



移动阅读

张晓博, 周萍, 张焜, 等, 2023. 基于 GIS 和 CF-Logistic 回归模型地质灾害易发性评价: 以青海湟中县为例[J]. 沉积与特提斯地质, 43(4): 797–807. doi: [10.19826/j.cnki.1009-3850.2020.12003](https://doi.org/10.19826/j.cnki.1009-3850.2020.12003)

ZHANG X B, ZHOU P, ZHANG K, et al., 2023. Evaluation method of geological hazard susceptibility: A case study on GIS and CF-Logistic regression model in Huangzhong, Qinghai[J]. Sedimentary Geology and Tethyan Geology, 43(4): 797–807. doi: [10.19826/j.cnki.1009-3850.2020.12003](https://doi.org/10.19826/j.cnki.1009-3850.2020.12003)

基于 GIS 和 CF-Logistic 回归模型地质灾害易发性评价：以青海湟中县为例

张晓博^{1,3}, 周萍², 张焜^{1,3}, 张兴^{1,3}, 刘宝山^{1,3}, 邓辉^{2*}

(1. 青海省地质调查院, 青海省青藏高原北部地质过程与矿产资源重点实验室, 青海 西宁 810012; 2. 成都理工大学地球科学学院, 四川 成都 610059; 3. 自然资源青海卫星应用技术中心, 青海 西宁 810012)

摘要: 研究地质灾害易发性的评价方法, 对地质灾害防治有着非常重要的现实意义。本文以青海省西宁市湟中县为研究区域, 选取高程、坡度、坡向、地形起伏度、距河流距离、距断层距离和工程岩组 7 个评价因子, 利用确定性系数与逻辑回归模型相结合的方法计算出每个单元格地质灾害发生的概率。同时利用 ROC 曲线和 AUC 值对模型分类精度进行验证, 最终得到 AUC 值为 0.863, 说明该方法对湟中县地质灾害易发性评价具有较强的适用性和客观性。本文研究表明, 高层、坡向、距河流距离和工程岩组 4 个因子对研究区地质灾害的影响最为显著。从地质灾害的空间分布来看, 该方法计算结果表明极高、高易发区主要分布在湟水河及其干流两侧低山丘陵地区, 低易发区主要分布在研究区北部和西南地区。从评价因子的角度分析, 高易发区主要分布在离河流 500 m 的松散冲洪积岩层和软弱层状碎屑岩岩层上。以上研究结果表明, 基于 CF-Logistic 回归模型对研究区地质灾害易发性评价有较强的参考价值, 能为研究区地质灾害的防治工作提供理论依据及方法。

关键词: 地质灾害; GIS; 易发性评价; 确定性系数; Logistic 回归模型; 湟中县

中图分类号: P642.23

文献标识码: A

Evaluation method of geological hazard susceptibility: A case study on GIS and CF-Logistic regression model in Huangzhong, Qinghai

ZHANG Xiaobo^{1,3}, ZHOU Ping², ZHANG Kun^{1,3}, ZHANG Xing^{1,3},
LIU Baoshan^{1,3}, DENG Hui^{2*}

(1. Key Laboratory of the Northern Qinghai-Tibet Plateau Geological Processes and Mineral Resources, Qinghai Geological Survey Institute, Xining 810012, China; 2. College of Earth Sciences, Chengdu University of Technology, Chengdu 610059, China; 3. Qinghai Satellite Application Technology Center of Natural Resources, Xining 810012, China)

Abstract: The evaluation method for the susceptibility of geological hazards has very important practical significance for the prevention and control of geological hazards. This paper takes Huangzhong County, Xining City, Qinghai Province as the research area to investigate the most suitable assessment method for the area's geological hazard susceptibility. The results show that using seven evaluation factors (including elevation, slope, aspect, topographic undulation, distance from river, distance from fault, and

收稿日期: 2020-11-16; 改回日期: 2021-04-16; 责任编辑: 郭秀梅

作者简介: 张晓博(1983—), 男, 主要从事地质灾害评价研究。E-mail: 170421709@qq.com

通讯作者: 邓辉(1984—), 男, 主要从事地质灾害隐患早期识别与评价。E-mail: dengh@cdu.edu.cn

资助项目: 青海省科技厅重点研发与转化计划(2019-SF-130); 青海省科技厅应用基础研究(2018-ZJ-737)

engineering rock group) and the certainty factor and logistic regression model to evaluate the geological hazard susceptibility of Huangzhong County is suitable and reliable. Using the determinant coefficient and logistic regression model, we calculate the probability of geological hazard in each cell. Subsequently, we use the ROC curve and the AUC value to verify the classification accuracy of the model. The AUC value is 0.863, indicating that our method has good applicability for evaluating the susceptibility of geological disasters in the study area. The four factors of high-rise, aspect, distance from river, and engineering rock group have the most significant impacts on geological disasters in the study area. From the perspective of the spatial distribution of geological hazard susceptibility, extremely high and high susceptibility areas calculated using this method are mainly distributed in the low mountain and hilly areas on both sides of the Huangshui River and its main stream, while low susceptibility areas are mainly distributed in the northern and southern regions of the study area. From the perspective of geological hazard evaluation factors, high susceptibility areas are mainly distributed on loose alluvial strata and weak clastic rock strata 500 m away from the river. The research results indicate that the CF-Logistic regression model has strong reference value for evaluating the susceptibility of geological hazards in the study area and can provide theoretical and methodological basis for the prevention and control of geological hazards in the study area.

Key words: geological hazard; GIS; susceptibility evaluation; certainty coefficient; logistic regression model; Huangzhong County

0 引言

在地球自身运动或者人类的工程实践等因素的影响下,发生的危害人类生命财产、生产生活活动或破坏人类赖以生存与发展的资源与环境的地质事件被称为地质灾害。中国是全球地质灾害高发区之一,地质灾害不但会严重破坏该区域的生态环境,而且还会严重威胁人类的生命财产安全。近年来,国内外众多学者对地质灾害易发性评价模型进行了深入研究,并取得了丰硕成果。贺鹏等(2016)利用层次分析法对西藏札达地区重大滑坡灾害进行易发性评价;洪增林等(2020)利用主成分分析法对榆神府地区进行地质灾害危险性分级评估;向喜琼和黄润秋(2000)运用人工神经网络的方法,结合GIS对长江三峡示范区(巴东—新滩)地区的地质灾害进行了危险性区划;饶品增等(2017)利用地理加权回归模型对云南省地质灾害进行易发性评价;胡凯衡等(2012)采用聚类分析和最大似然法对汶川重灾区震后地质灾害进行易发性评价;沈玲玲等(2016)应用模糊逻辑法、信息量模型及熵改进的信息量模型,对岷县地震诱发的滑坡灾害进行了易发性评价;刘艳芳等(2014)基于确定性系数分析方法对秭归县滑坡灾害进行易发性评价;覃乙根(2020)等采用确定性模型和Logistic回归模型两种模型耦合的方法对贵州省开阳县的地质灾害易发性进行评价。各项研究表明,多模型耦合在评价精度、合理性等方面比单一模型具有更明显的

