

彭贵芬, 刘盈曦. 基于模糊信息概率区间数的突发地质灾害降水因子评价研究[J]. 灾害学, 2015, 30(1): 1-4. [Peng Guifeng and Liu Yingxi. Study on Sudden Geological Disaster Precipitation Factors Evaluation Based on Fuzzy Information Probability Interval Numbers [J]. Journal of Catastrophology, 2015, 30(1): 1-4.]

基于模糊信息概率区间数的突发地质灾害降水因子评价研究^{*}

彭贵芬¹, 刘盈曦²

(1. 云南省气象台, 云南 昆明 650034; 2. 云南财经大学, 云南 昆明 6502213)

摘 要: 根据 2010、2012 两年云南省的 820 个地质灾害监测记录, 基于模糊信息概率区间数的大小来开展降水量因子的评价研究。结果表明: 突发地质灾害发生前 10 d 的单日降水量与灾害的关系很小, 基本不能用于预警模型的建立和风险评价分析; 累积降水量与突发地质灾害关系明显, 特别是 5~10 d 的累积降水量预测概括率已达 60% 以上, 可以用于突发地质灾害预警模型的建立和实际的监测预警; 加衰减系数的有效累积降水量因子对突发地质灾害预测效果并无提升, 与普通累积降水量因子相比并无优势。

关键词: 模糊信息; 概率区间数; 突发地质灾害; 降水因子; 评价

中图分类号: X43; P468 **文献标志码:** A **文章编号:** 1000-811X(2015)01-0001-04

doi: 10.3969/j.issn.1000-811X.2015.01.001

对于滑坡泥石流崩塌之类的突发地质灾害来说, 得到精确灾害记录和直接致灾因子——降水量的原始数据是很困难的。首先, 突发地质灾害的记录不完整, 不是一出现突发地质灾害就有记录, 而是造成了人员、房屋、庄稼损失或影响了交通、水利设施安全的突发地质灾害才被记录; 其次, 发灾地通常在高山深谷, 这些自然环境十分恶劣的地方, 不一定能收集到降水量观测记录, 或者观测点远离发灾地, 能够获得的降水量不一定能代表发灾地的实际降水量。也就是说无论是预报对象还是预报因子, 都具有相当的不确定性特征, 这给两者关系的研究和预警模型的建立带来不确定性。针对这种不确定性, 一方面需要进一步改进观测手段收集更多的相关信息, 另一方面则需要改进分析方法, 以提高因子分析的科学性、准确性和实用性。

近年来对不确定性事件进行概率分析并建立概率预测模型已在地震^[1-2]、油田开发^[3]、航空装备战斗损伤^[4]、疾病复发^[5]、交通拥堵预测^[6]研究和应用中取得成果; 而概率区间数用于不确定性问题的分析研究和预测成果在以下方面取得: 针对属性权重难以用有效手段确定的多属性决策问题, 分析了概率区间和可行性区间的性质之后, 结合模糊综合评判将决策问题转化为区间数的排序问题^[7]; 条件属性值缺失时规则的不确定性度

量用概率区间来表述^[8]; 机器人研究领域利用概率区间对角色不确定性动作建模的方法^[9]; 研究了负二项分布的成功概率的区间估计, 给出了成功概率的精确置信区间、不依赖于大样本的近似置信区间以及依赖于大样本的近似置信区间^[10]; 基于专家风险概率区间估计的供应链风险综合评价模型^[11]; 不确定环境下定期人寿保险的破产概率区间的计算, 用泊松分布得到其近似计算方法^[12]; 基于概率区间的成都市大气 PM₁₀ 浓度的特征分析^[13]; 自然状态发生的概率不能完全确定, 但可以估计其所在区间的决策问题及其性质^[14]; 强降雨是滑坡、泥石流发生的重要因素已有较多的分析结果^[15-19]。本文根据 2010、2012 两年云南省地质环境监测院的地质灾害监测记录, 基于模糊信息概率区间数的大小来开展降水量因子的评价研究, 为突发地质灾害与降水这种不确定性关系问题的解决提供较为科学、准确的依据。

1 资料与方法

资料: 所用突发地质灾害资料为云南省 2010、2012 两年的 820 个地质灾害记录(由云南省地质环境监测院提供), 定位精确到村。降水量资料为云南省气象局信息中心提供的发灾县(市、区)气象站的日降水量和多种组合累积降水量(地质灾害发

