

中国地质灾害伤亡事件的空间格局及影响因素

王 瑛^{1,2}, 林齐根^{1,2}, 史培军^{1,2}

(1. 北京师范大学环境演变与自然灾害教育部重点实验室, 北京 100875;

2. 北京师范大学减灾与应急管理研究院, 北京 100875)

摘要: 对中国2000-2012年造成人员伤亡的地质灾害事件进行分析, 其空间分布格局受地形等自然环境要素的影响, 南多北少, 主要位于川西山区和云贵高原地区, 东南丘陵地区, 北方黄土丘陵, 以及祁连山脉和天山山脉等地区, 但局部地区的分布格局表明其还受到人为因素影响。构建基于二元Logistic回归的中国地质灾害伤亡事件发生概率模型(CELC), 定量分析自然、人为因素的影响程度, 结果表明GDP增长率是仅次于地形起伏度的第二大影响因素, GDP增长率每增加2.72%, 地质灾害伤亡事件发生的概率变为原来的2.706倍。此外还有多年平均降水、植被覆盖度、岩性、土壤类型、断裂带、产业类型和人口密度等因素。将CELC模型应用于中国县域, 计算各个县的地质灾害伤亡事件概率, 发现尚未发生但概率较高的县有27个, 或为贫困县、或为矿产工业县域, 或为房产过度开发县, 它们是未来中国需要重点防范地质灾害的县域。

关键词: 地质灾害; 人员伤亡事件; 空间格局; 影响因素; 县域; 中国

DOI: 10.11821/dlxb201705011

1 引言

地质灾害极易造成人员伤亡, 中国2000-2012年的13年里共发生地质灾害33.9万起, 伤亡45381人, 年均伤亡约3500人, 仅次于地震、洪水灾害^[1]。地质灾害种类较多, 但根据2004年国务院颁发的《地质灾害防治条例》^[2]规定, 地质灾害主要指危害人民生命和财产安全的崩塌、滑坡、泥石流、地面塌陷、地裂缝、地面沉降六种与地质作用有关的灾害, 其中崩塌、滑坡、泥石流(简称崩滑流)是引发人员伤亡的三大主要地质灾害, 本文的地质灾害主要指崩滑流灾害。

2011年, 中国地质环境监测院根据环境地质调查数据编制了中国崩塌滑坡灾害图^[3]、中国泥石流灾害图^[4], 并指出, 中国崩滑流灾害主要分布于东北、华北丘陵山地, 东南丘陵山地, 川西北龙门山区, 秦巴山区, 三峡库区沿岸, 黔西南山区、黄土高原, 西北山地盆地高原区, 青藏高原。但是上述灾害图中, 只有该灾害事件的强度, 没有相关的灾情记录。而对于灾害风险防范工作, 造成人员伤亡的灾害事件, 更应引起重视。

国内外研究表明, 地质灾害的发生受到自然环境类因素影响, 包括地形、岩性、地质、土壤、植被、降水^[5-11]等。但是, 地质灾害与地震、水旱灾害的不同之处还在于它们的发生与人类自身活动对环境的影响之间有着紧密的联系。例如中国20世纪的地质灾害

收稿日期: 2016-08-01; 修订日期: 2017-01-15

基金项目: 国家自然科学基金项目(41271544); 国家重点研发计划专项项目课题(2016YFA0602403); “十二五”科技支撑计划项目(2012BAK10B03) [Foundation: National Natural Science Foundation of China, No.41271544, National Key Research and Development Program Project, No.2016YFA0602403, National Key Technology R & D Program of the Twelfth Five-Year Plan of China, No.2012BAK10B03]

作者简介: 王瑛(1974-), 女, 云南曲靖人, 教授, 主要从事灾害风险评估和灾后恢复研究。E-mail: wy@bnu.edu.cn

中, 70%以上的大型滑坡与人类活动有密切关系, 其中 50%的滑坡就是人类活动直接触发的^[12]。日本学者 Ayalew 等^[13]以道路网络、多个地形因素为自变量, 对日本中部山脉进行滑坡危险性评估, 发现道路网络是影响滑坡发生和分布重要的决定因素; Jadda 等^[14]采用概率频率比模型和地理信息技术, 综合考虑道路、地质、土壤、地形、水系等因素进行伊朗厄尔布尔士山脉滑坡危险性制图, 结果为地质、道路影响最大, 距离道路越近滑坡危险性越高; Eeckhaut 等^[15]的欧洲滑坡危险性分析制图中将土地利用类型分为农田、森林、草地、裸露地、城市用地和常被淹没土地, 共 6 类, 以农田为参考类别, 森林比农田更容易发生地质灾害, 而城市用地则相反。但是, 总的来看, 上述研究的主要关注点是局部山区, 以及道路、土地利用这 2 个直接触发因素, 没有探讨人类综合活动因素对地质灾害的影响。

因此, 本文以中国县域为研究单元, 基于 2000-2012 年的主要地质灾害伤亡事件数据, 从宏观尺度上, 对中国地质灾害人员伤亡事件的空间格局进行研究, 定量评估自然因素、人类活动因素对地质灾害伤亡事件的贡献程度。这对于认识中国地质灾害造成人员伤亡事件的主要规律, 制定相应的防灾减灾措施极为重要。

2 研究区与研究方法

2.1 研究区和数据

中国地质和地理环境复杂, 气候条件时空差异大, 地质灾害分布广、危害大^[15]。但是, 中国的地质灾害事件统计发布, 自 2011 年以后才逐渐规范完善, 由中国国土资源部通过《地质灾害灾情险情报告》^[16]实时报道发布。年度灾情由地质环境监测院采用《全国地质灾害通报》^[17]的方式统计公布。但 2011 年之前的地质灾害事件数据就较为分散。因此作者分别从民政部国家减灾中心、国土资源部、新闻网站, 收集了这些部门统计的 2000-2012 年共 13 年造成人员伤亡事件的崩滑泥灾害案例。根据灾害发生的时间、地点, 进行校对排重, 将每一起地质灾害的类型、发生时间、地点、人员伤亡、灾害损失等信息进行统一整理, 建成《中国主要地质灾害伤亡数据库》, 共计 576 起地质灾害伤亡事件, 具体数据来源如表 1 所示。

