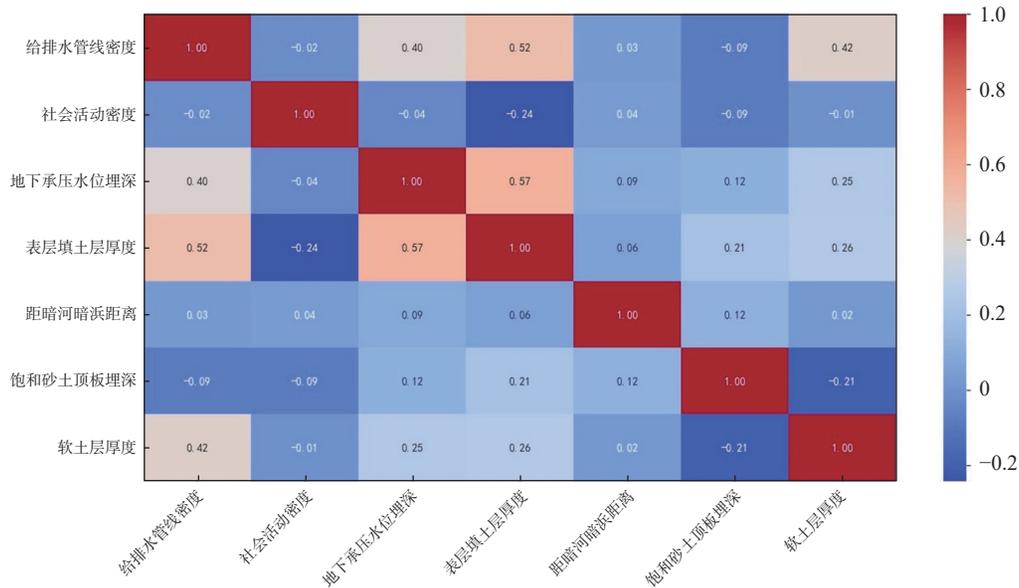


图 7 评价因子相关性热力图

Fig. 7 Heatmap of correlation of evaluation factors

机森林单模型准确度提高了 17.8%, 说明 RF-BP 神经网络集成模型可以有效提高地面塌陷易发性的预测精度, 在该研究区应用良好。

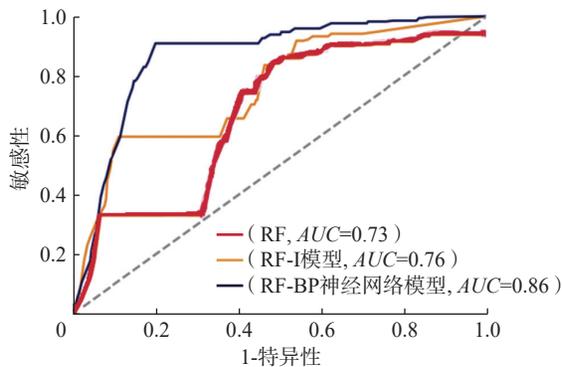


图 8 3 种评价模型的 ROC 曲线对比

Fig. 8 Comparison of ROC curves of three evaluation models

4.2 基于 RF-BP 神经网络模型的因子重要性分析

在已经确定 RF-BP 神经网络模型的预测精度最高后, 为了进一步确认各个因子对研究区地面塌陷灾害发生的贡献程度, 以便在该研究区地面塌陷灾害预防以及治理时对贡献程度较大的因子进行重点关注, 故需要对评价因子进行重要性分析。根据图 9 可以看出, 表层填土层厚度对地面塌陷易发性的发生贡献程度最大, 其次是饱和砂土顶板埋深, 这与研究区地面塌陷点基本在填土范围内, 且发生塌陷点地表以下 5 ~ 10 m 岩性组合为“填土-粉土”, “填土-粉质黏土”, 即填土以及粉土的不稳定是研究区地面塌陷的主要控制因素的实地勘察结

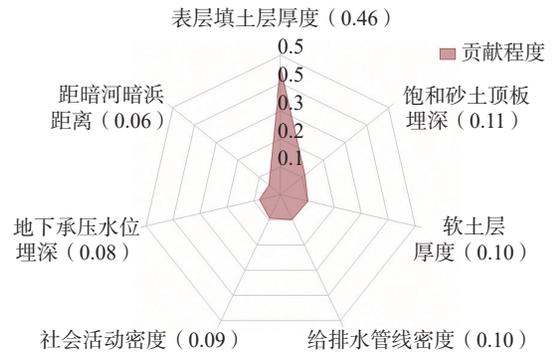


图 9 研究区地面塌陷易发性因子贡献程度雷达分布图

Fig. 9 Radar distribution map of contributing factors to ground collapse susceptibility in the study area