^{*} 收稿日期: 2014-07-25

修回日期: 2014-09-22

基金项目: 国家自然科学基金项目(41165004); 云南财经大学科学研究基金(YC2013D11)

作者简介: 彭贵芬(1955-), 女, 云南昆明人, 研究员级高级工程师, 主要从事气象灾害研究. E-mail: gfpeng0725@163.com

通讯作者: 刘盈曦(1984-), 女, 云南昆明人, 博士, 讲师, 主要从事经济理论研究. E-mail: yingxi.liu.0711@gmail.com

生前当日、当日 + 前 1 d、当日 + 前 2 d、当日 + 前 4 d、当日 + 前 5 d、当日 + 前 9 d、带衰减率的当日 + 前 9 d)。

方法: 设 R 是实数集, $\underline{c}, \bar{c} \in R$ 且 $\underline{c} \leq \bar{c}$, 称闭区间 $[\underline{c}, \bar{c}]$ 为区间数, 并记作 $C = [\underline{c}, \bar{c}]$, 在因子评价分析中, 综合评价值是存在于区间 $[\underline{c}, \bar{c}]$ 的定值。在实际问题中, 区间数的综合评价在区间上往往并不服从均匀分布, 甚至也不服从标准正态分布、指数分布、对数分布等非均匀分布, 因此有必要提供各区间数的综合评价在该区间上的概率分布情况, 根据较为符合实际分布的概率分布函数的密度函数求取区间概率密度, 通过概率密度的比较来进行区间数的比较, 从而得出较为科学准确的综合评价结果。

2 突发地质灾害降水量概率区间数的计算

设带有综合评价值的概率分布函数 $F(x)$ 的区间数 C 为带概率分布区间数, 记作 $(C, F(x))$, 以下简记为 C 。根据突发地质灾害记录从降水量数据库中抽取出发灾县发灾前 10 d 逐日降水量和多时间尺度累积降水量 R , 如果 R 与地质灾害关系紧密, 则 R 值将集中在某区间内, 因此可以通过对 R 的概率分布区间数的比较分析, 得出 R 与突发地质灾害的相互关系, 找出各时间尺度降水量的分布区间, 可得出临界降水量指标。相互比较关系还能确定各时间尺度降水量与突发地质灾害的相关程度(因子的重要性), 用于确定预警模型中各因子的权重, 提高预警模型的精度。因此突发地质灾害降水量区间概率的计算分两个步骤进行: 第一步, 计算概率分布密度函数; 第二步, 计算区间概率密度。

2.1 基于模糊信息分布的概率密度计算

根据模糊信息分布概率能将因子值根据其距相邻等级中心点的距离而被模糊地分配到不同的区间中, 从而能得到事先不假定分布函数, 只根据实际数据按期间进行模糊信息分配, 从而得到 R 的总体概率分布估计的折线图或软直方图, 通过对小区间概率值的求和代替对概率密度函数的积分得到带概率分布区间数。因此基于模糊信息分布概率计算得到各模糊信息分布概率可按式(1)^[20-21] 计算得出。

$$q_{ij} = \begin{cases} 1 - \frac{|x_i - u_j|}{\Delta}, & |x_i - u_j| \leq \Delta; \\ 0, & \text{其他。} \end{cases} \quad (1)$$

式中: q_{ij} 是第 i 个指数值在第 j 个等级的模糊概率数; x_i ($i = 1, 2, \dots, n$) 是样本组中的第 i 个样本值; u_j 是某个等级区间的中点值; $\Delta = u_j - u_{j-1}$ ($j = 2, 3, \dots, m$) 是信息分配期间长度; m 为等级个数, 而