需要指出的是, 中国会出现一些山洪灾害与泥石流灾害并发的情况, 例如 2005 年 6 月黑龙江宁安沙兰镇小学山洪泥石流灾害、2009 年舟曲特大山洪泥石流灾害等, 此类山洪泥石流并发灾害的灾情都十分巨大, 形成机理与单发泥石流灾害又有较大差别, 故本文的数据库不包括此类并发型灾害事件。

根据地质灾害伤亡事件所在的县域, 编制中国地质灾害伤亡县分布图 (2000-2012 年) (图 1), 共有 334 个县域发生过地质灾害人员伤亡事件, 这些县简称地质灾害伤亡县。

分析图 1, 中国地质灾害伤亡县的分布总体呈南多北少的格局, 主要分布于四大区域: 一是中国西南部的四川盆地和云贵高原地区, 约占 39%; 二是东南丘陵地区, 约占 34%; 三是位于北方的黄土丘陵; 约占 12%; 四是西北祁连山脉和天山山脉等地区, 约

表 1 中国主要地质灾害伤亡数据库的数据来源

Tab. 1 Data sources of China's geological disaster casualty database

数据来源	编制单位	记录数	获得途径
地质灾害灾情险情报告 ^[16]	中国国土资源部	95	国土资源部网站
全国地质灾害通报 ^[17]	中国地质环境监测院	81	中国地质环境信息网
昨日灾情报告 ^[18]	民政部国家减灾中心	228	民政部国家减灾中心
Web 新闻搜索	新闻网站	172	网络收集整理

10%。与中国崩塌滑坡灾害图^[3]、中国泥石流灾害图^[4]相比,中国华北、东北等平原地区以及青藏高原等人烟稀少地区只有很少的伤亡事件分布。

综合中国自然环境特征,通过多个因子进行分析,本文选择地形、断裂带、降水、植被、岩性、土壤类型作为影响地质灾害伤亡事件的6类主要自然环境类因素。通过GIS空间分析提取各县域内平均地形起伏度(图2a)、县域的断裂带情况(图2b)、年平均降水(图2c)、植被覆盖度(图2d)、县域内的主要岩性类型(图2e)、县域内的主要土壤类型(图2f)这6个指标来表征上述因子(数据来源为地理空间数据云^[19]和地球系统科学数据共享平台^[20])。对比图1和图2可以发现,这些自然因素和地质灾害伤亡事件有较大的空间相关性:地形起伏偏大、断裂带密集、降水偏多、植被覆盖偏高的地区,地质灾害伤亡事件偏多。

但是,并不是所有具有上述特征的县域都会发生地质灾害伤亡事件,例如云南、福建等省的一些县域,虽然自然环境因素和周围县域很相似,但并未发生伤亡事件,这个原因正如前所述,与人类活动有关。因此,地质伤亡县域的分布格局还应考虑人口密度(图3a)、人类活动强度等因素。

经济是反映人类活动强度最综合的指标。2000-2012年中国经济始终保持快速增长,平均每年GDP增速达10.14%(按可比价格计算),但是一些地区经济的快速增长往往是以疯狂挖掘资源、环境破坏为代价的。世界银行的经验数据表明社会黄金发展期,也是各类灾难事故风险高发期^[21]。因此,本文采用GDP增速(图3b)、产业类型(图3c)2个指标来反映人类活动强度。GDP增速是指2000-2012年各县的年均GDP增长速度;产业类型是指2010年各县的第一产业占GDP的比例,可以采用文献^[22-23]的自然断点法(Nature Break)将各县分为3类:第一产业优势县、中等县和弱势县。

从宏观格局看,图3中各分图与图1地质灾害伤亡分布的相关性不高,人口密度呈东西分布、GDP增速较快的县零散分布于各个省区、产业类型县在中国东中地区混杂分布,因此,还需要采用回归模型对这些数据进行进一步的分析。

2.2 研究方法

Logistic回归模型是一种概率型非线性回归模型,是研究影响因素与因变量之间关系的常用方法^[24-26]。近年来,该模型被广泛应用于地质灾害的危险性、敏感性评估与制图^[5-7, 9, 13, 27-29]。二元Logistic回归方程表达式如下:

$$P = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)} \quad (1)$$

式中: P 为因变量,是自变量因子相对于某一事件的发生概率,取值范围为 $[0, 1]$; x_i 是自变量因子($i = 1, 2, \dots, k$),是影响事件发生的因素; k 为自变量个数; β_i 是偏回归系数,反映自变量因子 x_i 对 P 的影响程度大小。

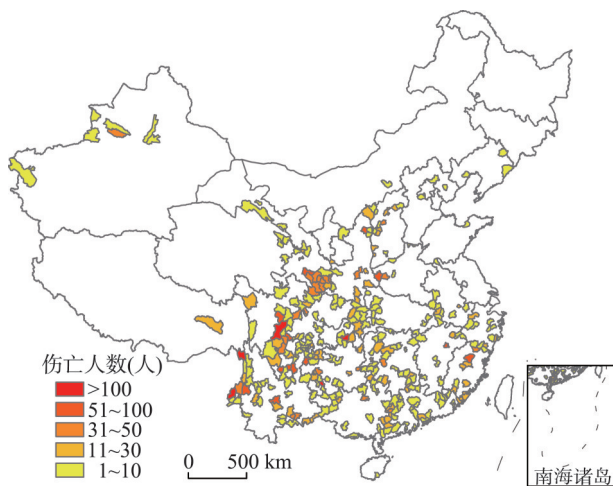


图1 中国地质灾害伤亡县空间分布(2000-2012年)

Fig. 1 Distribution of geological disaster casualty counties in China (2000-2012)