优越性。

东接青藏高原,西连黄土高原,地处二者结合部的湟中县生态环境极为脆弱,而青藏高原独特的地质构造以及湟中县所处地段的特殊性,使得此处外力地质作用极易形成并诱发各种地质灾害。由外力地质作用所诱发的灾害现象主要包括崩塌、滑坡、泥石流、不稳定斜坡等。本研究区地处青海省东部,人口密集。该区域的人类工程经济活动强烈,导致湟水河谷及其较大支沟谷与丘陵交接部位地质灾害频繁发生,形成这一地区最为严重的环境地质问题。本文结合湟中县地质灾害的空间分布状况,共选取高程、坡度、坡向、地形起伏度、工程地质岩组,以及断裂距离、距水系距离7个地质灾害孕灾因子,并运用CF模型和Logistic回归模型对研究区内的地质灾害易发性进行分析,最终得到湟中县地质灾害易发性分级图。合理判断湟中县地质灾害的易发程度,可以有效地规避地质灾害发生的风险,同时可以为该区域的地质灾害防治与管理提供可靠的科学理论依据(陈绪钰等,2019)。

1 研究区概况

本研究以青海省西宁市湟中县(以下简称湟中县)为研究区域。该研究区地处青海省东部,地理坐标为 $36^{\circ}13'32''\sim 37^{\circ}03'19''N$, $101^{\circ}09'32''\sim 101^{\circ}54'50''E$,总面积达 $2\,570\text{ km}^2$ 。湟中县西起扎麻隆峡口,东至小南川,北起娘娘山,南至拉脊山,总体地势呈东西走向,海拔向东降低,同时南北两侧高于中央腹

地,地形条件复杂,高差大。从山区到河谷平原,现代地貌景观反映了强烈的上升运动和地区性垂直气候分布而产生的深切作用。海拔 3 900 m 以下为水网强烈切割,河谷呈“V”字形,峡谷,谷坡呈直线形;基岩山区和中高山、低山丘陵区交接线外观十分平直;中高山除大的河谷形成狭长冲积平原外,均为冈峦起伏的丘陵。该区的地貌可划分为剥蚀中高山区、侵蚀低山丘陵区、侵蚀堆积平原区三个不同的地貌单元。其中剥蚀中高山集中分布在县境南北两侧拉脊山和娘娘山一带,海拔在 3 500 m 以上,主要由元古界白云岩、石英岩、片岩、片麻岩和花岗岩组成,山体侵蚀强烈,沟谷形态为“V”型谷,由于本区人类工程活动较少,地质灾害以局部崩塌为主,规模较小;剥蚀低山丘陵区主要分布于拉脊山北侧、娘娘山南侧的中高山山前地带,主要由古近纪砂岩组成,上覆风积黄土。丘陵区蒸发量大,地形极不平坦,沟壑纵横,水土流失比较严重。总体上,该地人类活动较强烈,植被覆盖率低,是崩塌、滑坡和泥石流等地质灾害最容易发生的地区;侵蚀堆积平原主要分布于湟水、西纳川、康城川和南川等河谷区,地势相对低平,植被较发育。本区地质灾害不发育,但在高阶地前缘,崩塌、滑坡灾害较常见。县域内大部分乡、镇分布在湟水河谷及一级支流谷底与山间丘陵,属于典型的山间河谷型县区。

湟中县内地质灾害十分严重,主要的地质灾害为滑坡、崩塌、泥石流等,如图 1 所示,区内地质灾害以滑坡和崩塌为主,泥石流发育较少,该区域共有地质灾害 417 处,其中滑坡 374 处,崩塌 29 处,泥石流 14 处。研究区滑坡和崩塌分为土质和岩质两类,其中以土质滑坡为主,共 355 处,占总滑坡数量的 94.92%,岩质滑坡共 19 处,其中有 213 处滑坡处于不稳定状态,161 处滑坡有轻微活动特征,处于基本稳定状态;崩塌也以土质崩塌为主,规模以中小型为主,小型崩塌 14 处,中型崩塌 11 处。

2 研究方法

逻辑回归模型在对地质灾害易发性评价时能够根据评价因子和灾害点之间的关系确定评价因子的权重,但在变量处理时存在很大的主观性;确定性系数模型能够解决多因子数据的量化问题,并且计算出评价因子各分级的相对权重。前人经验表明,将两种方法结合能得到更客观准确的结果,

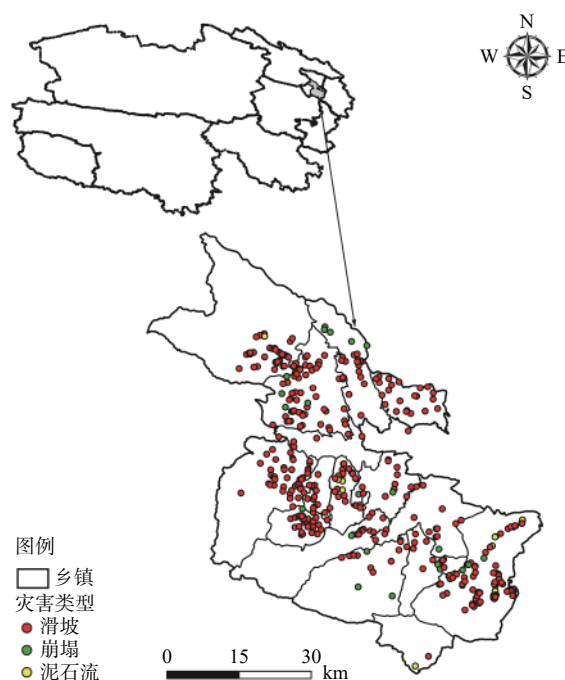


图 1 研究区位置示意图

Fig. 1 The location of research area

因此本文采取两种模型结合的方法,将计算得到的 CF 值作为自变量代入逻辑回归公式中进行二元逻辑回归分析,得到湟中县地质灾害的易发性概率,以此作为易发性分区的依据。