$\frac{q_{ij}}{\sum_{j=1}^m \sum_{i=1}^n q_{ij}}$ 则为第 i 个指数值在第 j 个等级的模糊概

率。取 $m = 50, \Delta = 5$, 即将地质灾害突发前各日降水量和多时间尺度累积降水量以 5 mm 为小区间, 分别计算 50 个小区间的模糊概率, 5 d、10 d 累积降水量的概率分布情况见图 1。突发地质灾害 5 d 的累积降水量分布在 15 ~ 175 mm 之间, 为主、次峰型分布, 主峰在 45 mm 附近, 分布概率为 5.49%, 次峰在 145 mm 附近, 分布概率为 3.59% (图 1a); 10 d 的累积降水量分布在 40 ~ 200 mm 之间, 也为主、次峰型分布, 主峰在 165 mm 附近, 分布概率为 5.34%, 次峰在 75 mm 附近, 分布概率为 4.32% (图 1b)。由两张概率分布图可见, 突发地质灾害前期降水所对应的概率分布较为复杂, 难以用通常的标准正态分布及指数分布、对数分布等非均匀分布函数来描述, 而采用模糊信息分配理论计算得出的概率分布情况能较为客观地描述实际的概率分布情况。

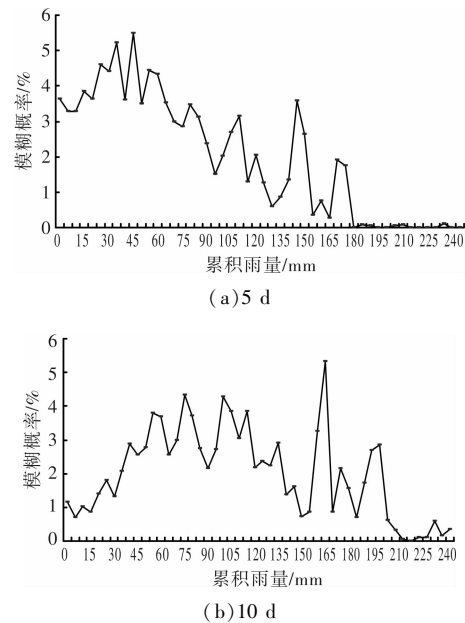


图1 突发地质灾害前 5、10 d 累积降水量概率分布图

2.2 基于模糊信息分配的概率区间数计算与估计

突发地质灾害是指由强降雨或连续性强降雨引发的滑坡、泥石流、崩塌等地质灾害, 先前的研究得到其与当地前 10 ~ 12 d 的降水量密切相关^[22-23], 而通过研究得出发灾当天到前 10 d 逐日及多时间尺度累积降水量与突发地质灾害的较为客观准确的定量关系是本文的目的。突发地质灾害是由降雨引发的, 即在某个时间段内, 降雨量越大发灾的风险越大, 以前人研究得出的日临界雨量 40 mm 为分界线, 将雨量分为两个区间 C_1 和 C_2 , $C_1 = [0, 40]$, $C_2 = [45, \infty]$, 计算前期降雨量的区间分布概率密度数 X , 以 X 来定量分析各尺度累积雨量与突发地质灾害的关系。对于突发地质灾害对应的降水量因子 R 来说, 可以将分布在 C_1 区间的概率数之和称为无效概率 $X_{\text{无效}}$, 意义为 40 mm 以下的降雨量与地质灾害的发生关系不大; 分布在 C_2 区间的概率数之和称为有效概率 $X_{\text{有效}}$, 意义为 40 mm 以上的降雨量可能引发地质灾害。降水量区间概率估计数与理想数接近度 D 按下式计算:

表 1 突发地质灾害前单日降水量区间概率数统计表

降水量 区间 /mm	单日降水量区间概率数 /%									
	R_0	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
0	26.62	36.15	47.18	43.27	45.6	51.61	46.36	55.5	54.01	57.12
5	12.73	14.26	17.2	21.18	16.67	15.74	18.1	15.54	17.65	15.85
10	8.98	8.48	10.66	11.21	7.9	9.09	8.21	8.58	7.58	7.61
15	9.27	8.17	6.07	5.59	8.08	5.16	11.69	7.32	4.9	4.85
20	8.61	6.34	5.02	6.12	6.6	5.32	4.81	3.63	5.89	4.74
25	5.51	4.6	3.52	3.5	4.85	4.71	3.59	1.45	3.38	3.79
30	2.46	5.8	2.08	1.55	1.78	1.99	1.43	2.5	1.94	1.63
35	2.29	2.35	1.35	1.74	1.83	0.97	1.26	2.32	1.24	1.36
40	2.59	4.78	2.53	2.14	1.8	1.34	1.2	0.97	0.9	0.52
0 ~ 40($X_{\text{有效}}$)	79.06	90.93	95.61	96.3	95.11	95.93	96.65	97.81	97.49	97.47
> 40($X_{\text{无效}}$)	20.94	9.07	4.39	3.7	4.89	4.07	3.35	2.19	2.51	2.53

$$D = \begin{cases} \frac{X_{\text{有效}}}{X} (\text{越接近 } 100\% \text{ 越好}); \\ \frac{X_{\text{无效}}}{X} (\text{越接近 } 0 \text{ 越好}). \end{cases} \quad (2)$$