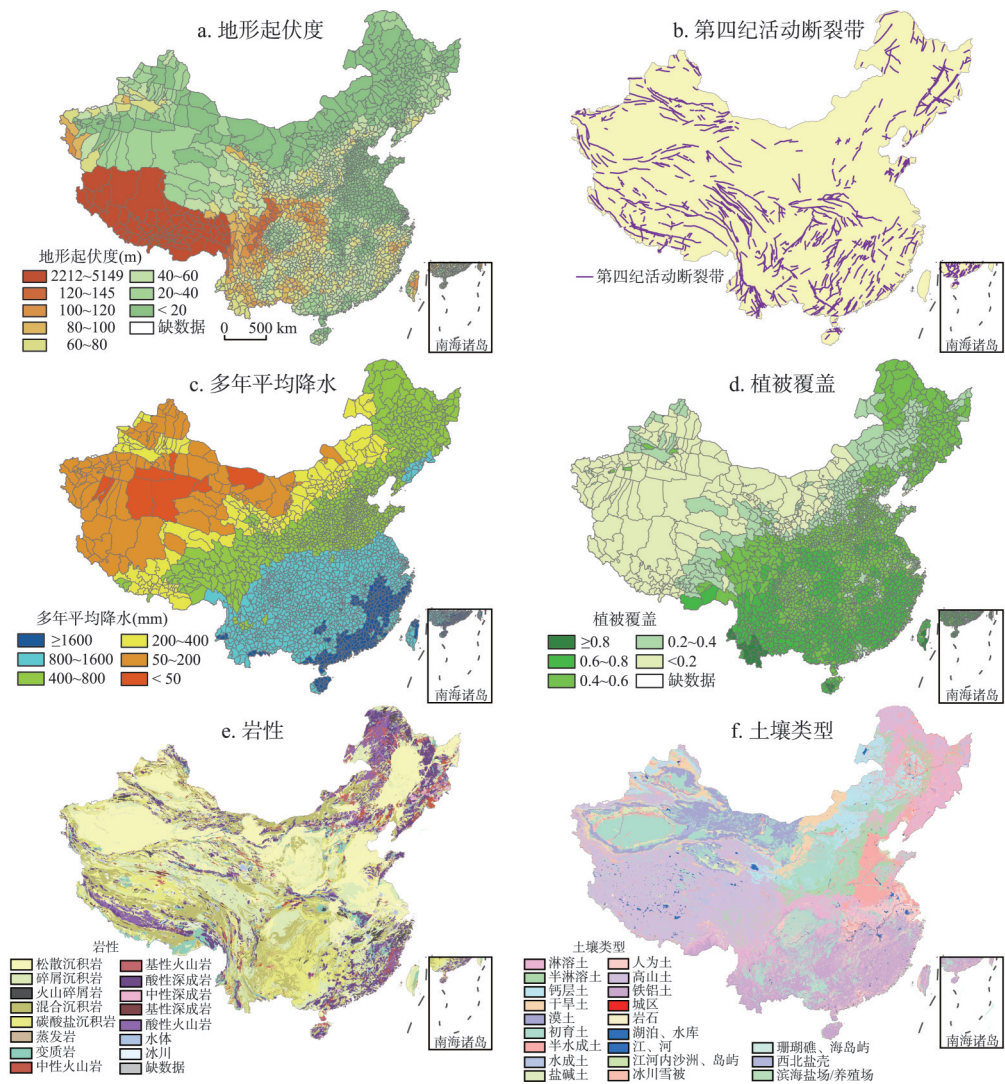


图2 中国自然环境类因子分布图
Fig. 2 Distribution maps of natural environment factors of China

Logistic 回归模型中，自变量有连续变量和分类变量两种。连续变量可以直接放入式 (1)，回归分析后，得到偏回归系数 β ，其意义是自变量连续变化对 P 大小的影响程度。分类变量，首先要根据数值情况，进行分类，分类后赋上相应的数值，代入式 (1)，所得到的 β 的意义是：与参考类因子相比，该类因子对事件的发生是否存在显著差异，如果存在，则说明该因子对因变量有影响。

根据《中国主要地质灾害伤亡数据库》中有无伤亡记录，对全国 2215 个县域进行分类，由于北京、天津和上海三市的市辖区以及台湾等县域缺少数据，西藏自治区整体地形较为特殊，故上述地区的 7 个伤亡县未进入模型计算，伤亡县数变为 327 个。将有伤亡记录的县域， P 设置为 1；无伤亡记录的县， P 设置为 0。再将这些县域对应的人口经济数据、自然数据代入式 (1)，即可根据偏回归系数 β 的情况，来判断各个因素与地质灾害伤亡事件的关系。

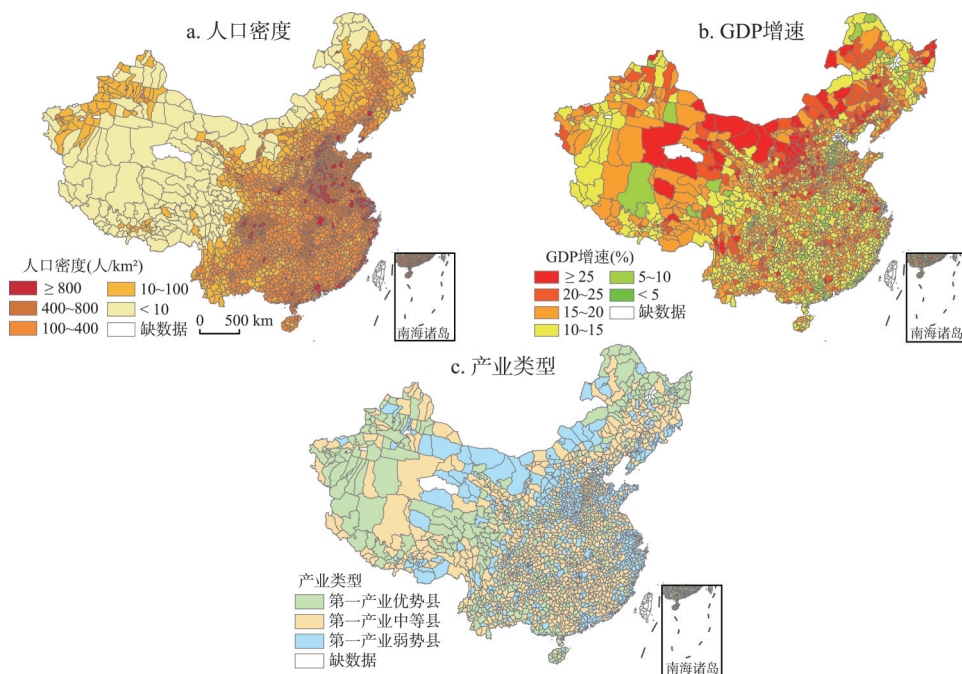


图3 中国人类经济活动类因子分布图

Fig. 3 Distribution maps of human-economic activity factors of China

本文“1”的县数为327;“0”的县数为1888。0的样本过大。在Logistic回归模型的应用中,King等^[30]认为“1”样本过少,会使得该类预测概率被严重低估,建议将两者的比率控制在0.2~1之间。因此,为了准确预测伤亡县“1”的概率,同时全面考虑非伤亡县“0”的样本信息,将非伤亡县“0”的全部样本与4倍伤亡县“1”的样本一起引入模型分析,即伤亡县“1”的样本为 $327 \times 4 = 1308$ 个;非伤亡县“0”为1888,1:0的比率为0.69,从而避免模型过于偏向非伤亡县。上述方法称为过抽样方法,目前在机器学习中被广泛使用,以实现训练样本数据均衡^[31]。

