

基于 Logistic 模型的城市土地抗震适宜性 评价方法研究

袁 媛, 陈新民, 袁宗盼

(南京工业大学交通学院, 南京 210009)

[摘要] 城市土地抗震适宜性评价可以通过土地发生地震灾害的概率大小来划分等级,而土地地震灾害的发生与否,是典型的二分类变量问题。运用 Logistic 回归模型处理这类(0,1)区间的概率问题,可以得到较理想的拟合效果,为城市土地抗震适宜性评价提供较合理的量化指标。而进行 Logistic 回归分析需要较大的样本数据,原始资料数据显然不能满足分析要求,这就需要利用 GIS 的空间分析功能进行空间插值,得到足够的样本数据,运用 SPSS 数值运算软件,进行二分 Logistic 回归分析。通过得到的城市土地抗震适宜性的量化评价方法,以南京市为例,分析南京市具体地质情况,得到评价体系,并对评价指标赋值,通过运用 Logistic 模型得到评价指标权重和城市土地地震灾害概率公式,利用计算得到的地震灾害概率,实现定量化评价土地抗震适宜性的目的。

[关键词] 抗震适宜性; Logistic 回归模型; GIS; SPSS

中图分类号: TU435 文献标识码: A 文章编号: 1002-848X(2013) S1-1479-04

Research on evaluation method for urban land aseismic suitability based on the logistic model

Yuan Yuan, Chen Xinmin, Yuan Zongpan

(College of Transportation, Nanjing University of Technology, Nanjing 210009, China)

Abstract: Evaluation on urban land seismic suitability can be graded by the probability of earthquake disaster. Earthquake disaster of land occur or not is a typical problem of binary variables. Using logistic regression model to deal with this kind of (0,1) inter probability problems can get ideal fitting effect, which can provide more reasonable quantitative indicators about evaluation on urban land aseismic suitability. Logistic regression analysis requires a large sample data. The original data obviously can not meet the requirements for analysis, which requires the use of spatial analysis capabilities of GIS spatial interpolation to obtain a sufficient data. SPSS numerical computing software is used to analyze binary logistic regression model. The quantitative evaluation method of urban land seismic suitability can be used with Nanjing as a example. Firstly, the evaluation system and evaluation index assignment of Nanjing can be get through studying the specific geological conditions of Nanjing. Secondly, the weight of evaluation indicators and the formula of the probability of the earthquake disaster on urban land can be obtained by using logistic model. Finally, the calculated probability of earthquake disaster was used to achieve the purpose of quantitative evaluation of land seismic suitability.

Keywords: seismic suitability; logistic regression model; GIS; SPSS

0 引言

随着社会的发展,城市已成为人类聚集和建设的主要场所,随之而来的却是地震灾害威胁的加大。地震灾害在地质方面一般表现为软土震陷、砂土液化、滑坡崩塌等,不同的土地地质情况下,发生地震灾害的可能性也就不同。作为一项有效降低城市土地地震灾害风险的基础工作,城市土地抗震适宜性评价对城市的规划与建设有很重要的指导作用,在国内外的城市规划中都已得到普遍关注。其评价手段一般为建立评价体系,利用 GIS 划分研究区栅格单元,通过栅格运算得到适宜性评价结果。而评价方法选择较多,为提高评价的准确度,人们已更倾向于采用量化的评价方法,如 Logistic 模型法、层次分析法等。

土地的抗震适宜性评价^[1]是指土地在地震发生时,灾害预防的适宜性情况,与土地出现地震灾害的概率大小直接相关。而土地地震灾害发生与否,是一个典型的二分类变量问题,因此,本文考虑采用二分类变量 Logistic 回归模型^[2,3]对城市土地地震灾害概率进行预测,通过灾害发生概率与土地抗震适宜性等级间的关系,实现量化的城市土地抗震适宜性评价。

1 Logistic 回归模型

Logistic 回归模型作为二分类变量非线性回归模型,已多次用于滑坡^[4]、基坑沉降^[5]等灾害预测中,

作者简介:袁媛,硕士,Email:358751744@qq.com。

并得到了较好的预测效果。在城市土地抗震适宜性评价过程中,土地是否会发生地震灾害可看作因变量,0代表不发生地震灾害,1表示发生地震灾害,而各个土地地震灾害影响因子可看作自变量。由于地震灾害发生与否不是连续事件,故此因变量与自变量的关系需用非线性回归表示。因此本文采用 Logistic 非线性回归模型进行回归分析,其表达式为:

$$P(y = 1 | x_{1i}, x_{2i}, \dots, x_{ki}) = \frac{1}{1 + e^{-\varepsilon_i}} \quad (1)$$