2.1 确定性系数模型

确定性系数模型(CF)是 1975 年由 Shortliffe 和 Buchanan(1975)提出的一个概率函数,Heckerman (1985)对其进行改进。同其他统计模型一样,确定性系数模型假设其将来发生地质灾害的原因和过去发生地质灾害的原因相同。CF 计算公式为:

$$CF = \begin{cases} \frac{PP_a - PP_s}{PP_a(1 - PP_s)}, & PP_a \geq PP_s \\ \frac{PP_a - PP_s}{PP_s(1 - PP_a)}, & PP_a < PP_s \end{cases} \quad (1)$$

式中: PP_a 为事件(地质灾害)在因子分类数据 a 里面发生的概率,在实际的研究中通常用因子分类 a 中的地质灾害个数(或面积)与数据分类 a 的面积比值表示; PP_s 为整个研究区的灾害总个数(或面积)与研究区总面积的比值。

CF 代表地质灾害发生的概率,其值在 [-1, 1] 范围内变化。若 CF 为正值,则意味着事件发生确定性的增长,即地质灾害发生的确定性高,说明地质环境条件易于发生地质灾害;负值则意味着事件发生的确定性降低,即地质灾害发生的确定性低,不易发生地质灾害, -1 表示不发生地质灾害。

2.2 逻辑回归模型

逻辑回归模型是一种二分类因变量常用的统计分析方法。它描述的是二元因变量(通常,0代表地质灾害不存在,1则代表地质灾害存在)和多个致灾因子($x_1, x_2 \dots, x_n$)之间的关系。该模型不限定自变量是否连续,并且不要求自变量满足正态分布。逻辑回归模型公式如下:

$$P(y=1) = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \dots + \beta_n x_n)}} \quad (2)$$

式中: P 为滑坡发生概率,范围为0~1; α 为逻辑回归计算出的一个常数项; β 为逻辑回归计算而得出的回归系数; i 为评价因子种类数目。

将式(2)两边取自然对数,等式左边作为因变量,将影响因子 $x_i (i=1, 2, \dots, n)$ 作为自变量,得:

$$\ln(P/(1-P)) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n = \alpha + \beta x \quad (3)$$

3 数据来源和评价因子的选择及分级

3.1 数据来源

本研究所用到的样本地质灾害点数据包括滑坡149处、崩塌9处、泥石流14处、不稳定斜坡225处,共计417个地质灾害点,地质灾害点空间分布位置如图1。除了样本数据以外,地质灾害易发性评价的基础数据有湟中县DEM数据、地质数据和水文数据。本研究区的坡度是根据对地观测与数字地球中心(<http://www.eorc.jaxa.jp/ALOS/en/aw3d30/>)下载的分辨率为30 m的DEM数据生成。

3.2 评价因子的选择和分级

3.2.1 高程

高程是区域宏观地貌的一个表征,很多研究表明地质灾害与高程分布具有明显的区域规律(田春山等,2016)。湟中县地势大体呈东西走向,海拔由西向东降低,同时南北两侧高于中央腹地,地形条件复杂,高差大。如图3(a)所示。考虑到研究尺度和湟中县地貌实际情况,运用自然断点法7个等级,如图2(a)所示,从图中可以看出,地质灾害点主要分布在2285~2745 m之间,占总灾害点数量的85.61%。

3.2.2 坡度因子

坡度是影响地质灾害发生的一个重要因素,其大小决定了地表松散物质发生位移并且形成地质灾害的可能性,同时也在很大程度上确定了斜坡造成破坏的形式和机制(杨德宏和范文,2015;赵艳南

和牛瑞卿,2010)。基于DEM数据,利用ArcMap10.5中的Slope命令生成坡度分布图,如图3(b)所示。根据地质灾害在每个坡度段内的分布特点,把坡度按照5°为间隔分为9个级别,由于40°以上都地质灾害点极少,故划分为一个等级。利用ArcGIS的空间分析功能,统计地质灾害在各个坡度段的分布规律,如图2(b)所示,从研究区坡度图来看,区内坡度陡峭,区内坡度分布在0°~73.04°之间。整个研究区内地质灾害总数为417个,从地质灾害与坡度的分布来看,地质灾害主要发生在5°~25°之间,灾害点个数为322个,占研究区内灾害点的77.22%。

3.2.3 坡向因子

坡向是指某一地面点处高程变化量最大的方向(汤国安等,2016)。坡向主要影响山坡面水热的大小,以及小气候作用的差异。向阳一面坡体水热条件充足,自然活动频繁,因而沟谷发育多陡坡,风化作用强烈,导致岩体裂隙发育易破裂,而阴坡土层多易于累积堆积。由图2(c)和图3(c)可知,地质灾害主要发生在东—南坡向范围内,地质灾害占总数的55.64%,由此推出该区域断层走向为东西向。

3.2.4 地形起伏度因子

地形起伏度是指某一确定区域内海拔最高点与海拔最低点的差值(程维明等,2009)。地形起伏度和坡度相互补充,两者在地质灾害评价中都发挥着独具风貌的重要作用。这两个因素均与地质灾害具有良好的相关性,特别是在高山峡谷和深切河谷地区,地质灾害容易发生。本次研究中的地形起伏度计算是借助ArcGIS软件,对DEM数据进行邻域分析和栅格计算器完成的。将地形起伏度按自然间断点法分为4类,研究区内的平坦起伏(0~15 m)地区面积占总面积的30.73%,小起伏(15~30 m)地区面积占总面积的35%,中起伏(30~50 m)地区面积占总面积的24.26%,山地起伏(50~205 m)地区面积占总面积的10.01%。这些统计结果如图3(d)所示,研究区内地形集中以平坦起伏和小起伏变化呈现,同时在小范围面积内分布着山地起伏地形。研究表明,平坦起伏和小起伏地区是地质灾害发生的主要区域,占研究区灾害点的79.38%。

