

doi: 10.19388/j.zgdzdc.2023.06.11

引用格式: 冯寿娟. 基于网格化 Logistic 回归模型的区域突发环境事件应急决策方法[J]. 中国地质调查, 2023, 10(6): 95-100. (Feng S J. Emergency decision method of regional environmental emergencies based on gridding logistic regression model [J]. Geological Survey of China, 2023, 10(6): 95-100.)

基于网格化 Logistic 回归模型的区域突发环境事件应急决策方法

冯寿娟

(河北德禹环保科技有限公司, 河北唐山 064400)

摘要: 针对突发环境事件不确定性较高, 应急决策方案复杂, 实行成本较高的问题, 提出一种基于网格化 Logistic 回归模型的区域突发环境事件应急决策方法。利用德尔菲方法选取区域突发环境事件影响变量, 以变量为输入, 构建基于 Logistic 回归模型的区域突发环境事件发生概率计算模型, 将研究区域划分为单元格形式, 划分区域等级, 明确安全区范围。在多个约束条件下, 利用蚁群算法求取目标函数, 得出最优人员疏散路径, 完成应急决策方案设计。实验结果表明: 该方法能够有效应对突发环境事件, 且从建模至完成人员疏散的时间约 0.5 h, 总耗时较短, 实现了区域突发环境事件应急决策, 从而保障了人员和财产安全。研究对区域突发环境事件的防控具有指导意义。

关键词: 网格化; Logistic 回归模型; 区域突发环境事件; 疏散路径; 应急决策

中图分类号: TP214.33 **文献标志码:** A **文章编号:** 2095-8706(2023)06-0095-06

0 引言

突发环境事件是指由于环境污染、自然灾害等因素的影响而突然发生的且造成重大损失的事件。近年来突发环境事件频繁发生, 造成了严重的人员伤亡和财产损失。与一般环境事件相比, 突发性环境事件重点在于“突发”两个字, 不确定性更高, 造成的损失和伤亡也更大^[1-2]。在此背景下, 如何进行有效地应急防控成为当下研究的重点。

为解决上述问题, 很多专家和学者都进行了相关研究。例如: 魏斌等^[3]以甘肃省为例, 总结了 2011—2019 年期间该地区发生的 80 多起事件的起因, 并对这些事件的时空特征进行了分析, 最后以污染物特征为参考提出了具体的防控策略; 张飞等^[4]以江苏工业园区为研究区, 提取该地区的 3 个一级指标及 25 个二级指标, 以这些指标为输入, 利用层次分析法建立了评估模型, 对该地区突发环

境事件风险分级进行了划分; 王思等^[5]在其研究中以珠三角某石化区及周边区域为例, 通过网格化技术将该区域划分为若干基本单元, 然后计算环境风险受体易损性, 确定了网格风险值, 完成了风险区划分, 确定了重点防控区域, 为差异化管理提供参考。结合前人研究成果, 本文提出一种基于网格化 Logistic 回归模型的区域突发环境事件应急决策方法, 以为突发环境事件的处理提供参考, 降低突发事件带来的损失和伤亡。

1 应急决策方案

1.1 影响变量的选择

影响变量, 即影响因子, 也可以说是诱因, 在影响变量的综合作用下, 会导致区域突发环境事件的发生, 因此, 要想确定区域突发环境事件的发生概率, 划分风险等级区域, 确定影响变量是十分重要的^[6-7]。

基于上述分析,本文利用德尔菲方法来进行影响变量的选择。德尔菲方法是一种通过计算3个专家系数而进行变量筛选的方法。该方法经过多轮的征询,计算专家意见协调程度^[8]。

专家积极系数计算公式为

$$\mu_i = \frac{\sqrt{\sum_{k=1}^m x_{ki}}}{m} \quad (1)$$

式中: μ_i 为第 i 个影响变量的专家积极系数; x_{ki} 为第 k 个专家对第 i 个影响变量的评分; m 为专家人数,个; n 为影响变量的数量,个。