地形起伏度、年平均降水、植被覆盖度、GDP增速和人口密度5个影响因素均为连续变量。由于各影响因素的数值差别较大,因此首先进行自然对数标准化处理,再进入Logistic回归模型。

断裂带、岩性、土壤类型和产业类型是分类变量。断裂带变量,依据县域内是否有第四纪活动断裂带来赋值。如果有,赋值为1,否则为0。岩性参照Hartmann等^[32]的世界岩性图分类,将岩性分为松散沉积岩、碳酸盐沉积岩、混合沉积岩、碎屑沉积岩、蒸发岩、火山碎屑岩、变质岩、酸性深成岩、中性深成岩、基性深成岩、酸性火山岩、中性火山岩、基性火山岩、冰川和水体15类,以松散沉积岩为参考类别,将岩性变量转化为14个二分类虚拟变量;如某县岩性为碳酸盐沉积岩的面积最多,则该县在碳酸盐变量里是1,在其他岩性变量里都是0。土壤类型数据从中国科学院资源环境数据中心获得,采用“土壤发生分类”系统方法划分土壤类型,分为12个土纲,包括淋溶土、半淋溶土、钙层土、干旱土、漠土、初育土、半水成土、水成土、盐碱土、人为土、高山土和铁铝土;同岩性变量类似,以淋溶土为参考类别,其余转化为11个虚拟变量进入模型。产业类型变量,用自然断点法将各县分为3类,第一产业占GDP的比例 $\geq 31.21\%$,为第一产业优势县;第一产业占GDP的比例为 $14.70\% \sim 31.21\%$,就是第一产业中等县;其他为第

一产业弱势县；以第一产业优势县为参考类别，其余2类转化为2个虚拟变量代入模型计算。

3 结果分析

表2为Logistic回归模型中的变量，由表2可知，地形起伏度、GDP增速、年平均降水量、断裂带、岩性、土壤、产业类型、人口密度和植被覆盖度9个变量的Sig. < 0.05，说明这些变量对地质灾害伤亡事件确实存在影响。

影响程度的大小，由Exp(β)的大小决定，Exp(β)是优势比（简称OR），表示自变量每变化一个单位，引起事件发生概率相应变化的比例。Exp(β)值大于1，说明是正向影响；Exp(β)值小于1，说明是负向影响。Exp(β)距离1越大，说明该因子的影响程度越大。因此，地形起伏度变大、GDP增速变大、年平均降水量变大、断裂带经过、人口密度变大，都会增加该县域地质灾害伤亡发生的可能性，但植被覆盖度的变大会降低可能性，这说明虽然从全国尺度看，伤亡县主要分布在植被覆盖相对偏高的南方地区，但在各种自然因素一致的情况下，增加植被覆盖度可以降低地质灾害伤亡事件的发生。与淋溶土相比，干旱土、漠土、初育土和人为土会增加地质灾害伤亡事件发生的可能；与松散沉积岩相比，碳酸盐沉积岩和酸性火山岩会减少地质灾害伤亡事件发生的可能；与第一产业优势县相比，中等和弱势县会增大地质灾害伤亡事件发生的可能。

将表2中的参数代入式（1），结果为式（2）、（3）：

$$Z = -11.913 + 1.922X_1 + 0.996X_2 + 0.535X_3 - 0.333X_4 + 0.374X_5 + 0.317X_6 - 0.649X_{71} - 0.448X_{72} - 0.173X_{73} - 0.572X_{74} - 1.574X_{75} - 0.866X_{76} - 19.936X_{77} - 0.204X_{78} - 20.172X_{79} - 19.612X_{710} + 0.238X_{711} - 17.337X_{712} - 0.396X_{81} + 0.567X_{82} + 1.183X_{83} + 2.083X_{84} + 0.447X_{85} + 0.857X_{86} - 17.918X_{87} + 0.893X_{88} + 0.467X_{89} + 0.524X_{810} + 0.536X_{91} + 0.352X_{92}$$

$$P = \frac{\exp(Z)}{1 + \exp(Z)} \tag{3}$$

表2 Logistic回归模型中的变量
Tab.2 Variables applied in the logistic regression model

变量	β	S.E.	Sig.	Exp(β)
ln地形起伏度	1.922	0.117	0.000	6.834
lnGDP增速	0.996	0.209	0.000	2.706
ln年平均降水量	0.535	0.178	0.003	1.707
ln植被覆盖度	-0.333	0.159	0.037	0.717
断裂带	0.374	0.093	0.000	1.453
ln人口密度	0.317	0.063	0.000	1.373
岩性 ^a			0.000	
碎屑沉积岩	-0.649	0.171	0.000	0.523
火山碎屑岩	-0.448	0.330	0.174	0.639
混合沉积岩	-0.173	0.168	0.302	0.841
碳酸盐沉积岩	-0.572	0.178	0.001	0.564
酸性火山岩	-1.574	0.408	0.000	0.207
中性火山岩	-0.866	0.475	0.069	0.421
基性火山岩	-19.936	9093	0.998	0.000
酸性深成岩	-0.204	0.184	0.268	0.815
中性深成岩	-20.172	12480	0.999	0.000
基性深成岩	-19.612	40190	1	0.000
变质岩	0.238	0.309	0.442	1.268
水体	-17.337	27290	0.999	0.000
土壤类型 ^a			0.000	
半淋溶土	-0.396	0.241	0.1	0.673
钙层土	0.567	0.31	0.068	1.762
干旱土	1.183	0.477	0.013	3.263
漠土	2.083	0.532	0.000	8.026
初育土	0.447	0.169	0.008	1.564
半水成土	0.857	0.468	0.067	2.357
盐碱土	-17.918	12030	0.999	0
人为土	0.893	0.284	0.002	2.443
高山土	0.467	0.3	0.119	1.595
铁铝土	0.524	0.157	0.001	1.69
产业类型 ^a			0.000	
第一产业中等县	0.536	0.126	0.000	1.71
第一产业弱势县	0.352	0.145	0.015	1.422
Constant	-11.913	1.344	0.000	0.000