3.2.5 水系因子

地质灾害的发育会受到水系发育程度和分布密度的影响,因此水文条件也是影响地质灾害形成

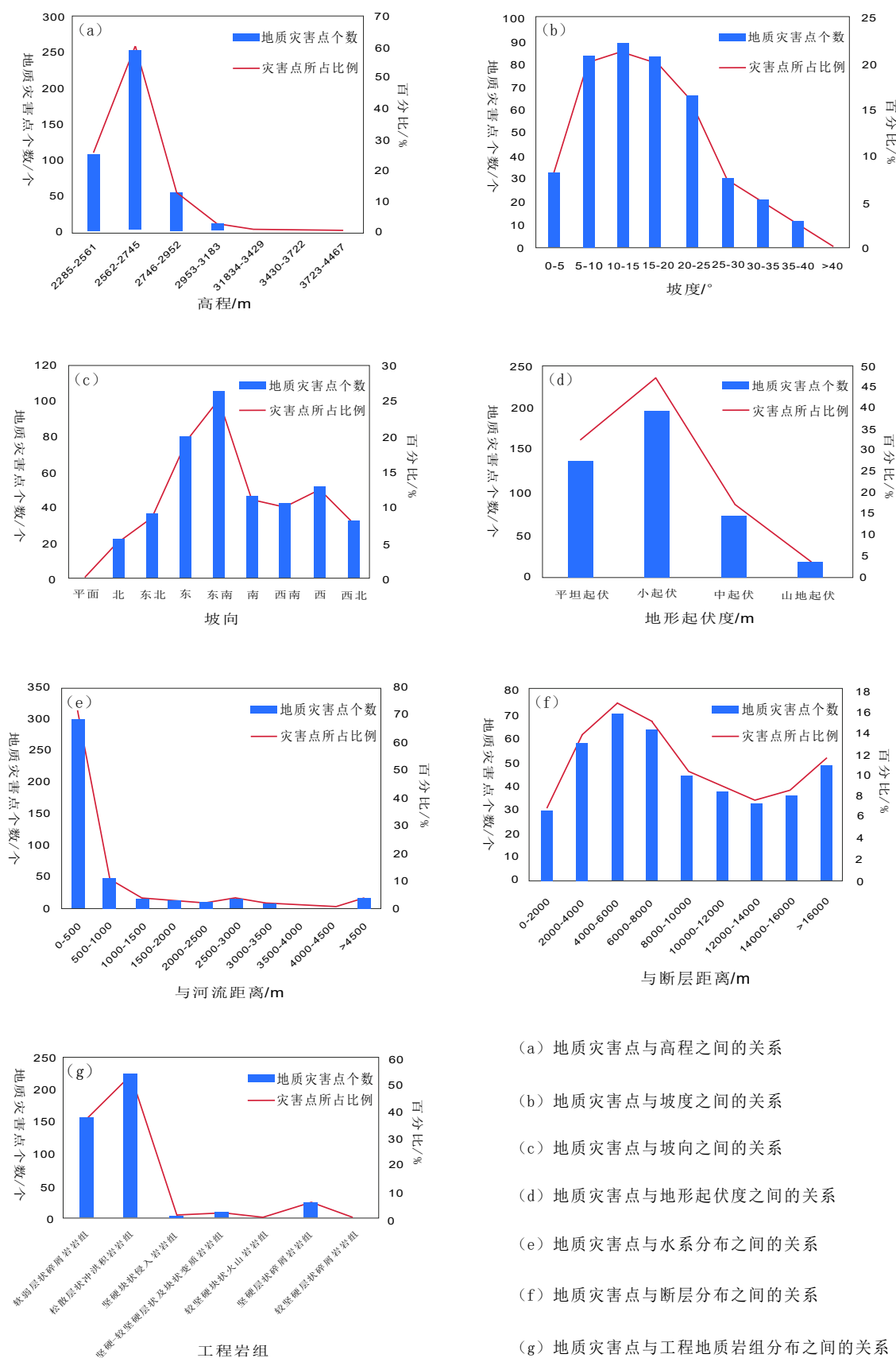


图2 地质灾害点与各评价因子之间的关系

Fig. 2 Relation between geological hazards and evaluation factors

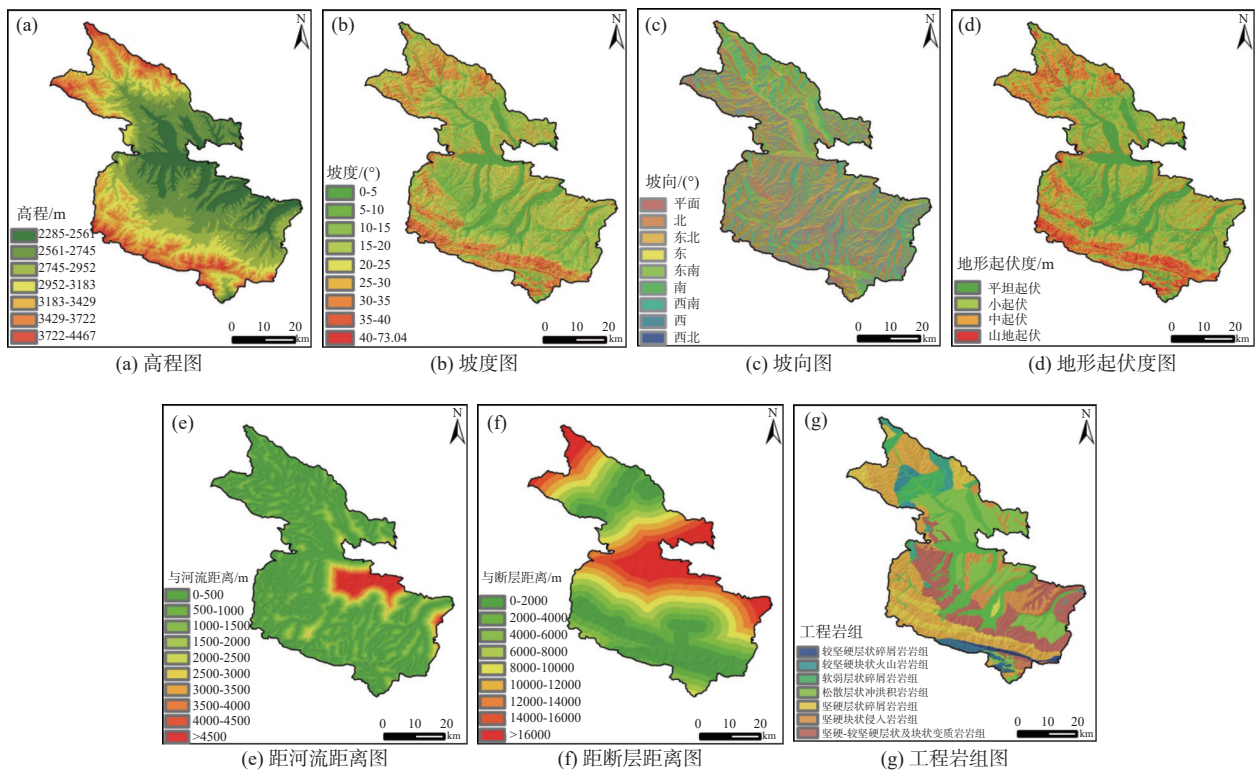


图3 地质灾害易发性评价因子分级

Fig. 3 Classification of susceptibility evaluation factors of geologic hazard in Huangzhong