专家意见权威系数计算公式为

$$\eta_i = \frac{\sum_{k=1}^m (x_{ki} - M_i)^2}{m - 1} \quad (2)$$

式中: η_i 为第 i 个影响变量的第 k 个专家意见权威系数; M_i 为第 i 个影响变量的经验性评分结果。

专家意见协调系数计算公式为

$$Y_i = \frac{\eta_i}{\mu_i} \quad (3)$$

式中: Y_i 代表第 i 个影响变量的专家意见协调系数。其他符号含义同式(1)、(2)。

当 $\mu_i \geq 0.7$, $\eta_i \leq 0.25$ 以及 $Y_i < 0.5$ 时,认为该影响变量是重要的^[9]。

选取出来的指标由于量纲不同,不能用于同一个模型的输入,因此需要进行标准化处理,处理公式为

$$x'_i = \frac{x_i - \min_{1 \leq i \leq n} x}{\max_{1 \leq i \leq n} x - \min_{1 \leq i \leq n} x} \quad (4)$$

式中: x'_i 为标准化处理后的第 i 个影响变量; x_i 为原始影响变量; $\max_{1 \leq i \leq n} x$ 、 $\min_{1 \leq i \leq n} x$ 为影响变量中的最大和最小值。

1.2 事件发生概率计算与安全区划定

Logistic 回归模型是一种定量分析事件发生概率的大小计算模型。Logistic 回归模型是由 logit 变换 + 几率计算公式演化而成^[10]。

logit 变换模型为

$$\log(y) = \zeta_0 + \zeta_1 x_1 + \zeta_2 x_2 + \dots + \zeta_n x_n \quad (5)$$

式中: y 为事件发生概率; $\zeta_0, \zeta_1, \zeta_2, \dots, \zeta_n$ 代表模型的参数,即逻辑回归系数; x_1, x_2, \dots, x_n 为影响变量。

几率是指事件发生的概率与不发生的概率之比。假设一个事件发生的概率为 p , 则相应的不发

生的概率则为 $1 - p$, 那么这个事件的整体几率 y 就可以通过下述公式计算出来^[11], 即

$$y = \frac{p}{1 - p} \quad (6)$$

概率和几率是对等关系, 因此, 由上述 logit 变换模型和几率表达式构建 Logistic 回归模型。根据 1.1 节选出的变量, 采集研究区的历史变量数据, 以此作为输入, 构建基于 Logistic 回归模型的事件发生概率计算模型^[12]。该模型原始表达式为

$$\log\left(\frac{p}{1 - p}\right) = \zeta_0 + \zeta_1 x_1 + \zeta_2 x_2 + \dots + \zeta_n x_n \quad (7)$$

式中: 各符号含义同式(5)(6)。

将式(7)进行转换, 即可得到区域突发环境事件发生概率 p 的计算公式, 即

$$p = \frac{e^{\zeta_0 + \zeta_1 x_1 + \zeta_2 x_2 + \dots + \zeta_n x_n}}{1 + e^{\zeta_0 + \zeta_1 x_1 + \zeta_2 x_2 + \dots + \zeta_n x_n}} \quad (8)$$

式中: 各符号含义同式(7)。

在式(7)中, 由于逻辑回归系数 $\zeta_0, \zeta_1, \zeta_2, \dots, \zeta_n$ 是不确定的, 因此模型并不能直接应用, 还需要从研究区突发环境事件历史资料中提取变量对应的数据及其结果代入式(8)中。利用 SPSS 软件来反演求解各个逻辑回归系数, 使得 $\zeta_0, \zeta_1, \zeta_2, \dots, \zeta_n$ 已知^[13]。最后进行两种检验, 即对回归系数进行显著性检验, 对回归模型进行拟合度检验, 从而保证模型的合理性。