论相吻合。

4.3 基于 RF-BP 神经网络模型结果的研究区易发性评价

在已经确定 RF-BP 神经网络模型的预测精度最高后, 以该集成模型的预测结果为基准, 将全域样本点导入 ArcGIS 软件并利用等间距法^[19-20]将研究区的地面塌陷灾害的概率(0 ~ 1 范围内)进行重分类, 最终划分为 4 类, 分别为低易发(0 ~ 0.25)、中易发(0.25 ~ 0.5)、高易发(0.5 ~ 0.75), 极高易发(0.75 ~ 1), 其中低易发区面积占研究区总面积的 37.87%、中易发区面积占研究区总面积的 8.68%、高易发区面积占研究区总面积的 35.52%、极高易发区面积占研究区总面积的 17.93% (图 10), 研究区各地面塌陷隐患区分布情况与易发等级分区见表 2, 通过表 2 可以得出 21 个隐患区中有 20

个分布于高、极高易发区内, 仅有 1 个隐患区位于中易发区, 隐患区的分布情况与机器学习模型的预测分区结果重合度较高(表 2)。

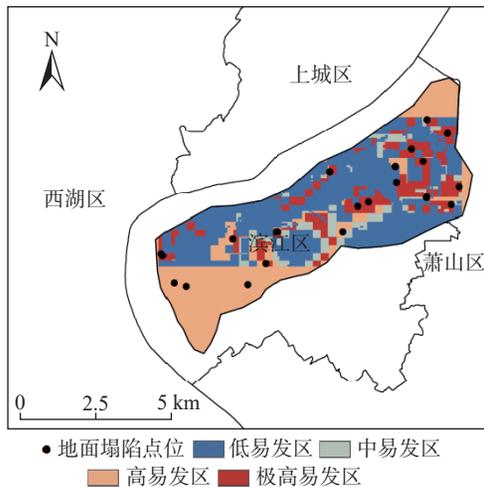


图 10 RF-BP 神经网络模型研究区地面塌陷灾害易发性因子分区图
Fig. 10 RF-BP neural network model study area ground collapse disaster susceptibility factor zoning map

表 2 研究区地面塌陷灾害易发性分区表

Table 2 Susceptibility zoning table for ground collapse disasters in the study area

滑坡灾害 易发性等级	分区面积 /km ²	占总面积比 /%	地面塌陷 隐患区数量/个	占总灾害 数量比/%
低易发	13.97	37.87	0	0
中易发	3.20	8.68	1	4.76
高易发	13.10	35.52	5	23.81
极高易发	6.61	17.93	15	71.43

5 结论

(1) 本文选取杭州市填土-粉砂土典型区为研究区, 针对机器学习应用于城市地面塌陷易发性评价的空白, 通过对评价因子进行相关性检验, 最终选择了给排水管线密度、社会活动密度、地下承压水位埋深、表层填土层厚度、与暗河暗浜距离、饱和砂土顶板埋深、软土层厚度与 7 个评价因子。

(2) 本文利用 Python 构建了 RF、RF-I 集成模型、RF-BP 神经网络集成模型, 并分别利用模型对研究区发生地面塌陷的概率进行预测, 最终应用于该研究区的地面塌陷易发性评价, 其 AUC 值分别为 0.73、0.76、0.86, I-RF 模型、RF-BP 神经网络模型相较于单体模型的预测精度分别提高了 4%、17.8%。说明了集成模型, 尤其是将 BP 神经网络模型与 RF 模型耦合能够有效使预测精度得到提高, 并且选取准确度最高的 RF-BP 神经网络集成模型为研究对象确定了各因子对发生地面塌陷灾

害的影响贡献, 最终得到表层填土厚度的贡献最大, 饱和砂土顶板埋深其次, 这与该区塌陷点基本分布于道路人工填土区域且与砂土层相关性较大的实地勘察结果相吻合。

(3) 利用预测准确度最高的 RF-BP 神经网络集成模型进行易发性分区评价, 将研究区划分为 4 个等级, 分别占比为低易发区面积占 37.87%、中易发区面积占 8.68%、高易发区面积 35.52%、极高易发区面积占 17.93%。最终 95% 的地面塌陷点都分布于高、极高易发区内, 说明易发性评价结果较为准确。