和发育的一个重要因素(李鑫, 2009)。湟中县属黄河一级支流湟水流域, 主要河流有湟水、康城沟、西纳川河等 15 条, 其中多年平均流量 $0.5 \text{ m}^3/\text{s}$ 以上的河流有 10 条, 河流众多, 河网密布。本文中水系是通过 DEM 数据提取得到的, 利用 ArcGIS 软件中欧式距离分析, 以 500 m 为间隔生成水系因子图, 如图 2(e) 和图 3(e) 所示, 地质灾害主要分布在距离河流 1000 m 内, 该区域共有地质灾害 341 个, 占总灾害数量的 81.77%, 距离河流越远, 地质灾害发育越少。

3.2.6 断裂构造因子

地质灾害的发育一般都与断裂构造密切相关, 断裂构造影响到区域的稳定性, 尤其是在断层附近, 地质作用较活跃, 岩土体容易遭到破坏, 松散的岩土体为地质灾害的形成和发育提供有力的构造条件和物质来源(郭芳芳等, 2008; 黄润秋等, 2008)。湟中县的中部地区是新生代西宁断陷盆地, 南侧为拉脊山, 北侧为娘娘山背斜褶皱带, 受达坂山南麓断裂和拉脊山北麓断裂控制。本文根据距离断层的距离来刻画断层对地质灾害易发的影响, 把距断层距离以 2 km 为间距作距离分析, 共分为 9 级, 如图 3(f) 所示, 从图 2(f) 中可以看出, 研究区地质灾

害主要分布在断层附近 10 km 的范围内, 一共发育地质灾害 264 处, 占区域地质灾害总数的 63.31%。

3.2.7 工程地质岩组因子

工程地质岩组影响着地质灾害的发育和形成, 以往研究成果表明, 地层岩性是产生地质灾害的基本物质, 岩性的类型不同, 其具有的物理力学性质也不同, 因此地质灾害发生的类型及规模也往往各不相同, 而且在不同地质年代中, 地层的岩性与地质灾害也高度相关(冯新科, 2014)。本文根据岩石成因和性质将湟中县的岩石划分成 7 个岩组大类: 以元古界花岗岩、花岗闪长岩、变质岩为主的坚硬块状侵入岩岩组; 以碎屑岩和变质岩为主的坚硬-较坚硬层状及块状变质岩岩组; 以三叠纪砂岩为主的较坚硬块状变质岩岩组; 以古近系砾岩、砂岩、泥岩等组成的坚硬层状碎屑岩岩组; 由白垩系砾岩、砂砾岩等组成的较坚硬层状碎屑岩岩组; 以古近纪和新近纪泥岩、石膏岩等组成的软弱层状碎屑岩岩组; 由风积黄土、砾石土、崩滑堆积体等组成的松散层状冲洪积岩组。从图 2(g) 中可以看出, 研究区地质灾害主要分布在软弱层状碎屑岩岩组和松散层状冲洪积岩岩组。

4 基于 CF 和 Logistic 模型的地质灾害易发性评价

本文对青海省湟中县地质灾害易发性的评价,首先基于统一的矢量范围,对所有的评价因子栅格数据进行裁剪。同时,根据像元大小设置每个因子图层的栅格大小为30 m×30 m。采用 CF 模型计算出高程、坡度、坡向、地形起伏度、距河流距离、距断层距离、工程岩组 7 个影响因子各分级的权重(即 CF 值),随后把各评价因子的 CF 值视为自变量,把是否发生地质灾害视为因变量(1 代表发生地质灾害,0 代表不发生地质灾害),采用 SPSS25.0 软件进行二元逻辑回归分析,把分析结果中得出的各因子的回归系数作为该因子的权重,从而计算出湟中县地质灾害的易发程度。具体计算步骤如下:

(1)确定性系数模型(式 1)计算出 7 个影响因子各分类级别的 CF 值。

根据灾害编目中的 417 个地质灾害点在各因子分级中的分布,利用 CF 模型可以计算出各分类级别在因子中的权重,该权重不仅可以比较同一评价因子下各分类级别的相对重要性,还可以在不同评价因子的分类级别之间进行比较,各分类级别的 CF 值见表 1。

(2)各因子之间独立性检验

为了使选择的因子更加合理,需要对各影响因子之间的独立性进行分析,剔除关联性较大的因子,各因子之间的相关矩阵见表 2。表中数据显示,地形起伏度与高程、坡度和工程岩组具有较高的相关系数,其值均大于 0.3,其余各因子之间的相关系数基本则不足 0.3。因此,舍去地形起伏度因子,保留剩余 6 个因子。

(3)利用逻辑回归模型得到各因子权重

将各影响因子分级中的 CF 值作为自变量,是否发生地质灾害作为因变量。本文将研究区内 417 个地质灾害点作为训练样本,并在距灾害点 1 km 范围外随机选取 417 个点作为非地质灾害样本,这样一共得到 834 个具有独立属性的样本,根据灾害点与非灾害点样本进行各种因子的综合评价。回归结果见表 3。

在本次研究过程中, B 代表各因子权重的大小,由 sig 值来判断自变量在方程中的重要程度,若 sig 值大于 0.05,则没有意义。从回归结果可以看

出,地形起伏度的 sig 值为 0.164,大于 0.05,说明该因子是无效的,所以剔除地形起伏度这一因子。

(4)地质灾害易发性的计算

基于各因子的回归系数,根据式(2)和式(3)可以得到逻辑回归方程:

$$Y = -0.663 + 1.993x_1 + 1.283x_2 + 1.338x_3 + 1.539x_4 + 0.281x_5 + 0.858x_6$$

$$P = \frac{1}{1 + e^{-Y}} \quad (4)$$

其中, P 表示发生地质灾害的概率; x_1 为高程因子中各分级的 CF 值; x_2 为坡度因子中各分级的 CF 值; x_3 为坡向因子中各分级的 CF 值; x_4 为距河流距离中各分级的 CF 值; x_5 为距断层距离中各分级的 CF 值; x_6 为工程岩组中各分级的 CF 值。