1 km × 1 km 的单元网格可提供合理的精度, 揭示研究区的空间变化情况, 同时也可以保持计算复杂度在可接受范围内。在城市规划、自然灾害风险评估中, 1 km × 1 km 的单元网格尺度已经被证明在环境风险评估中具有较好的适用性。因此, 将研究区划分为 1 km × 1 km 的单元网格, 然后利用式(8)计算每一个单元区域的突发环境事件的发生概率^[14-15]。发生概率取值为 0 ~ 1, 为提高可操作性, 减少分类和决策选择, 将突发环境事件发生概率划分为 3 个等级(表 1)。

表 1 突发环境事件发生概率等级划分

Tab.1 Classification of occurrence probability for environmental emergencies

风险等级	概率
高风险	[0.8, 1.0]
中风险	[0.6, 0.8)
低风险	[0.0, 0.6)

在低风险区中划分安全区,当需要进行人员疏散或者撤离时,将这些安全区作为目的地,规划应急撤离方案。

1.3 应急决策方案设计

以安全区为参考,确定为可供选择的安全地点,规划区域突发环境事件应急决策方案,即确定人员疏散路径。应急决策方案设计流程如下。

- (1)对疏散问题作出基本假设,确定安全点。
- (2)建立以最短距离为目标的函数模型。该模型可描述为

$$\min L = \sum_{i=1}^m \sum_{j=1}^n H_{ij} G_{ij} Z_{ij} \quad (9)$$

其中

$$H_{ij} = w \cdot l_{ij} \quad (10)$$

$$Z_{ij} = \frac{R_{ij} - \log p_{\min}}{\log p_{\max} - \log p_{\min}} \quad (11)$$

式中: $\min L$ 为最短疏散路径距离, m ; H_{ij} 为待疏散人群集中点 i 至安全点 j 的当量长度, m ; G_{ij} 为需要从点 i 疏散到点 j 的人数, 个; Z_{ij} 为从点 i 疏散到点 j 的道路阻断风险效用函数; R_{ij} 为从点 i 疏散到点 j 的通行概率; p_{\min} 、 p_{\max} 为疏散网络的最小通行概率和最大通行概率; m 为待疏散人群集中点数量, 个; n 为安全点数量, 个; w 为不同区域的危险系数, 高风险取值 1.0, 中风险取值 0.5, 低风险取值 0.2; l_{ij} 为待疏散人群集中点 i 至安全点 j 的疏散路线距离, m 。

- (3)设置约束条件。计算公式如下

$$\sum_{i=1}^m G_{ij} \leq A_j \quad (12)$$

$$\sum_{j=1}^n G_{ij} \leq B_i \quad (13)$$

$$G_{ij} \geq 0 \quad (14)$$

式中: A_j 为待疏散人群集中点需要疏散人数, 个; B_i 为各安全点容量。

- (4)利用寻优算法,如粒子群算法进行求解,得出最优人员疏散路径,完成应急决策方案设计^[16-17]。

基于粒子群算法计算变异粒子 ζ 的适应值,则有

$$f(\zeta) = \frac{\sum_{l=1}^{|\zeta|} f(\zeta_l)}{|\zeta|} \quad (15)$$

式中: $f(\zeta)$ 为 ζ 中所有疏散路径距离的平均适应值, m ; $l=1, 2, \dots, |\zeta|$, 为第 l 个疏散路径; $|\zeta|$ 为 ζ 中的疏散路径数量, 个。

对适应度低的粒子进行替换,选择出适应度最优的个体粒子 L_1 及全局粒子 L_2 。

首先对 L_1 和 L_2 的位置进行更新,则有

$$V_i^{g+1} = V_i^g + v_1 \varepsilon_1 (L_1 - a_i^g) + v_2 \varepsilon_2 (L_2 - a_i^g) \\ a_i^{g+1} = a_i^g + V_i^{g+1} \quad (16)$$