本问题中的降水量在整个分析区间中的 X 为 100%, 而 $X_{\text{有效}}$ 和 $X_{\text{无效}}$ 为两个区间的区间概率数, 单位为 %, 所以可以直接用 $X_{\text{有效}}$ 和 $X_{\text{无效}}$ 作为因子的接近度。做因子评价时, $X_{\text{无效}}$ 小、 $X_{\text{有效}}$ 大说明指标对突发地质灾害敏感性好。

3 基于模糊信息区间概率数的致灾因子分析及评价

3.1 单日降雨量因子分析及评价

将发灾当天的降雨量记为 R_0 , 前灾前 n 天的降雨量记为 R_n , 表 1 所示为突发地质灾害前单日降水量区间概率数, 表 1 中单日的 $X_{\text{无效}}$ 值均在 75% 以上, $\min(X_{\text{无效}}) = 79.06\%$, $\max(X_{\text{无效}}) = 97.49\%$; $X_{\text{有效}}$ 均在 25% 以下, $\min(X_{\text{有效}}) = 2.19\%$, $\max(X_{\text{有效}}) = 20.94\%$; 说明单日降雨量因子对突发地质灾害不敏感, 虽然发灾当天的降雨量与地质灾害关系最大, 也仅能概括 21% 的突发地质灾害, 因此, 单日降雨量基本不能用于预警模型的建立和风险评价分析。

3.2 累积降雨量因子分析及评价

用 R_{0-1} 、 R_{0-2} 、 R_{0-3} 、 R_{0-4} 、 R_{0-9} 分别代表发灾当天 + 前 1 d、发灾当天 + 前 2 d、发灾当天 + 前 3 d、发灾当天 + 前 4 d 和发灾当天 + 前 9 d 的累积降水量, 表 2 中的数为各时间尺度累积降水量的模糊概率区间数。由表 2 可见, 地质灾害突发当日的 $X_{\text{无效}}$ 为 79.03%, 在各时间尺度中数值最大; 而 $X_{\text{有效}}$ 为 20.97%, 在各时间尺度中数值最小, 说明地质灾害突发当日的单日雨量与突发地质灾害关系较差, 如果以 40 mm 单日降水量作为临界指标的话, 只能概括约 21% 的地质灾害, 预警效果很差。随着累积日数的增加, $X_{\text{有效}}$ 逐渐增大, $X_{\text{无效}}$ 逐渐减小, 有明显的准线性增大或减小的态势。 R_{0-3} 以上的 $X_{\text{有效}}$ 已在 50% 以上, R_{0-4} 的 $X_{\text{有效}}$ 增大到 64.5%、 $X_{\text{无效}}$ 减小到 35.5%, 说明地质灾害发生前 5 d (包括当日) 的累

积降水量预警概括率已达约 65%, 单指标已有一定的预警价值。而 R_{0-9} 的 $X_{\text{有效}}$ 达到 86.77%、 $X_{\text{无效}}$ 减小到 13.23%, 说明地质灾害发生前 10 d (包括当日) 的累积降水量预警概括率高达 87%, 单指标已有相当的预警价值, 这个结果说明地质灾害突发前单日降水量对地质灾害并不敏感, 多时间尺度累积降水量对突发地质灾害要敏感得多, 而 5 d 以上的累积降水量因子是突发地质灾害预警建模及风险分析评价要考虑的主要因子。