$$\varepsilon_i = \alpha + \sum_{k=1}^k \beta_k x_{ki} \quad (2)$$

式中: ε_i 表示影响因子的线性函数。事件发生的概率 $P(y = 1 | x_{ki})$ 与自变量 x_{ki} 之间是非线性单调函数, x_{ki} 为事件的影响因子,在城市土地抗震适宜性评价中也称为评价指标。随着 x_{ki} 的增加 $P(y = 1 | x_{1i}, x_{2i}, \dots, x_{ki})$ 也单调增加,或随着 x_{ki} 的减少 $P(y = 1 | x_{1i}, x_{2i}, \dots, x_{ki})$ 也单调减少,这类曲线与随机变量的累积分布曲线相似,取值范围在(0,1)区间内,且呈S形曲线。利用 Logistic 回归模型进行地震灾害发生概率计算时,开始各评价指标变化很小,地震灾害发生的概率变化也很小,在中间阶段随着评价指标间的差异加大,对应的概率值出现波动,但当评价指标样本量累积到一定程度后,地质灾害发生概率就会保持在同一数值。

在此,设 P 为地震灾害发生概率,其取值范围为(0,1)。因变量为地震灾害发生概率 P ,自变量为各评价指标 $x_{1i}, x_{2i}, \dots, x_{ki}$,建立 Logistic 非线性回归方程:

$$P = \frac{e^{\alpha + \sum_{k=1}^k \beta_k x_{ki}}}{1 + e^{\alpha + \sum_{k=1}^k \beta_k x_{ki}}} \quad (3)$$

式(3)中, α 为常数,其意义是指在没有任何评价指标的干扰下,地震灾害发生与不发生的概率之比的对数值; β_k 为逻辑回归系数,表示某一个评价指标发生变化时,地震灾害发生与不发生的概率之比的对数变化值,实际意义就是土地抗震适宜性评价各指标的权重值。

由式(3)可以看出,如果逻辑回归系数 β_k 确定,就可以通过各个评价指标值计算出某一区域的地震灾害发生概率 P ,由 P 值大小与土地适宜性等级间的关系,就可以实现定量化对某地区的土地抗震适宜性评价。因此,问题就转化为,通过历史地震灾害发生情况和研究区土地抗震适宜性评价指标 $x_{1i}, x_{2i}, \dots, x_{ki}$ 带入 Logistic 模型,反演出逻辑回归系数 β_k ,即得到各评价指标的权重值。

2 数据预处理

2.1 评价指标的确立

南京市规划区占地 4386.6km²,处于长江中下游,为长江河谷的一部分,地貌上属低山丘陵区,长江自西南至东北流经市境中部。全市低山岗地占土地总面积的60%以上,对于坡度大于20%的地形,均不同程度存在滑坡和崩塌的可能。南京市规划区内断裂构造较为复杂,主要断裂有近十条,其中四条仍有明显活动迹象,因此断层分布对城市土地抗震适宜性的影响不容忽视。另外,南京长江两岸软土广布,这些岸段在长江水流作用下容易发生侵蚀崩岸,地震作用则会进一步加剧江岸的崩滑,且软弱地基在地震中可能产生较大的沉陷,这对建筑物和生命线工程的安全影响较大。南京市沿长江两岸及主城区的古河道还广泛分布有砂土层,容易发生地震砂土液化。此外,值得注意的是,城市因其特殊性,各种工程活动,例如抽取地下水、加载等,都会对土地的稳定造成影响。根据地震灾害固有影响因素和南京市具体情况,建立南京市土地抗震适宜性评价指标体系,如图1所示。