式中: V_i^{g+1} 为粒子更新速度, m/s ; a_i^{g+1} 为粒子更新后位置, m ; V_i^g 为粒子在 V_i^{g+1} 前的更新速度, m/s ; v_1 为 L_1 的加速因子; v_2 为 L_2 的加速因子; ε_1 、 ε_2 均为 0~1 之间的随机数; a_i^g 为粒子更新前位置, m 。

根据适应度值判断粒子迭代是否停止。若适应度值为最优,则停止迭代,输出最优疏散路径和路径距离,否则重复更新粒子,直至满足停止迭代条件。

通过上述过程,完成了基于网格化 Logistic 回归模型的区域突发环境事件应急决策方法研究。

2 算例分析与测试

2.1 研究区概况

假设研究区内某化工园区中一化工厂发生毒气泄漏问题(图1),引发区域突发环境事件,需要进行应急决策,将人员疏散至安全区。

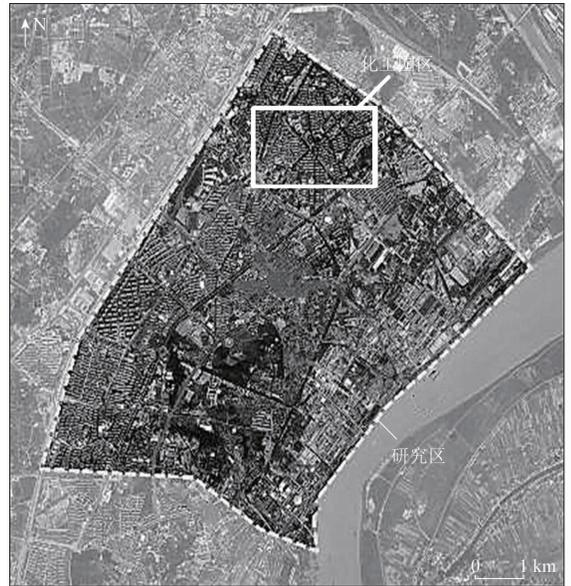


图1 研究区示意图

Fig.1 Sketch of the study area

2.2 影响变量选取结果

基于 1.1 节的研究,选取该研究区内突发环境事件的影响变量,建立变量体系,如表 2 所示。

表 2 影响变量选取结果

Tab.2 Results of the impact variable selection

类别	影响变量
环境风险源强度	研究区化工企业数量
	危险废品的产生强度
	风险物质总量
	5 年来发生突发环境事件的频率
环境风险受体脆弱性	居民区污染风向频率
	常住人口数量
	人均 GDP 水平
	流通渠道类别
环境风险防控与应急能力	环境风险受体数量
	监测预警能力
	应急人员数量
	物资储备情况
	应急预案的演练情况

从表 2 可以看出,总共选出 13 个变量,作为回归模型的输入指标。

2.3 回归系数计算及模型构建

基于表 2 变量,采集该区域历史数据,以此作为样本,输入公式(8)中,计算得到回归方程回归系数,并进行检验,计算得出 Logistic 回归系数,结果如表 3 所示。

表 3 Logistic 回归系数

Tab.3 Logistic regression coefficient

名称	参数值
ζ_0	12.588 2
ζ_1	1.584 3
ζ_2	0.221 4
ζ_3	-1.256 5
ζ_4	0.268 7
ζ_5	0.425 5
ζ_6	0.123 6
ζ_7	3.555 5
ζ_8	3.201 1
ζ_9	-1.254 8
ζ_{10}	-3.545 4
ζ_{11}	0.558 2
ζ_{12}	0.745 5
ζ_{13}	2.574 8

将表 3 中的 Logistic 回归系数对应地代入到公式(8)中,完成基于 Logistic 的突发环境事件发生概率计算模型的构建。

2.4 单元网格划分结果

假设研究区内承灾体的性质、规模和易损性均一致,利用 ArcGIS 的 create fishnet 模块对图 1 研究区进行单元网格划分,计算每一个网格的突发环境