参考文献 (References):

- [1] 刘伟. 我国地质灾害调查统计与分析 [J]. 采矿技术, 2021, 21(5): 100 - 103. [LIU Wei. Investigation, statistics and analysis of geological disasters in China [J]. Mining Technology, 2021, 21(5): 100 - 103. (in Chinese with English abstract)]
- [2] 邵泽斐, 马青山, 葛伟亚, 等. 杭州市地面沉降重点区土体压缩和固结特性研究——以试验孔 HZ22J06 为例 [J/OL]. 华东地质, 2024: 1 - 10. (2024-10-22). [SHAO Zefei, MA Qingshan, GE Weiya, et al. Study on characteristics of soil compression and consolidation in key areas of land subsidence in Hangzhou: Taking experimental hole HZ22J06 as an example [J/OL]. East China Geology, 2024: 1 - 10. (2024-10-22). <https://link.cnki.net/doi/10.16788/j.hddz.32-1865/P.2023.05.001>.]
- [3] 雷廷, 马青山, 葛伟亚, 等. 地面沉降钻孔全断面光纤多参量精细化监测——以杭州萧山区山末址村为例 [J]. 华东地质, 2024, 45(4): 452 - 465. [LEI Ting, MA Qingshan, GE Weiya, et al. Full section multi-parameter precise monitoring of fiber optic in land subsidence borehole: A case study of Shanmozhì Village, Xiaoshan District of Hangzhou [J]. East China Geology, 2024, 45(4): 452 - 465. (in Chinese with English abstract)]
- [4] 于翠翠, 刘兰玉, 李传生, 等. 基于 AHP 法的鲁中某高铁岩溶塌陷易发性评价 [J]. 山东国土资源, 2023, 39(2): 42 - 49. [YU Cuicui, LIU Lanyu, LI Chuansheng, et al. Karst collapse risks evaluation of a high-speed railway in central Shandong based on AHP method [J]. Shandong Land and Resources, 2023, 39(2): 42 - 49. (in Chinese with English abstract)]
- [5] XIE Yanhua, ZHANG Binghui, LIU Yuxin, et al. Evaluation of the karst collapse susceptibility of subgrade based on the AHP method of ArcGIS and prevention measures: A case study of the Quannan expressway, section K1379+300 - K1471+920 [J]. Water, 2022, 14(9): 1432.
- [6] 庄树裕. 基于模糊层次分析法的龙岩市永定区岩溶塌陷