3.3 累积降水量与有效累积降水量因子分析及评价——以 10 d 累积降水量为例

10 d 累积降水量由发灾当天的降水量与发灾前 9 d 的降水量进行简单的求和得出, 而 10 d 有效累积降水量则为加上衰减系数后的累积降水量, 按下式计算得出^[24]:

$$R_{10\text{有效}} = R_0 + R_1 + R_2 + \sum_{i=3}^9 \alpha^{i-2} R_i \quad (3)$$

式中: α 为衰减系数, 通常取 0.84; $R_0, R_1, R_2, \dots, R_9$ 代表发灾当天、前 1 d、前 2 d、……、前 9 d 的降水量。有效累积降水量是按时间越接近发灾日降水量与灾害发生的关系越大的思路来设计的, 理论上来说应该是更科学、更有效的。表 2 最后 1 列是用式 (2) 对 820 个地质灾害发生前 10 d (包括当天) 单日降雨量计算得到的有效累积降水量 $R_{0-9\text{有效}}$ 的区间概率数, 用其做预警因子应该比 R_{0-9} (表 2 倒数第二列) 相比更有效, 但计算结果却得出了相反的结论。 $R_{0-9\text{无效}}$ 在 C_1 区间的概率数为 23.3%, R_{0-9} 则为 13.2%;

表 2 突发地质灾害前累积降水量区间概率数统计表

降水量 区间 /mm	累积降水量区间概率数 /%						
	R_0	R_{0-1}	R_{0-2}	R_{0-3}	R_{0-4}	R_{0-9}	$R_{0-9\text{有效}}$
0	26.53	13.95	7.52	4.87	3.64	1.15	1.35
5	12.75	8.14	5.78	4.14	3.29	0.7	1.38
10	8.99	7.53	5.7	4.36	3.27	1.01	1.77
15	9.28	9.44	8.71	6.78	3.85	0.86	2.22
20	8.62	8.08	7.34	7.2	3.62	1.41	2.85
25	5.52	5.58	5.53	5.17	4.6	1.81	2.74
30	2.46	4.4	5.2	4.31	4.41	1.32	3.42
35	2.29	3.61	5.35	4.02	5.21	2.08	4.68
40	2.59	4.35	3.99	4.92	3.61	2.89	2.89
0 ~ 40($X_{\text{无效}}$)	79.03	65.08	55.12	45.77	35.5	13.23	23.3
> 40($X_{\text{有效}}$)	20.97	34.92	44.88	54.23	64.5	86.77	76.7

R_{0-9} 有效 在 C_2 区间的概率数为 76.7%, R_{0-9} 则为 86.8%; 计算结果说明用 R_{0-9} 有效 ≥ 40 mm 作为临界雨强预测突发地质灾害的概括率为 76.7%, 而 $R_{0-9} \geq 40$ mm 作为临界雨强概括率可达到 86.8%, 也就是说直接用 10 d 的累积降水量比用 10 d 的有效累积降水量做突发地质灾害预警概括率高 10% 左右, 说明加衰减系数的有效累积降水量因子对突发地质灾害预测效果并无提升, 与普通累积降水量因子相比并无优势, 将其用于监测、预警及评估要谨慎。

4 结论

(1) 突发地质灾害发生前 10 d 的单日降雨量与灾害的关系很小, 基本不能用于预警模型的建立和风险评估分析。

(2) 累积降水量与突发地质灾害关系明显, 特别是 5 ~ 10 d 的累积降水量预测概括率已达 60% 以上, 可以用于突发地质灾害预警模型的建立和实际的监测预警。

(3) 加衰减系数的有效累积降水量因子对突发地质灾害预测效果并无提升, 与普通累积降水量因子相比并无优势, 将其用于监测、预警及评估要谨慎。

参考文献:

- [1] 苏有锦, 李忠华. 云南地区 6 级以上强震时间分布特征及其概率预测模型研究[J]. 地震研究, 2011, 34(1): 1-7.
- [2] 李俏, 逯绍峰. 辽宁地区 5 级以上地震复发间隔及地震概率预测模型研究[J]. 防灾减灾学报, 2012, 28(2): 29-33.
- [3] 钟仪华, 徐瑶, 王昱. 油田开发指标系统的概率模拟预测[J]. 数学的实践与认识, 2011, 41(10): 80-86.
- [4] 杨新广, 陈云翔, 费文. 航空装备战斗损伤概率预测模型研究[J]. 数学的实践与认识, 2011, 41(18): 118-122.
- [5] 马获, 赵振娟, 张金丽. 经皮冠状动脉介入治疗患者术后 2 年心脏不良事件发生概率预测模型的建立[J]. 护理学报,