5 评价结果及检验

5.1 易发性评价结果分析

根据式(4)可以得到研究区地质灾害发生的概率 P 。而地质灾害易发性概率可以依照自然间断点法归类为五个级别:极低易发区、低易发区、中易发区、高易发区和极高易发区。通过对研究区地质灾害易发性分区的面积和灾害点数量进行统计分析,如表 4 所示,研究区内地质灾害极高、高度易发区的总面积为 978.046 km²,占研究区总面积的 38.051%,其中极高易发区分布 216 个地质灾害点,占研究区总数的 51.799%,高易发区分布 156 个地质灾害点,占研究区总灾害点的 37.410%,而极低、低易发区面积较大但灾害点分布极少,仅占地灾点的 2.158%。通过统计发现灾害点的分布密度随着易发性等级的升高而增加,两者成正相关关系,说明地质灾害的易发性分区与灾害点的拟合程度较高,表明采用 CF-Logistic 回归模型进行湟中县地质灾害易发性评价能较好地反映区内地质灾害易发性的空间分布情况。

研究区内极高和高易发区主要分布在湟水干流南北两岸较大支流两侧低山丘陵地区,总体呈南北向展布,如图 4 所示,研究分析发现,极高和高易发区地层岩性以全新统、上、中更新统时期的冲洪积岩土为主,该土体具有大孔隙及垂直节理,其力学性质较差,在临空条件下发育大量滑坡、崩塌等地质灾害。从河流水系来看,高易发区和极高易发区河网水系密布,常年性河流和季节性河流数量较多,水资源丰富,河流对斜坡坡脚和坡面的冲刷、侵蚀作用明显,易发育不稳定斜坡和泥石流等地质

表 1 各影响因子分类级别的 CF 值
Table 1 Values of CF for the seven factors influencing geologic hazard in Huangzhong

评价因子	分类级别	类别面积/km ²	灾害点/个	CF	评价因子	分类级别	类别面积/km ²	灾害点/个	CF
高程/m	2 285~2 561	409.002	106	0.446 449	距河流 距离/m	0~500	1 336.213	297	0.322 411
	2 562~2 745	716.86	251	0.640 582		500~1 000	651.169	44	-0.625 782
	2 746~2 952	418.673	52	-0.267 671		1 000~1 500	228.218	15	-0.636 713
	2 953~3 183	355.947	8	-0.881 270		1 500~2 000	94.841	12	-0.251 972
	3 184~3 429	329.91	0	-1.000 000		2 000~2 500	54.777	7	-0.243 409
	3 430~3 722	239.063	0	-1.000 000		2 500~3 000	39.933	16	0.710 336
	3 723~4 467	100.912	0	-1.000 000		3 000~3 500	30.792	7	0.341 811
坡度/°	0~5	372.734	33	-0.498 401	距断层 距离/m	3 500~4 000	25.795	3	-0.320 384
	5~10	320.274	84	0.455 303		4 000~4 500	21.518	2	-0.470 853
	10~15	396.794	89	0.330 288		>4 500	87.111	14	-0.011 157
	15~20	388.188	83	0.287 955		0~2 000	550.712	29	-0.712 956
	20~25	341.139	66	0.192 716		2 000~4 000	359.757	58	-0.007 449
	25~30	302.074	30	-0.430 602		4 000~6 000	281.879	70	0.413 851
	30~35	256.745	20	-0.563 754		6 000~8 000	238.574	63	0.460 318
	35~40	144.7	11	-0.575 143		8 000~10 000	220.945	44	0.221 240
坡向/°	>40	47.719	1	-0.889 468		10 000~12 000	213.104	37	0.078 309
	平面	2.247	0	-1.000 000	工程岩组	12 000~14 000	211.939	32	-0.081 653
	北	357.756	22	-0.661 639		14 000~16 000	192.899	36	0.156 013
	东北	404.466	36	-0.495 471		>16 000	300.557	48	-0.018 560
	东	398.679	80	0.228 596		坚硬块状侵入岩岩组	175.234	3	-0.910 053
	东南	300.292	106	0.645 050		较坚硬块状火山岩岩组	927.608	0	-1.000 000
	南	233.271	46	0.211 630		坚硬层状碎屑岩岩组	98.676	25	0.429 305
	西南	244.869	42	0.064 628		软弱层状碎屑岩岩组	674.074	154	0.346 023
	西	290.841	52	0.110 546		较坚硬层状碎屑岩岩组	18.65	0	-1.000 000
地形起伏度/m	西北	337.945	33	-0.441 177		松散层状冲洪积岩岩组	845.456	225	0.465 994
	平坦起伏	789.873	135	0.060 620		坚硬-较坚硬层状 及块状变质岩岩组	730.669	10	-0.928 345
	小起伏	899.687	196	0.304 749					
	中起伏	623.485	70	-0.346 908					
	山地起伏	257.322	16	-0.657 623					

表 2 影响因子间的相关系数矩阵
Table 2 Correlation matrix of seven factors

因子	高程	坡度	坡向	地形起伏度	距河流距离	距断层距离	工程岩组
高程	1	0.241	0.105	0.311	0.134	0.115	0.182
坡度		1	-0.024	0.693	0.052	0.132	0.296
坡向			1	-0.024	0.077	0.137	0.101
地形起伏度				1	0.036	0.153	0.390
距河流距离					1	0.057	0.124
距断层距离						1	0.225
工程岩组							1

灾害。北部和西南地区主要为低易发区和极低易发区,分布面积为 1 105.909 km²,占研究区总面积的 43.025%,海拔较高,主要为坚硬岩和较坚硬岩

岩组,抗风化能力强,河流主要为支流,河网稀疏,没有发生地质灾害的条件。中易发区主要分布在高易发区两侧,地质灾害发育较少。

表 3 逻辑回归分析结果
Table 3 Results of logistic regression analysis

回归项	<i>B</i>	<i>SE</i>	<i>wals</i>	<i>df</i>	<i>Sig</i>
高程	1.993	0.229	75.665	1	0.000
坡度	1.283	0.305	17.641	1	0.000
坡向	1.338	0.224	35.719	1	0.000
距河流距离	1.539	0.229	45.258	1	0.000
距断层距离	0.281	0.294	0.914	1	0.000
工程岩组	0.858	0.303	8.008	1	0.005
常量	−0.663	0.126	27.818	1	0.000

表 4 地质灾害易发性分区统计表
Table 4 Results of geological hazard susceptibility in different districts