注：a：岩性、土壤类型和产业类型的参考类分别为松散沉积岩、淋溶土和第一产业优势县。

式中： P 为地质灾害伤亡发生事件概率； X_1 为ln地形起伏度； X_2 为lnGDP增速； X_3 为ln年平均降水量； X_4 为ln植被覆盖度； X_5 为断裂带； X_6 为ln人口密度； X_{71} 为碎屑沉积岩； X_{72} 为火山碎屑岩； X_{73} 为混合沉积岩； X_{74} 为碳酸盐沉积岩； X_{75} 为酸性火山岩； X_{76} 为中性火山岩； X_{77} 为基性火山岩； X_{78} 为酸性深成岩； X_{79} 为中性深成岩； X_{710} 为基性深成岩； X_{711} 为变质岩； X_{712} 为水体； X_{81} 为半淋溶土； X_{82} 为钙层土； X_{83} 为干旱土； X_{84} 为漠土； X_{85} 为初育土； X_{86} 为半水成土； X_{87} 为盐碱土； X_{88} 为人为土； X_{89} 为高山土； X_{810} 为铁铝土； X_{91} 第一产业中等县； X_{92} 第一产业弱势县。该式即为中国地质灾害伤亡事件发生概率模型（the Probability Model of Casualty Events Caused by Landslide in China, CELC模型）。

通过混淆矩阵和ROC曲线对CELC模型的精度进行评估。表3为CELC模型的预测混淆矩阵，即对是否为地质伤亡县两类情况的预测正确率，分割值为0.5， $P \geq 0.5$ 判断为地质灾害伤亡县， $P < 0.5$ 则判断为未发生地质灾害伤亡县。模型总的正确率为75.0%，其中非地质灾害伤亡县的正确率为75.0%，地质灾害伤亡县的预测正确率为74.9%。

图4中红色线为CELC模型的ROC曲线（Receiver Operating Characteristic），对ROC曲线下的面积（AUC）进行统计， $AUC = 0.826$ ，标准误差0.007。根据Swets^[33]的研究，ROC曲线下面积在0.5~0.7之间表示预测价值较低，在0.7~0.9之间表示预测价值中等，0.9以上表示预测价值高。因此，本模型具有相对较高的预测价值。

为了进一步验证本文模型的结果，参照Chung等^[34]和Poiraud^[35]的研究方法，对中国县域进行10次简单随机抽样，每次分别以70%的样本建立模型，30%的样本验证模型结果。因类别变量分类较多，在随机抽样验证过程中某些用于预测模型的类别在建立模型的过程中不存在，导致模型无法计算，因此，随机抽样交叉验证过程中不考虑岩性、土壤类型和产业类型变量。运用ROC曲线计算10次随机抽样建立模型的AUC值为0.811~0.831，相应的30%样本验证结果的AUC值为0.781~0.821（图4），并且10次试验结果的各个参数波动都较小。这说明，CELC模型中的各个因子具有较高的稳定性。

综上，在中国，地形起伏度、GDP增速、年平均降水、植被覆盖度、断裂带、人口密度、岩性、土壤类型和产业类型，对于地质灾害伤亡

表3 CELC模型的预测混淆矩阵
Tab. 3 Confusion matrix for the CELC model

实际是否伤亡	模型预测		正确率(%)
	是否伤亡		
	0	1	
0	1416	472	75.0
1	328	980	74.9
总体百分比			75.0

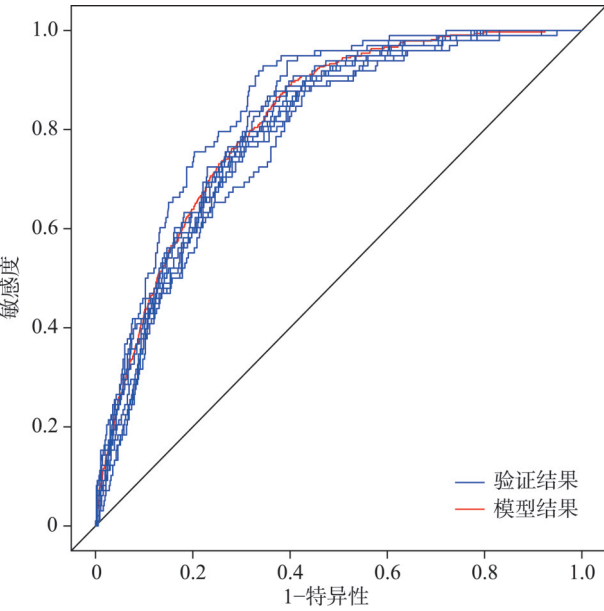


图4 中国地质灾害伤亡事件发生概率模型和70%样本建立的模型验证结果的ROC曲线
Fig. 4 ROC curve for probability model of casualty events caused by landslides in China and validation of models produced from 10 samples of 70% training data

事件的发生都有影响。

表2中,GDP增速的优势比OR值为2.706,这说明lnGDP增长率每增加1,即年均GDP增长率每增加2.72%,地质灾害伤亡事件发生的概率变为原来的2.706倍,由此可见,高速增长的GDP对于引发地质灾害伤亡事件有重要的影响。这从一个角度证实了,中国一些县域的GDP高速增长,是以破坏环境为代价的,继而引发了很多地质灾害伤亡事件。

除各县的GDP增速外,本文还分析了各县的GDP总值、人均GDP、地均GDP这些因子是否对地质灾害伤亡事件发生有影响,将这些因子引入模型进行计算,但结果表明它们对地质灾害伤亡事件发生都没有显著的影响。分析原因,GDP的大小与该地区当年的积累等诸多因素有关,并不能有效反映当地的人为活动强度。例如,一些前期经济较为发达县的GDP总值、人均GDP、地均GDP都会较高,但这些地区各种发展已经趋于稳定,对环境的破坏也相对较小,故模型结果为影响因子不显著。