- 2013, 20(10): 33-35.
- [6] 郑炜, 张科, 汪芳. 概率预测的网络数学模型及其应用[J]. 计算机工程与应用, 2009, 45(18): 59-61, 69.
- [7] 李铭洋, 曹萍萍. 基于可行性区间数的模糊综合决策模型[J]. 微计算机信息, 2009, 25(16): 284-286.
- [8] 任文隆, 孙宏宾, 职林, 等. 论企业价值评估的价值概率区间[J]. 经济论坛, 2009, 23(11): 90-93.
- [9] 李萍, 瞿英, 王芳, 等. 不完备信息系统中决策规则的不确定性表示[J]. 计算机科学, 2009, 36(9): 193-195.
- [10] 田尊华, 赵龙, 贾焰. 基于概率区间的不确定性动作建模研究[J]. 计算机科学, 2010, 37(1): 201-203, 207.
- [11] 王志祥. 负二项分布参数的区间估计[J]. 统计与决策, 2010, 26(24): 35-37.
- [12] 万中, 胡朝明, 尹伟. 不确定环境下定期人寿保险的破产概率区间的计算[J]. 应用概率统计, 2011, 27(4): 391-398.
- [13] 丁辉. 基于概率区间的成都市大气 PM_{10} 浓度的特征分析[J]. 安徽农业科学, 2011, 39(27): 16836-16842.
- [14] 高峰记, 王海英. 基于优先度的概率区间型决策[J]. 统计与决策, 2012, 28(4): 85-87.
- [15] 李向红, 赵洁妮, 伍静, 等. 桂林“5.9”山体滑坡的暴雨成因分析[J]. 灾害学, 2013, 28(3): 95-99.
- [16] 闵颖, 胡娟, 李超, 等. 云南省滑坡泥石流灾害预报预警模型研究[J]. 灾害学, 2013, 28(4): 216-220.
- [17] 马超, 胡凯衡, 赵晋恒, 等. 震后泥石流的激发雨量特征——以汶川地震和集集地震后泥石流为例[J]. 灾害学, 2013, 28(4): 89-94.
- [18] 铁永波, 周洪福, 倪化勇. 西南山区短时强降雨诱发型低频泥石流成因机制分析——以四川省宝兴县冷木沟泥石流为例[J]. 灾害学, 2013, 28(4): 110-113, 187.
- [19] 雷发洪, 胡凯衡, 胡云华, 等. 汶川地震灾区震后泥石流激发雨量研究[J]. 灾害学, 2014, 02: 199-203.
- [20] 黄崇富. 自然灾害风险评价——理论与实践[M]. 北京: 科学出版社, 2006.
- [21] 彭贵芬, 张一平, 赵宁坤. 基于信息分配理论的云南干旱风险评估[J]. 气象, 2009, 35(7): 79-86.
- [22] 陈静静, 姚蓉, 文强, 等. 湖南省降雨型地质灾害致灾雨量阈值分析[J]. 灾害学, 2014, 29(2): 42-47.
- [23] 彭贵芬, 段旭, 舒康宁, 等. 应用 KDD 技术分析气象地质灾害与降水的关系[J]. 气象科技, 2007, 35(2): 252-257.
- [24] 万石云, 李华宏, 胡娟. 云南省滑坡泥石流灾害危险区划[J]. 灾害学, 2013, 28(2): 60-64.

Evaluation on Sudden Geological Disaster Precipitation Factors Based on Fuzzy Information Probability Interval Numbers

Peng Guifeng¹ and Liu Yingxi²

(1. Meteorological Observatory of Yunnan Province, Kunming 650034, China; 2. Yunnan University of Finance and Economics, Kunming 650221, China)

Abstract: According to 820 monitoring records of geological disasters in 2010 and 2012 in Yunnan Province, study on evaluation of precipitation factors is carried out based on fuzzy information probability interval numbers. Results show that the daily rainfall 10 days before has weak relationship with disasters, so basically it cannot be used to build forewarning modes and in risk evaluation or risk analysis; accumulated rainfall has strong relationship with sudden geological disasters, especially that of the 5 - 10 days has reached prediction generalization rate of more than 60%, so it can be used to build forewarning modes and to monitor and to warn early sudden geological disasters practically; effective accumulated rainfall factor added with attenuation coefficient has no effect on improving prediction of sudden geological disasters, so it has no advantage compared with general accumulated rainfall factor.

Key words: fuzzy information; probability interval number; sudden geological disaster; precipitation factor; evaluation