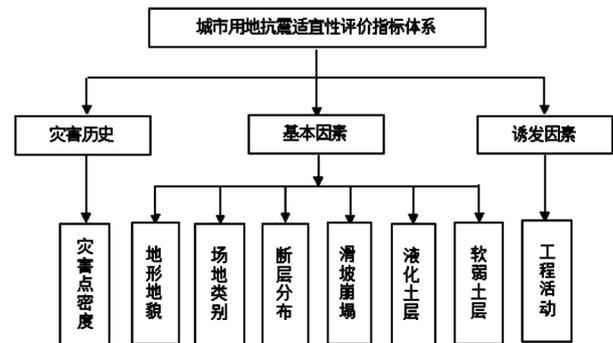


图1 南京市土地抗震适宜性评价指标体系

2.2 评价指标赋值

各评价指标的赋值则需要根据影响各指标的因素进行分级赋值,并将其统一控制在(0,1)区间内。例如,灾害点密度越大的区域,越容易再次发生灾害,故赋值越接近1;滑坡崩塌则在综合考虑坡型、坡度、坡高以及沿江边坡稳定性的基础上,划分等级,从而赋值;软弱土层震陷的危害性大小与软土震陷量 S_T 有关, S_T 越大,越易发生软土震陷,赋值就越接近1;液化指数 I_{LE} 越大,砂土液化也就越严重,赋值也越大;而人类工程活动的赋值,则与人类对土地的开发利用程度有关,人类工程活动越密集的区域,对土体的破坏也越大,越易发生人为因素造成的破坏。依据这一规律,得到各评价指标的赋值标准,见表1。

根据已有的七十余份勘察报告,可以直接得到有关区域的赋值,其余地区则运用 GIS 进行空间插值法

评价指标赋值标准 表 1

评价指标赋值	0	0~0.5(0.33)	0.5~1(0.67)	1
灾害点密度	无	稀疏	较密	密集
地形地貌	平坦地区	丘陵地区	河漫滩区	陡峭山地
场地类别	I 类场地	II 类场地	III 类场地	IV 类场地
断层分布	无活动性	活动性弱	活动性中	活动性强
滑坡崩塌	稳定	较稳定	有条件稳定	不稳定
软弱土层	$S_T=0$	$0 < S_T < 5$	$5 < S_T < 15$	$S_T > 15$
液化土层	$I_{LE}=0$	$0 < I_{LE} < 5$	$5 < I_{LE} < 15$	$I_{LE} > 15$
人类工程活动	各类保护区、禁区	二级保护区	保护区边缘区	非保护区、禁区

赋值^[6],从而得到整个研究区的赋值结果。而进行指标运算前需要将研究区划分成若干网格单元,运用同一网格单元内的指标进行灾害概率计算。利用 GIS 的网格划分功能,将研究区以 33m × 33m 大小^[7]划分单元格,共得到 4028099 个单元格,将赋值结果以单元格为单位转化为 Raster to ASC II 数据,由于数据量过于庞大,超出了软件的运算范围,故从中随机抽取 5000 例数据作为样本,输入 SPSS 19.0 数据分析软件中。

3 二分 Logistic 模型回归分析

在软件 SPSS19.0 中,导入因变量和自变量数据集,设置灾害发生情况为因变量,其余为自变量,运用分析模块中的多项 Logistic 回归工具进行运算,结果如表 2 所示。

多项 Logistic 回归计算结果 表 2

评价指标	回归系数	标准误差	卡方值	显著水平	发生比率	Exp(B) 的置信区间 95%	
	β_k	S.E	Wald	P	OR	下限	上限
地形地貌	0.469	0.203	5.336	0.012	0.423	0.02	0.68
断层分布	0.638	0.573	1.239	0.002	2.126	0.90	3.72
场地类别	0.749	0.245	9.345	0.031	0.926	0.18	2.63
滑坡崩塌	1.083	0.532	4.145	0.021	2.693	1.38	5.31
软弱土层	1.362	0.285	22.83	0.018	0.163	0.03	0.21
液化土层	1.159	0.346	11.22	0.009	2.692	0.83	3.91
灾害点密度	0.581	0.169	11.83	0.002	1.724	1.27	2.95
人类工程活动	1.347	0.924	2.126	0.039	0.726	0.21	1.19
Constant	7.388	3.277	68.087	0.134	11.470		

表格中 P 为显著水平,是指检验统计量拒绝方向的概率^[8],可以计算出自变量取一定值时因变量为 1 的预测概率值 $P < a$,则拒绝 0 假设,即各组的划分与因变量的实际取值相关,意味着模型的拟合优度较高;反之较低。在许多研究领域 0.05 的 P 值通常被

认为是可接受的边界水平。表中各系数 $P < 0.05$,说明所建立的模型达到了很好的拟合效果,具有很好的统计学意义,可以很好的预测事件发生概率。所建立的回归方程为:

$$P = \frac{e^{7.388+0.469x_1+0.638x_2+0.749x_3+1.083x_4+1.362x_5+1.159x_6+0.782x_7+1.347x_8}}{1 + e^{7.388+0.469x_1+0.638x_2+0.749x_3+1.083x_4+1.362x_5+1.159x_6+0.782x_7+1.347x_8}} \quad (4)$$