4 讨论

将中国2215个县域的自变量数值,放入CELC模型中计算,可以得到各县地质灾害伤亡事件发生概率的大小。由于目前国内外对于地质灾害易发性、危险性概率的分级(包括分多少级、分级的阈值等)尚没有统一的标准,因此借鉴相关的地理制图数据分级方法^[22-23],根据上述2215个县数据的分布规律进行分级,分为5级,对应发生概率极高、高、中、低、极低,再对阈值进行适当的取整处理,最终结果如图5所示。从图上可以看到,中国地质灾害伤亡事件发生概率高的地区与图1的分布规律基本相同,主要集中于前述的4大区域。

但是,图5中发生概率>0.78的县域共有80个,与图1相比,其中42个县已经发生过地质灾害伤亡事件,38个县尚未发生过地质灾害。38个县中的27个县GDP增速超过本省GDP平均增速,我们认为,这27个县是中国未来可以通过加强监管减低地质灾害风险的区域,这些县的自然环境条件,如地形起伏度、降水、植被覆盖等,都极易引发地质灾害,而且近10年来,这些县的经济始终高速发展。依据其经济发展的不同特点,分为贫困县、矿产工业县域、房产过度开发县3类(表4)。

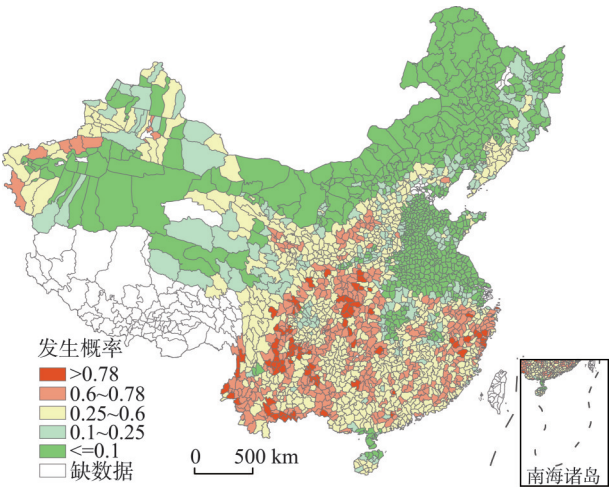


图5 中国各县域地质灾害伤亡事件发生概率
Fig. 5 Geological disaster casualty occurrence probability distribution in China

表4 地质灾害风险重点防范县	
Tab. 4 Counties, stressing geological disaster risk prevention	
类别	县 域
贫困县	金阳县、永善县、黔江区、泸定县、茂县、冕宁县、镇康县、石柱县、西昌市、会理县、鲁甸县、墨江县、屏边县、河口县、元阳县
	凤县、靖西县、米易县、峨眉山市、洪雅县、会东县、水富县、商洛市辖区
房产过度开发县	丽水市辖区、舟山市辖区、青田县、武夷山市

其中, 贫困县大部分是国家扶贫工作重点县, 受益于国家扶贫工作计划的支持, 近年来的经济发展较快, 但其本身经济水平相当落后, 使得灾害防御工程也相对匮乏。因此, 这些县在经济脱贫的同时要增大对地质灾害风险的防范力度。

矿产工业县依托丰富的矿产资源, 发展采矿冶金等工业, 对GDP高速增长具有极大贡献。例如, 陕西省凤县是全国四大铅锌矿基地之一, 其铅锌矿贮量350万t, 约占全省的80%。近10年来, 凤县平均GDP增速高达本省平均增速的两倍, 这是当地大力发展矿业取得的成果。但与此同时该县的植被、生态环境也都遭到了严重破坏, 增大了其未来发生地质灾害伤亡事件的可能性。因此, 应适量开发矿产资源, 同时注重对生态环境的保护, 做好地质灾害风险防范措施。

房地产过度开发县主要位于中国东部地区, 主要为市辖区, 经济水平相对较高, 工业发展对原材料的依赖程度有所下降, 向深加工发展, 第三产业比重不断增加, 具有充足的防灾减灾经济实力。但房地产业、旅游业等的大力发展需不断向山区扩张土地, 如浙江省地形素有“七山一水两分田”之说, 山地丘陵面积占70.4%, 平原面积仅有23.4%, 而位于其西南山区的丽水市山地面积更是占全市88.42%。近年来房地产业发展迅速, 2012年房地产投资额达到84.99亿元, 房地产施工面积逐年增长, 由2003年的540.55万 m^2 , 增至2012年的1913.46万 m^2 , 不断向山区扩张的土地利用使得更多的人分布在具有灾害风险山区。因此, 对于中国东部经济发达县, 由于城市建设扩张所带来的地质灾害风险必须重点防范。

5 结论

本文基于《中国地质灾害伤亡数据库》的县级人员伤亡数据, 采用Logistic回归模型, 对中国地质灾害人员伤亡事件的空间格局进行研究, 并定量评估自然因素、人类活动因素对地质灾害伤亡事件的贡献程度。

中国地质灾害伤亡县的分布总体呈南多北少的格局, 主要分布于西南部的四川盆地和云贵高原地区, 东南丘陵地区, 北方的黄土丘陵, 西北祁连山脉和天山山脉等地区。

通过构建CELC模型, 本文发现中国地质灾害伤亡事件受自然与人类活动因素的共同影响, 地形起伏度、GDP增速、多年平均降水、植被覆盖度、断裂带、岩性、土壤类型、产业类型和人口密度对中国地质灾害伤亡事件的发生都有影响。连续变量中按影响程度从大到小分别为地形起伏度、GDP增速、多年平均降水、植被覆盖度和人口密度。GDP增长率每增加2.72%, 地质灾害伤亡事件发生的概率变为原来的2.706倍, 定量证明中国快速增长的经济破坏了自然环境, 触发了较多的地质灾害。表3中, 常数项对应的显著性水平 <0.05 , 即常数项为0的假设检验不通过。这说明, 除本文考虑的因素外, 还存在着一些其他的因素对地质灾害伤亡事件的发生有影响。因此, 将来的工作是, 获取更多、更精细的数据, 来考虑自然因素、人类经济活动强度因素, 进一步提高地质灾害伤亡事件发生概率的预测精度。此外, 本文在构建模型时, 采用了过抽样方法, 即将伤亡县“1”样本进行4倍复制后引入模型分析, 提高了CELC模型的精度, 但这种人为增加样本的方法, 是否适用于其他logistic模型, 有待进一步研究。