事件发生概率,并划分风险等级,获取低风险区,结果如图 2 所示。



图 2 单元网格划分

Fig.2 Cell grid division

2.5 应急决策方案

选取蚁群算法作为求取算法,求取最优人员疏散路径,结果如表 4 所示。

表 4 蚁群算法基本参数设置

Tab.4 Basic parameter setting of the ant colony algorithm

参数	参数值
种群的规模/个	10
最大迭代次数/次	1 000
信息素重要程度因子	1.0
启发函数重要程度因子	2.0
信息素浓度常数	5.0
信息素初始值	0.000 2
全局信息素挥发系数	0.1
转移概率	0.5

在表 4 给定参数设置下,蚁群算法可在不足 10 min 的时间内找到满意的解决方案。

在公式(12)~(14)等约束条件下,利用蚁群算法求取公式(9)目标函数,得出最优人员疏散路径,完成应急决策方案设计,如图 3 所示。

安全区的选择不仅要离待疏散人群集中点较近且具有足够容量,而且要考虑其可达性和逃生路径的畅通程度。共有 3 条最优人员疏散路径,分别对应 3 个安全区作为人员集中点,其中:路径 1 总

长度为 2 365.48 m, 人员疏散预计花费 25.47 min 完成; 路径 2 总长度为 1 687.95 m, 人员疏散预计花费 18.25 min 完成; 路径 3 总长度为 1 892.35 m, 人员疏散预计花费 20.40 min 完成。

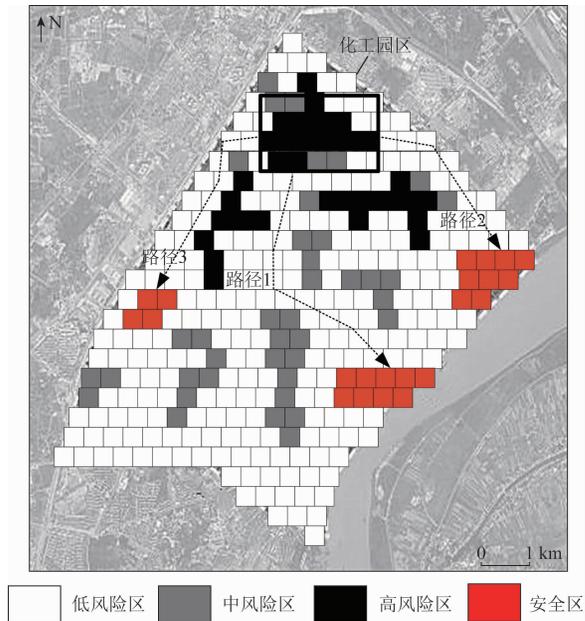


图 3 最优人员疏散路径

Fig. 3 The optimal evacuation route

3 结论

(1) 利用 ArcGIS 的 create fishnet 模块进行单元格划分, 通过计算每一个网格的突发环境事件发生概率, 确定风险等级, 可有效获取安全区。

(2) 通过蚁群算法, 可在 10 min 内规划 3 条最优人员疏散路径, 每条路径的人员疏散预计时间较短, 均未超过 26 min, 有效完成了应急决策方案的制定。

(3) 研究仍需进一步改进, 如在突发因素的考虑上, 疏散过程中可能遇到原本可以通行的路径突然无法通行的情况, 在路径设计上, 要考虑此种情况的出现, 进行优化。

参考文献 (References):

[1] 刘红岩, 张小趁, 李俊峰, 等. 基于突发地质灾害应急防治的边(滑)坡分类方法[J]. 防灾减灾工程学报, 2021, 41(2): 193-202, 237.
Liu H Y, Zhang X C, Li J F, et al. Slope (Landslide) classification method based on emergency prevention of sudden geological disaster[J]. Journal of Disaster Prevention and Mitigation Engi-