易发性分级	面积/km ²	占研究区面积比例/%	灾害点个数/个	占灾害点总数比例/%	地灾密度
极低易发区	516.611	20.099	3	0.719	0.006
低易发区	589.298	22.927	6	1.439	0.010
中易发区	486.413	18.924	36	8.633	0.074
高易发区	606.095	23.580	156	37.410	0.257
极高易发区	371.951	14.471	216	51.799	0.581

5.2 精度分析

ROC 曲线常被用来对易发性评价实验结果做定性分析,AUC 曲线常被用来度量分析模型的有效性,其大小代表曲线下方的面积,AUC 值越大,表明模型的拟合度越高(钟宙锋,2019)。本文用 ROC 曲线和 AUC 对该模型进行验证,ROC 曲线横轴表示易发性面积比的累积量,纵轴表示灾害点面积百分比的累积量,结果如上图 5 所示。验证结果 AUC 值为 0.863,表明 CF-Logistic 回归模型能够较为客观准确地对湟中县地质灾害易发性进行评价。

6 结论

本文以青海省湟中县为例,通过分别计算高程、坡度、坡向、地形起伏度、工程岩组等地质灾害影响因子的确定性系数,并利用 Logistic 回归模型得到地质灾害发生的概率,以此为基础,对研究区进行地质灾害易发性评价,主要结论有:

- (1)本文结合 CF 模型和 Logistic 回归模型对湟中县地质灾害易发性进行评价。确定性系数模型和逻辑回归模型的结合能够实现复杂多因子数据同区间的定量化,解决各因子之间存在的共线性问题,而且能够客观地计算出各因子的权重和各因子分类级别的权重,具有较强的客观性。
- (2)通过 SPSS 和 ArcGIS 计算,剔除地形起伏

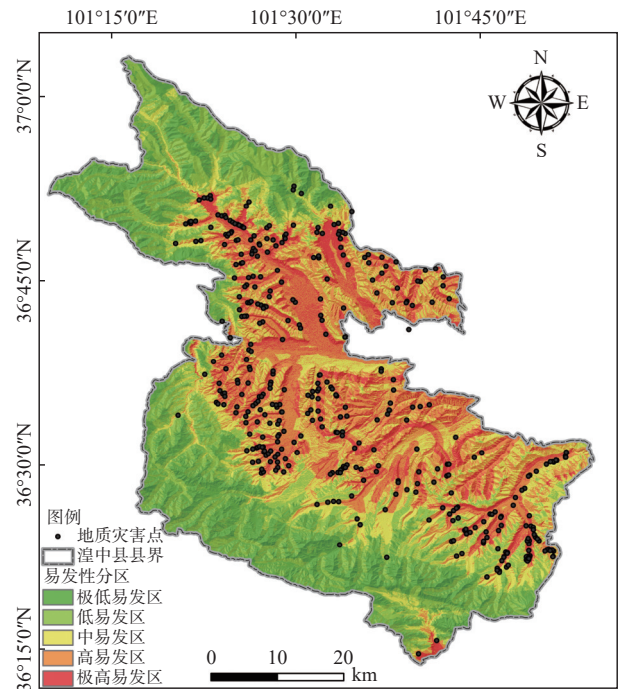


图 4 地质灾害易发性评价图
Fig. 4 Assessment chart of geological hazard susceptibility

度因子,消除了因子共线性。从剩余的 6 个地质灾害影响因子的各分类级别的 CF 值和各因子的逻辑回归系数可以看出,高程、距河流距离、坡向和工程岩组对湟中县地质灾害易发程度贡献相对较大。特别是在海拔 2 560~2 750 m 内、坡向为东、

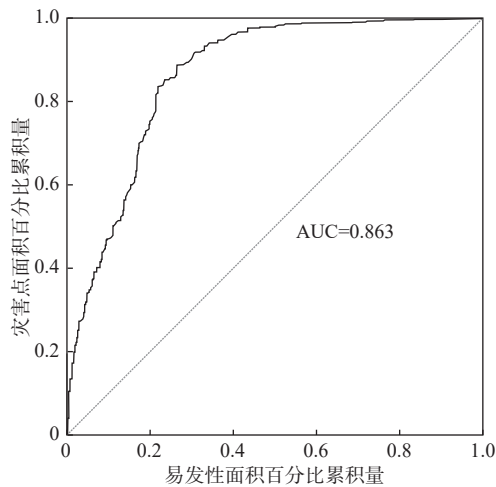


图5 逻辑回归模型的 ROC 曲线

Fig. 5 ROC curve of Logistic Regression model

东南方向、距河流 500 m、松散层状冲洪积岩组和软弱层状碎屑岩岩组是最容易发生地质灾害的区域。

(3)从各易发程度分布状况来看,湟中县地质灾害高易发区和极高易发区主要分布在湟水河干流南北两岸及其较大支流两侧的低山丘陵地区,面积为 978.046 km²,占研究区总面积的 38.051%。低易发区和极低易发区主要分布在北部和西南高海拔地区,面积为 1 105.909 km²,占研究区总面积的 43.025%。

(4)基于 CF-Logistic 回归模型的易发性评价方法可以客观准确地计算出地质灾害影响因子与地质灾害的发生之间的确定性关系。而 ROC 曲线和 AUC 值则说明该方法可以较为准确地对湟中县地质灾害进行易发性评价。