CELC模型计算结果还表明, 中国地质灾害伤亡事件发生概率高, 但尚未发生的38个县域中, 大约有71%县的经济都处于高速发展, 其GDP增长率远高于本省的GDP增长率, 而且这些县可以分为3类: 贫困县、矿产工业县域、房产过度开发县。它们是中国需要加强监管, 减少人为破坏环境, 降低地质灾害风险的区域。

由于中国主要地质灾害伤亡数据库中的承灾体信息尚不完备, 故本文采用了二元 logistic 回归方法进行研究, 其结果只能用于计算某县域发生地质灾害伤亡事件, 未来随着数据库的完善, 还可以采用地理探测器等方法, 对比分析地质灾害伤亡率的空间影响因素。

参考文献(References)

- [1] Sheng Laiyun, Wang Wenbo, Zhong Shouyang. China Statistical Yearbook. Beijing: China Statistics Press, 2013. [盛来运, 王文波, 钟守洋. 中国统计年鉴. 北京: 中国统计出版社, 2013.]
- [2] The State Council. Geological Disaster Prevention Regulations. 2004. [国务院. 地质灾害防治条例. 2004.]
- [2] China Institute for Geo-Environment Monitoring. Rockfall and Landslide Disaster Map of China. Beijing: SinoMaps Press, 2007. [中国地质环境监测院. 中国崩塌滑坡灾害图. 北京: 中国地图出版社, 2007.]
- [4] China Institute for Geo-Environment Monitoring. Debris Flow Disaster Map of China. Beijing: SinoMaps Press, 2007. [中国地质环境监测院. 中国泥石流灾害图. 北京: 中国地图出版社, 2007.]
- [5] Eeckhaut M., Hervás J., Jaedicke C, et al. Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. *Landslides*, 2011, 9(3): 357-369.
- [6] Ramani S E, Pitchaimani K, Gnanamanickam V R. GIS based landslide susceptibility mapping of Tevankarai Ar Sub-watershed, Kodaikkanal, India using binary logistic regression analysis. *Journal of Mountain Science*, 2011, 8(4): 505-517.
- [7] García-Rodríguez M J, Malpica J A, Benito B, et al. Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression. *Geomorphology*, 2008, 95(3/4): 172-191.
- [8] Guzzetti F, Peruccacci S, Rossi M, et al. Rainfall thresholds for the initiation of landslides in central and southern Europe. *Meteorology and Atmospheric Physics*, 2007, 98(3/4): 239-267.
- [9] Ohlmacher G C, Davis J C. Using multiple logistic regression and GIS technology to predict landslide hazard in Northeast Kansas, USA. *Engineering Geology*, 2003, 69(3/4): 331-343.
- [10] Atkinson P.M., ssari R. Generalised linear modelling of susceptibility to landsliding in the Central Apennines, Italy. *Computers & Geosciences*, 1998, 24(4): 373-385.
- [11] Ayalew L, Yamagishi H, Ugawa N. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. *Landslides*, 2004, 1(1): 73-81.
- [12] Huang Runqiu. Large-scale landslides and their sliding mechanisms in China since the 20th century. *Chinese Journal of Rock Mechanics and Engineering*, 2007, 26(3): 433-454. [黄润秋. 20 世纪以来中国的大型滑坡及其发生机制. 岩石力学与工程学报, 2007, 26(3): 433-454.]
- [13] Ayalew L, Yamagishi H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, central Japan. *Geomorphology*, 2005, 65(1/2): 15-31.
- [14] Jadda M, Shafri H Z M, Mansor S B. PFR model and GiT for landslide susceptibility mapping: A case study from central Alborz, Iran. *Natural Hazards*, 2011, 57(2): 395-412.
- [15] Li Yuan, Meng Hui, Dong Ying, et al. Main types and characteristics of geo-hazard in China-based on the results of geo-hazard survey in 290 counties. *The Chinese Journal of Geological Hazard and Control*, 2004, 15(2): 29-34. [李媛, 孟晖, 董颖, 等. 中国地质灾害类型及其特征: 基于全国县市地质灾害调查成果分析. 中国地质灾害与防治学报, 2004, 15(2): 29-34.]
- [16] Ministry of Land and Resources of China. Report on Geological Disaster Situation. <http://www.mlr.gov.cn/dzhj/dzzh/zqxqbg/>, Accessed September, 2014. [中国国土资源部. 地质灾害灾情险情报告. <http://www.mlr.gov.cn/dzhj/dzzh/zqxqbg/>, 访问时间: 2014 年 9 月.]
- [17] China Institute for Geo-Environment Monitoring. China Geological Hazard Bulletin (2004-2012). China Geological Environmental Information Site. <http://www.cigem.gov.cn/>, Accessed September, 2014. [中国地质环境监测院. 全国地质灾害通报(2004-2012). 中国地质环境信息网. <http://www.cigem.gov.cn/>, 访问时间: 2014 年 9 月.]
- [18] Ministry of Civil Affairs National Disaster Reduction Center. Yesterday's Disaster (2004-2012). [民政部国家减灾中心. 昨日灾情(2004-2012).]
- [19] Computer Network Information Center of CAS. Geospatial Data Cloud. <http://www.gscloud.cn/>, Accessed August, 2015. [中国科学院计算机网络信息中心. 地理空间数据云. <http://www.gscloud.cn/>, 访问时间: 2015 年 8 月.]
- [20] Institute of Geographic Sciences and Natural Resources Research of CAS. Data Sharing Network of Earth System