neering, 2021, 41(2): 193-202, 237.
[2] 方然可, 刘艳辉, 苏永超, 等. 基于逻辑回归的四川青川县区域滑坡灾害预警模型[J]. 水文地质工程地质, 2021, 48(1): 181-187.
Fang R K, Liu Y H, Su Y C, et al. A early warning model of regional landslide in Qingchuan County, Sichuan Province based on logistic regression [J]. Hydrogeology & Engineering Geology, 2021, 48(1): 181-187.
[3] 魏斌, 梁佳, 薛丽洋, 等. 甘肃省 2011-2019 年突发环境事件特征及防控对策[J]. 西北师范大学学报: 自然科学版, 2021, 57(6): 124-130.
Wei B, Liang J, Xue L Y, et al. Characteristics of environmental emergencies in Gansu Province from 2011 to 2019 and prevention and control countermeasures [J]. Journal of Northwest Normal University (Natural Science Edition), 2021, 57(6): 124-130.
[4] 张以飞, 付丽洋, 毕军, 等. 江苏工业园区突发环境事件风险评估体系研究[J]. 环境污染与防治, 2020, 42(11): 1430-1435.
Zhang Y F, Fu L Y, Bi J, et al. Research on the risk assessment system of environmental emergencies in industrial parks of Jiangsu Province [J]. Environmental Pollution & Control, 2020, 42(11): 1430-1435.
[5] 王思, 张志娇, 余璇, 等. 基于网格化的区域突发环境事件风险分区研究[J]. 环境科学与技术, 2021, 44(5): 220-228.
Wang S, Zhang Z J, Yu X, et al. Risk zoning of regional environmental emergency based on gridded method [J]. Environmental Science & Technology, 2021, 44(5): 220-228.
[6] 陈朝亮, 彭树宏, 钱静, 等. 基于 AHP-Logistic 熵权模型的西南浅丘区地质灾害分布特征研究——以内江市为例[J]. 长江科学院院报, 2020, 37(2): 55-61.
Chen C L, Peng S H, Qian J, et al. Distribution characteristics of geological hazards in southwestern shallow hill based on AHP-logistic entropy combined weight model: A case study of Neijiang City [J]. Journal of Yangtze River Scientific Research Institute, 2020, 37(2): 55-61.
[7] 许强, 彭大雷, 何朝阳, 等. 突发型黄土滑坡监测预警理论方法研究——以甘肃黑方台为例[J]. 工程地质学报, 2020, 28(1): 111-121.
Xu Q, Peng D L, He C Y, et al. Theory and method of monitoring and early warning for sudden loess landslide: A case study of Heifangtai Terrace [J]. Journal of Engineering Geology, 2020, 28(1): 111-121.
[8] 周定义, 左小清, 喜文飞, 等. 联合 SBAS-InSAR 和 PSO-BP 算法的高山峡谷区地质灾害危险性评价[J]. 农业工程学报, 2021, 37(23): 108-116.
Zhou D Y, Zuo X Q, Xi W F, et al. Combined SBAS-InSAR and PSO-BP algorithm for evaluating the risk of geological disasters in alpine valley regions [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(23): 108-116.
[9] 杨顺, 常昆瑜, 李孝波, 等. 基于 Logistic 回归模型的黄土斜坡地震稳定性快速评判方法[J]. 地震工程学报, 2020, 42(2): 512-516.

Yang S, Chang C Y, Li X B, et al. Quick assessment of the seismic stability of loess slopes based on a logistic regression model[J]. *China Earthquake Engineering Journal*, 2020, 42(2) :512 – 516.

[10] 覃乙根, 杨根兰, 江兴元, 等. 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*, 2020, 20(1) :96 – 103.

Qin Y G, Yang G L, Jiang X Y, et al. Geohazard susceptibility assessment based on integrated certainty factor model and logistic regression model for Kaiyang, China[J]. *Science Technology and Engineering*, 2020, 20(1) :96 – 103.

[11] 罗路广, 裴向军, 黄润秋, 等. GIS支持下CF与Logistic回归模型耦合的九寨沟景区滑坡易发性评价[J]. *工程地质学报*, 2021, 29(2) :526 – 535.