References

- Heckerman D. Probabilistic Interpretation for MYCIN's Certainty Factors[C]. Proceedings UAI, 1985.
- Shortliffe E H, Buchanan B G. A model of inexact reasoning in medicine[J]. Mathematical Biosciences, 1975, 23: 351–379.
- Chen X Y, Li M H, Wang D W, et al., 2019. Quantitative evaluation of geohazards susceptibility based on GIS and information value model for Emeishan City, Sichuan. Sedimentary Geology and Tethyan Geology, 39 (4): 100–112.
- Cheng W M, Zhou C H, Chai H X, et al., 2009. Quantitative extraction and analysis of basic morphological types of land geomorphology in China[J]. Journal of Geo-Information Science, 11 (6): 725–736 (in Chinese with English abstract).
- Feng X K, 2014. Risk assessment of earthquake-induced geological hazards based GIS -A Case Study of Yiliang County. Beijing: Institute of Geology, China Earthquake Administration.
- Guo F F, Yang N, Meng H, et al., 2008. Application of the relief amplitude and slope analysis to regional landslide hazard assessments[J]. Geology in China, 35 (1): 131–143 (in Chinese with English abstract).
- He P, Tong L Q, Guo Z C, et al., 2016. Evaluation research on the landslide disaster liability in Zha Da Region of Tibet[J]. Science Technology and Engineering, 16 (25): 193–200 (in Chinese with English abstract).
- Hong Z L, Li Y H, Zhang L Y, et al., 2020. A method of regional geological hazard assessment based on principle component analysis[J]. Journal of Catastrophology, 35 (1): 118–124 (in Chinese with English abstract).
- Hu K H, Cui F, Han Y S, et al., 2012. Susceptibility mapping of landslides and debris flows in 2008 Wenchuan earthquake by using cluster analysis and maximum likelihood classification methods. Science of Soil and Water Conservation, 10 (1): 12–18.
- Huang R Q, Xu X N, Tang C, et al., 2008. Assessment of the Geological Environment and Geo-hazard Management. Beijing: Science Press.
- Li X, 2009. Development mechanism of slope geological disasters and risk assessment of slopes in TaiShan. Guangzhou: Zhongshan University.
- Liu Y F, Fang J L, Chen X H, et al., 2014. Evaluation of landslide susceptibility in Zigui County based on certainty factor method[J]. Journal of Natural Disasters, 23 (6): 209–217 (in Chinese with English abstract).
- Rao P Z, Cao R, Jiang W G, 2017. Susceptibility evaluation of geological disasters in Yunnan Province based on geographically weighted regression model[J]. Journal of Natural Disasters, 26 (2): 134–143 (in Chinese with English abstract).
- Shen L L, Liu L Y, Xu C, et al., 2016. Multi-models based landslide susceptibility evaluation illustrated with landslides triggered by Minxian earthquake[J]. Journal of Engineering Geology, 24 (1): 19–28 (in Chinese with English abstract).
- Qin Y G, Yang G L, Jiang X Y, et al., 2020. Geohazard susceptibility assessment based on the integrated model of certainty factor model and logistic regression model: a case study in Kaiyang County, Guizhou Province[J]. Science Technology and Engineering, 20 (1): 96–103 (in Chinese with English abstract).
- Tang G A, Li F Y, Liu X J, 2016. Digital Elevation Model Tutorial[M]. 3rd ed. Beijing: Science Press.
- Tian C S, Liu X L, Wang J, 2016. Geohazard susceptibility assessment based on CF model and logistic regression models in Guangdong[J]. Hydrogeology & Engineering Geology, 43 (6): 154–161 (in Chinese with English abstract).
- Xiang X Q, Huang R Q, 2000. Application of GIS-based artificial neural networks on assessment of geohazards risk. The Chinese Journal of Geological Hazard and Control, 11 (3): 23–27.
- Yang D H, Fan W, 2015. Zoning of probable occurrence level of geological disasters based on ArcGIS: a case of Xunyang. The Chinese Journal of Geological Hazard and Control, 26 (4): 82–86.

- Zhao Y N, Niu R Q, 2010. Exploration of Landslide Hazard Zonation Based on the Weights of Evidence Method. *Geography and Geo-Information Science*, 26 (6): 19–23.
- Zhong M F, 2019. Geological Hazard Susceptibility Evaluation Based on Certainty Factor and Logistic Regression Model. *World Nonferrous Metals*, (23): 186–190.

附中文参考文献

- 陈绪钰, 李明辉, 王德伟, 等, 2019. 基于 GIS 和信息量法的四川峨眉山市地质灾害易发性定量评价[J]. *沉积与特提斯地质*, 39 (4): 100–112.
- 程维明, 周成虎, 柴慧霞, 等, 2009. 中国陆地地貌基本形态类型定量提取与分析[J]. *地球信息科学学报*, 11 (6): 725–736.
- 冯新科, 2014. 基于 GIS 的地震地质灾害危险性评价: 以彝良县为例[D]. 北京: 中国地震局地质研究所.
- 郭芳芳, 杨农, 孟晖, 等, 2008. 地形起伏度和坡度分析在区域滑坡灾害评价中的应用[J]. *中国地质*, 35 (1): 131–143.
- 贺鹏, 童立强, 郭兆成, 等, 2016. GIS 支持下基于层次分析法的西藏札达地区滑坡灾害易发性评价研究[J]. *科学技术与工程*, 16 (25): 193–200.
- 洪增林, 李永红, 张玲玉, 等, 2020. 一种基于主成分分析法的区域性地质灾害危险性评估方法[J]. *灾害学*, 35 (1): 118–124.
- 胡凯衡, 崔鹏, 韩用顺, 等, 2012. 基于聚类和最大似然法的汶川灾区泥石流滑坡易发性评价[J]. *中国水土保持科学*, 12–18.
- 黄润秋, 许向宁, 唐川, 等, 2008. 地质环境评价与地质灾害管理[M]. 北京: 科学出版社.
- 李鑫, 2009. 台山市崩滑地质灾害风险评价与风险管理研究[D]. 中山大学.
- 刘艳芳, 方佳琳, 陈晓慧, 等, 2014. 基于确定性系数分析方法的秭归县滑坡易发性评价[J]. *自然灾害学报*, 23 (6): 209–217.
- 饶品增, 曹冉, 蒋卫国, 2017. 基于地理加权回归模型的云南省地质灾害易发性评价[J]. *自然灾害学报*, 26 (2): 134–143.
- 沈玲玲, 刘连友, 许冲, 等, 2016. 基于多模型的滑坡易发性评价——以甘肃岷县地震滑坡为例[J]. *工程地质学报*, 24 (1): 19–28.
- 覃乙根, 杨根兰, 江兴元, 等, 2020. 基于确定性系数模型与逻辑回归模型耦合的地质灾害易发性评价——以贵州省开阳县为例[J]. *科学技术与工程*, 20 (1): 96–103.
- 汤国安, 李发源, 刘学军, 2016. 数字高程模型教程 (第3版) [M]. 北京: 科学出版社.
- 田春山, 刘希林, 汪佳, 2016. 基于 CF 和 Logistic 回归模型的广东省地质灾害易发性评价[J]. *水文地质工程地质*, 43 (6): 154–161.
- 向喜琼, 黄润秋, 2000. 基于 GIS 的人工神经网络模型在地质灾害危险性区划中的应用[J]. *中国地质灾害与防治学报*, 11 (3): 23–27.
- 杨德宏, 范文, 2015. 基于 ARCGIS 的地质灾害易发性分区评价——以旬阳县为例[J]. *中国地质灾害与防治学报*, 26 (4): 82–86.
- 赵艳南, 牛瑞卿, 2010. 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*, 26 (6): 19–23.
- 钟亩锋, 2019. 基于确定性系数与 Logistic 回归模型的地质灾害易发性评价[J]. *世界有色金属*, (23): 186–190.