- 灾害易发性评估分析 [J] . *福建地质*, 2022, 41(4): 323 - 331. [ZHUANG Shuyu. Evaluation of karst collapse disaster susceptibility in Yongding District, Longyan City based on fuzzy hierarchical analysis [J] . *Geology of Fujian*, 2022, 41(4): 323 - 331. (in Chinese with English abstract)]
- [7] 江思义, 吴福, 黄希明, 等. 基于专家-层次分析法的岩溶地面塌陷易发性评价——以广西平桂区为例 [J] . *矿产勘查*, 2021, 12(11): 2294 - 2302. [JIANG Siyi, WU Fu, HUANG Ximing, et al. Evaluation of vulnerability of karst collapse based on expert-analytic hierarchy process: A case study of Pinggui Distric, Guangxi [J] . *Mineral Exploration*, 2021, 12(11): 2294 - 2302. (in Chinese with English abstract)]
- [8] 王桂林, 强壮, 曹聪, 等. 基于地理探测器与层次分析法的岩溶地面塌陷易发性评价——以重庆中梁山地区为例 [J] . *中国岩溶*, 2022, 41(1): 79 - 87. [WANG Guilin, QIANG Zhuang, CAO Cong, et al. Evaluation of susceptibility to karst collapse based on the geodetector and analytic hierarchy method: An example of the Zhongliangshan area in Chongqing [J] . *Carsologica Sinica*, 2022, 41(1): 79 - 87. (in Chinese with English abstract)]
- [9] 陈芯宇, 师芸, 赵侃, 等. 基于 CF 融入 SSA 优化 SVM 和 RF 模型的滑坡易发性评价 [J] . *西安理工大学学报*, 2023, 40(1): 121 - 131. [CHEN Xinyu, SHI Yun, ZHAO Kan, et al. Landslide susceptibility evaluation based on CF integrated with SSA to optimize SVM and RF models [J] . *Journal of Xi'an University of Technology*, 2023, 40(1): 121 - 131. (in Chinese with English abstract)]
- [10] COROMINAS J, VAN WESTEN C, FRATTINI P, et al. Recommendations for the quantitative analysis of landslide risk [J] . *Bulletin of Engineering Geology and the Environment*, 2014, 73(2): 209 - 263.
- [11] 陈水满, 赵辉龙, 许震, 等. 基于人工神经网络模型的福建南平市滑坡危险性评价 [J] . *中国地质灾害与防治学报*, 2022, 33(2): 133 - 140. [CHEN Shuiman, ZHAO Huilong, XU Zhen, et al. Landslide risk assessment in Nanping City based on artificial neural networks model [J] . *The Chinese Journal of Geological Hazard and Control*, 2022, 33(2): 133 - 140. (in Chinese with English abstract)]
- [12] 屠水云, 张钟远, 付弘流, 等. 基于 CF 与 CF-LR 模型的地质灾害易发性评价 [J] . *中国地质灾害与防治学报*, 2022, 33(2): 96 - 104. [TU Shuiyun, ZHANG Zhongyuan, FU Hongliu, et al. Geological hazard susceptibility evaluation based on CF and CF-LR model [J] . *The Chinese Journal of Geological Hazard and Control*, 2022, 33(2): 96 - 104. (in Chinese with English abstract)]
- [13] 曾斌, 吕权儒, 寇磊, 等. 基于 Logistic 回归和随机森林的清江流域长阳库岸段堆积层滑坡易发性评价 [J] . *中国地质灾害与防治学报*, 2023, 34(4): 105 - 113. [ZENG Bin, LYU Quanru, KOU Lei, et al. Susceptibility assessment of colluvium landslides along the Changyang section of Qingjiang River using Logistic regression and random forest methods [J] . *The Chinese Journal of Geological Hazard and Control*, 2023, 34(4): 105 - 113. (in Chinese with English abstract)]
- [14] 冉星星. 基于机器学习的深圳市地面坍塌灾害风险评估 [D] . 荆州: 长江大学, 2023. [RAN Xingxing. Risk assessment of ground collapse in Shenzhen based on machine learning [D] . Jingzhou: Yangtze University, 2023. (in Chinese with English abstract)]
- [15] 王子童. 煤矿采空区地面塌陷危险性评价与三维可视化 [D] . 西安: 西安科技大学, 2021. [WANG Zitong. Risk evaluation and three-dimensional visualization of ground collapse in coal mine goaf [D] . Xi'an: Xi'an University of Science and Technology, 2021. (in Chinese with English abstract)]
- [16] NAKAMURA H, KUBOTA T, FURUKAWA M, et al. Unified construction of running track tunnel and crossover tunnel for subway by rectangular shape double track cross-section shield machine [J] . *Tunnelling and Underground Space Technology*, 2003, 18(2/3): 253 - 262.
- [17] 黄发明, 胡松雁, 闫学涯, 等. 基于机器学习的滑坡易发性预测建模及其主控因子识别 [J] . *地质科技通报*, 2022, 41(2): 79 - 90. [HUANG Faming, HU Songyan, YAN Xueya, et al. Landslide susceptibility prediction and identification of its main environmental factors based on machine learning models [J] . *Bulletin of Geological Science and Technology*, 2022, 41(2): 79 - 90. (in Chinese with English abstract)]
- [18] GONÇALVES L, SUBTIL A, OLIVEIRA M R, et al. ROC curve estimation: An overview [J] . *REVSTAT-Statistical journal*, 2014, 12(1): 1 - 20.
- [19] 吴宏阳, 周超, 梁鑫, 等. 基于 XGBoost 模型的三峡库区燕山乡滑坡易发性评价与区划 [J] . *中国地质灾害与防治学报*, 2023, 34(5): 141 - 152. [WU Hongyang, ZHOU Chao, LIANG Xin, et al. Assessment of landslide susceptibility mapping based on XGBoost model: A case study of Yanshan Township [J] . *The Chinese Journal of Geological Hazard and Control*, 2023, 34(5): 141 - 152. (in Chinese with English abstract)]
- [20] 牟家琦, 庄建琦, 王世宝, 等. 基于深度神经网络模型的雅安市滑坡易发性评价 [J] . *中国地质灾害与防治学报*, 2023, 34(3): 157 - 168. [MOU Jiaqi, ZHUANG Jianqi, WANG Shibao, et al. Evaluation of landslide susceptibility in Ya'an City based on depth neural network model [J] . *The Chinese Journal of Geological Hazard and Control*, 2023, 34(3): 157 - 168. (in Chinese with English abstract)]