Luo L G, Pei X J, Huang R Q, et al. Landslide susceptibility assessment in Jiuzhaigou scenic area with GIS based on certainty factor and logistic regression model[J]. *Journal of Engineering Geology*, 2021, 29(2) :526 – 535.

[12] 李利峰, 张晓虎, 邓慧琳, 等. 基于熵指数与逻辑回归耦合模型的滑坡灾害易发性评价——以蓝田县为例[J]. *科学技术与工程*, 2020, 20(14) :5536 – 5543.

Li L F, Zhang X H, Deng H L, et al. Assessment of landslide susceptibility based on coupling model of index of entropy and logistic regression: A case study of Lantian County[J]. *Science Technology and Engineering*, 2020, 20(14) :5536 – 5543.

[13] 田钦, 张彪, 郭建飞, 等. 基于信息量和逻辑回归耦合模型的滑坡易发性评价[J]. *科学技术与工程*, 2020, 20(21) :8460 – 8468.

Tian Q, Zhang B, Guo J F, et al. Landslide susceptibility assessment based on the coupling model of information value and logistic regression [J]. *Science Technology and Engineering*, 2020, 20(21) :8460 – 8468.

[14] 龚凌枫, 徐伟, 铁永波, 等. 基于数值模拟的城镇地质灾害危险性评价方法[J]. *中国地质调查*, 2022, 9(4) :82 – 91.

Gong L F, Xu W, Tie Y B, et al. Risk assessment on urban geological disaster based on numerical calculation[J]. *Geological Survey of China*, 2022, 9(4) :82 – 91.

[15] 铁永波, 孙强, 徐勇, 等. 南方山地丘陵区典型地质灾害成因机制与风险评价[J]. *中国地质调查*, 2022, 9(4) :1 – 9.

Tie Y B, Sun Q, Xu Y, et al. Genetic mechanism and risk assessment of typical geological hazards in mountainous and hilly areas of South China[J]. *Geological Survey of China*, 2022, 9(4) :1 – 9.

[16] 李怡静, 胡奇超, 刘华赞, 等. 耦合信息量和 Logistic 回归模型的滑坡易发性评价[J]. *人民长江*, 2021, 52(6) :95 – 102.

Li Y J, Hu Q C, Liu H Z, et al. Evaluation of landslide susceptibility based on coupling model of information value – logistic regression[J]. *Yangtze River*, 2021, 52(6) :95 – 102.

[17] 莫家乐, 叶脉, 解光武, 等. 突发环境事件应急演练场景设计探索[J]. *四川环境*, 2020, 39(4) :161 – 166.

Mo J L, Ye M, Xie G W, et al. Exploration on the scenario design of emergency drill for environmental emergencies[J]. *Sichuan Environment*, 2020, 39(4) :161 – 166.

Emergency decision method of regional environmental emergencies based on gridding logistic regression model

FENG Shoujuan

(Hebei Deyu Environmental Protection Technology Co., Ltd., Hebei Tangshan 064400, China)

Abstract: In view of the high uncertainty, complex emergency decision – making scheme and high implementation cost of environmental emergencies, the authors in this research proposed a emergency decision method for regional environmental emergencies based on grid logistic regression model. The Delphi method is used to select the impact variables of regional environmental emergencies, and a probability calculation model of regional environmental emergencies based on logistic regression model is constructed, with the variables as input. The study area was divided into cell forms, and the regional grades were divided to define the scope of the safety area. Under multiple constraints, the objective function was obtained by using the ant colony algorithm, and the optimal evacuation route was obtained to complete the emergency decision scheme design. The experimental results show that this research method is effective for environmental emergencies, and the duration between model establishment and people evacuation is about half an hour. This short time – consuming means the emergency decision is effective for regional environmental emergencies, which would ensure the safety of personnel and property. This research has guiding significance for the prevention and control of regional environmental emergencies.

Keywords: gridding; logistic regression model; regional environmental emergencies; evacuation route; emergency decision