- Science. <http://www.geodata.cn/Portal/index.jsp>, Accessed June, 2015. [中国科学院地理科学与资源研究所. 地球系统科学数据共享平台. <http://www.geodata.cn/Portal/index.jsp>, 访问时间: 2015年6月.]
- [21] Zhang Haibo, Tong Xing. Public policy in a high-risk society. *Journal of Nanjing Normal University: Social Science*, 2009(6): 23-28. [张海波, 童星. 高风险社会中的公共政策. *南京师大学报: 社会科学版*, 2009(6): 23-28.]
- [22] Hu Y, Wang J, Li X, et al. Geographical detector-based risk assessment of the under-five mortality in the 2008 Wenchuan earthquake, China. *PloS One*, 2011, 6(6): e21427.
- [23] Luo W, Jasiewicz J, Stepinski T, et al. Spatial association between dissection density and environmental factors over the entire conterminous United States. *Geophysical Research Letters*, 2016, 43(2): 692-700.
- [24] Li Hong, Gong Zhaoning, Zhao Wenji, et al. Driving forces analysis of reservoir wetland evolution in Beijing based on logistic regression model. *Acta Geographica Sinica*, 2012, 67(3): 357-367. [李洪, 宫兆宁, 赵文吉, 等. 基于 Logistic 回归模型的北京市水库湿地演变驱动力分析. *地理学报*, 2012, 67(3): 357-367.]
- [25] Qi Lili, Bo Yanchen. MAUP effects on the detection of spatial hot spots in socio-economic statistical data. *Acta Geographica Sinica*, 2012, 67(10): 1317-1326. [齐丽丽, 柏延臣. 社会经济统计数据热点探测的 MAUP 效应. *地理学报*, 2012, 67(10): 1317-1326.]
- [26] Liu Wangbao, Yan Xiaopei, Cao Xiaoshu. Housing type variation and its influencing factors in transitional urban China: Based on analysis of CGSS 2005. *Acta Geographica Sinica*, 2010, 65(8): 949-960. [刘望保, 闫小培, 曹小曙. 转型期中国城镇居民住房类型分化及其影响因素: 基于 CGSS(2005)的分析. *地理学报*, 2010, 65(8): 949-960.]
- [27] Chau K T, Chan J E. Regional bias of landslide data in generating susceptibility maps using logistic regression: case of Hong Kong Island. *Landslides*, 2005, 2(4): 280-290.
- [28] Lee S, Pradhan B. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides*, 2006, 4(1): 33-41.
- [29] Wang Y, Song C, Lin Q, et al. Occurrence probability assessment of earthquake-triggered landslides with Newmark displacement values and logistic regression: The Wenchuan earthquake, China. *Geomorphology*, 2016, 258: 108-119.
- [30] King G, Zeng L. Logistic regression in rare events data. *Political Analysis*, 2001, 9(2): 137-163.
- [31] Yen S J, Lee Y S, Lin C H, et al. Investigating the effect of sampling methods for imbalanced data distributions//2006 IEEE International Conference on Systems, Man and Cybernetics. *IEEE*, 2006(5): 4163-4168.
- [32] Hartmann J, Moosdorf N. The new global lithological map database GLiM: A representation of rock properties at the earth surface. *Geochemistry, Geophysics, Geosystems*, 2012, 13(12): 1-37.
- [33] Swets J A. Measuring the accuracy of diagnostic systems. *Science*, 1988, 240(4857): 1285-1293.
- [34] Chung C F, Fabbri A G. Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards*, 2003, 30(3): 451-472.
- [35] Poiraud A. Landslide susceptibility-certainty mapping by a multi-method approach: A case study in the tertiary basin of Puy-En-Velay (Massif central, France). *Geomorphology*, 2014, 216: 208-224.

Spatial pattern and influencing factors of casualty events caused by landslides

WANG Ying^{1,2}, LIN Qigen^{1,2}, SHI Peijun^{1,2}

(1. Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Beijing Normal University, Beijing 100875, China; 2. Academy of Disaster Reduction and Emergency Management, Beijing Normal University, Beijing 100875, China)

Abstract: The economy of China has maintained rapid growth with an average annual GDP growth rate of 10.14% (in comparable price) from 2000 to 2012. During this period, China witnessed frequent landslide disasters, including 338,964 identifiable individual landslide disasters that resulted in 45,381 casualties, including 9,928 deaths. Analysis of the casualty events caused by landslides from 2000 to 2012 revealed that the spatial pattern of the casualty

events was affected by terrain and other factors of the natural environment, which resulted in the distribution of casualty events being higher in the south region than in the north region. Hotspots of casualty events caused by landslides were in the western Sichuan mountain area and the Yunnan-Guizhou Plateau region, the southeast hilly area, the northern part of the loess hills, and the Qilian and Tianshan Mountains, among some others. However, their local distribution pattern indicated that they were also influenced by economic activity factors. To quantitatively analyze the influence of natural environment factors and human-economic activity factors, the binary logistic regression model was applied. The binary logistic regression model is a type of probabilistic nonlinear regression model describing the relationship between a binary dependent variable and a set of independent variables (explanatory factors). The explanatory factors used in this study included relative relief, mean annual precipitation, vegetation coverage, fault zones, lithology, soil type, GDP growth rate, industry type, and population density. The dependent variable used in this study was the presence (1) or absence (0) of casualty events caused by landslides in the county. For the logistic regression analysis, the continuous variables of relative relief, mean annual precipitation, vegetation coverage, GDP growth rate, and population density were substituted into the model. The categorical variables of fault zones, lithology, soil type, and industry type were transformed into binary dummy variables and then substituted into the model. The Probability Model of Casualty Events Caused by Landslide in China (CELC) was built based on the logistic regression analysis, and the confusion matrix and the receiver operating characteristic (ROC) curve were applied to assess the model performance. The results showed that all explanatory variables in the model were selected based on a significance level of 0.05. The coefficients of the explanatory variables showed that relative relief, GDP growth rate, mean annual precipitation, fault zones, and population density have a positive effect on casualty events caused by landslides. In contrast, vegetation coverage has a negative influence on casualty events caused by landslides. More specifically, the results showed that in terms of the influence degree of casualty events caused by landslides, the GDP growth rate ranks only second to relative relief. The probability of occurrence of casualty events caused by landslides will be 2.706 times that of the previous probability with an increase of GDP growth rate of 2.72%. In the evaluation of the model performance, the correct percentage in the confusion matrix is 75 % and the area under the ROC curve (AUC) is 0.826, revealing that the CELC model has good predictive ability. The CELC model was then applied to calculate the occurrence probability of casualty events caused by landslides for each county in China. The results showed that there are 27 counties with high occurrence probability but zero casualty events caused by landslides. The 27 counties can be divided into three categories: poverty-stricken counties, mineral-rich counties, and realty-overexploited counties, which are the key areas where great emphasis should be placed on landslides risk reduction.

Keywords: landslide; casualty event; spatial pattern; influencing factors; counties; China