从所得到的回归方程可以看出,各逻辑系数值 β_k 均为正值,说明因变量会随着自变量的增加而非线性增加,而 β_k 值越大,对因变量的影响力也越大。由 Logistic 模型计算结果可以看出,滑坡崩塌指标、软弱土层指标、液化土层指标和人类工程活动对南京市土地抗震适宜性评价的影响较大,这与南京市属于丘陵地区,存在滑坡不稳定地区,且沿江岸段在长江水流和地震作用下易发生崩滑的实际情况相符。此外,南京长江、秦淮河、滁河漫滩地区均广泛分布较厚软土层和可液化砂土层,使得软弱土层与液化土层对城市土地抗震性的影响较大,而随着南京市城区的发展扩大,人类工程活动对土地抗震防灾能力的影响也不容忽视。

城市土地适宜性评价各指标权重可以通过各逻辑系数与逻辑系数和之比得到,其公式为:

$$\alpha = \frac{B_i}{\sum_{i=1}^7 B_i} \quad (5)$$

式中: α 为各评价指标的权重值; B_i 为 Logistic 回归方程中各指标的回归系数。

计算得的各指标权重,如表 3 所示。

评价指标的权重值 表 3

指标名称	地形地貌	断层分布	场地类别	滑坡崩塌	软弱土层	液化土层	灾害点密度	人类工程活动
逻辑回归系数	0.469	0.638	0.749	1.083	1.362	1.159	0.581	1.347
指标权重	0.063	0.086	0.101	0.147	0.184	0.158	0.079	0.182

将灾害概率计算方程输入 GIS 中,通过各指标的赋值运算,得到南京市土地地震灾害发生概率图,如图 2 所示。建立地震灾害发生概率与适宜性评价等级的关系,如表 4 所示。由此,即得到南京市土地抗震适宜性评价的量化区划结果,如图 3 所示。

适宜性等级与灾害概率的关系 表 4

适宜性评价等级	适宜	较适宜	有条件适宜	不适宜
地震灾害发生概率	0.00~0.20	0.20~0.40	0.40~0.60	0.60~1.00

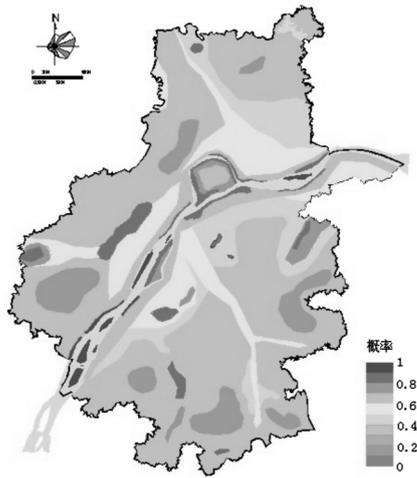


图2 南京市地震灾害发生概率图



图3 南京市土地抗震适宜性评价区划图

4 结语

本文根据南京市的实际情况,建立了南京市土地抗震适宜性评价指标体系,并以GIS为平台,利用

Logistic 回归方程确定指标权重,得到了南京市土地地震灾害概率的计算公式。实际证明,Logistic 模型可以很好地拟合地震灾害与各评价指标间的关系,是较为直观量化的土地抗震适宜性的评价方法。但需指出的是,城市土地抗震适宜性的影响因素很多,本文由于资料的限制,仅考虑了影响较大的8个因素,同时在对指标赋值时,有一定的主观性,如何确立一个统一的赋值标准还有待研究。

参考文献

- [1] 梁师俊. 基于GIS的可托理论在城市用地抗震适宜性评价中的应用[J]. 地理与地理信息科学, 2011, 27(6): 74-77.
- [2] Hosmer D W, Lemeshow S. Applied Logistic Regression [M]. New York: John Wiley and Sons, 1989.
- [3] 杨小平. 二分 Logistic 模型在分类预测中的应用分析[J]. 四川师范大学学报: 自然科学版, 2009, 32(3): 393-395.
- [4] 许湘华. 用 Logistic 回归模型编制滑坡灾害敏感性区划图的方法研究[J]. 铁道科学与工程学报, 2010, 7(5): 87-91.
- [5] 谭献良, 邓宗伟. Logistic 模型地基沉降预测中的应用[J]. 湖南城市学院学报(自然科学版), 2010, 19(2): 14-16.
- [6] 朱求安, 张万昌, 余钧辉. 基于GIS的空间插值方法研究[J]. 江西师范大学学报: 自然科学版, 2004, 28(2).
- [7] 崔党群. Logistic 曲线方程的解析与拟合优度测验[J]. 数理统计与管理, 2005, 24(1): 112-115.
- [8] 李军, 周成虎. 基于GIS栅格滑坡风险评价方法中格网大小选取分析[J]. 遥感学报, 2003, 7(2): 86-92.