

阵性泥石流的周期性分析

胡凯衡, 李泳

(中国科学院水利部成都山地灾害与环境研究所, 四川 成都 610041)

摘要:对东川蒋家沟发生的一次阵性泥石流序列进行周期性分析。然后根据分析结果,提出了泥石流发生相关的几个问题。

关键词:阵性泥石流;时间序列;隐含周期;周期图

中图分类号:0211.61 X4 **文献标识码:**A

对具体的泥石流沟而言,泥石流的发生、发展不仅是一个时间过程,而且带有一定的周期性。这一方面是由于影响泥石流形成的自然条件的周期性,更重要的则可能是由于泥石流运动的自身特性,因此泥石流的周期性分析也就具有特别的意义。泥石流是在松散物质、水量、坡度等外在条件成熟的情况下才发生的。我们把每一次条件的成熟比做一次策动。假如外界电流是周期变化的,则脉冲的变化也有周期性。但是,对于一个区域或某个具体流域来说,我们掌握的关于泥石流发生的序列资料是极其有限的,不足以建立足够长的分析序列。不过,我们有较长序列的阵性泥石流的时序资料。对于泥石流这样一种自然灾害,人们往往更关心其发生的时间,而限于发生的过程或性质。我们可以把一次阵性流看作一次突发事件,把事件本身的时间略去,抽象为时间轴上的一个点。这样整个序列就成为轴上的一系列点。相邻两点间的距离就是间歇值。通过研究间歇值的变化规律,我们有可能对下一次突发事件的坐标作出预报。当然也可以应用离散随机过程的思想,以间隔值为随机变量,把时间序列看作随机过程的一个现实,计算各种统计特征量来分析数据中的随机成分。

野外观测数据或实验数据大部分是依次在相继的时刻点上获取的。比如某次泥石流在 $t_1 < t_2 < \dots < t_{N-1} < t_N$ 时刻的流量分别为 $X(t_1)$ 、 $X(t_2)$... $X(t_n)$... $X(t_{N-1})$ 、 $X(t_N)$ 。这种以时间 t 为自变量的有序数集合 $\{X(t_n)\}$ 在数学上称为时间序列。象月降水量、河流流量、某个国家的人口总数都是这样的时

间序列。时间序列分析是数据分析的重要内容。回归分析,相关分析,平稳性分析,周期性分析,频谱分析都是时间序列分析的重要方法。对时间序列的动态分析能够在一定程度上反映系统的动力学性质。粗略地讲,时间序列有点类似简单的离散动力学过程。近年来,随着复杂性现象研究的深入,时间序列在研究混沌现象本质方面也得到重要的应用。

1 分析

我们选取 1989-8-3 13:18 在蒋家沟发生的第九场泥石流(编号为 8909)的阵性序列作周期性分析。因为这场泥石流历时长(9小时2分),阵次多(达 140 次),过程比较完整,原始数据见表 1^[1]。当然该序列不是一般所谓的时间序列。因为它的下标并不表示时间,而是阵次。但是由于阵次代表着在顺次相继的时刻点上发生的阵性流,因此可以把它看作广义的时间序列,在混沌时间序列分析中,把它称为尖峰时间序列。这种序列是一维非负的标量序列。并且数据是不可再现的。

1.1 数据的预处理

由于人为或仪器的影响,原始数据与真实值之间总会存在误差。误差较大时,必然会使数据分析的结果产生偏差。因此应该依据统计学的原理剔除序列中的数值异常点,而代之以内插值。常用的方法有 3δ , $53H$ 法^[2]。另外移动平均法^[2,3]也可用来平滑数据,消除或微弱一些随机干扰。我们现在使用的观测数据为时间值,可以认为实际偏差不大。

收稿日期:2000-05-22;改回日期:2000-12-08。

基金项目:中国科学院重大项目 B“中国山区城镇泥石流综合防治技术研究”(编号 ZK951-B₁-202)资助。

作者简介:胡凯衡(1975-),男(汉族),江西瑞金人,1997年毕业于四川大学数学系,现为山地所硕士研究生,主要从事地理过程的计算机模拟研究。

表 1 同一场泥石流相邻两次阵性流之间的时间间隔

Table 1 Time intervals between two consecutive occurrences in a spasmodic series of debris flow

阵次	间隔值(s)	阵次	间隔值(s)	阵次	间隔值(s)	阵次	间隔值(s)	阵次	间隔值(s)
1	21	29	106	57	87	85	159	113	54
2	70	30	55	58	36	86	91	114	69
3	72	31	83	59	100	87	156	115	128
4	51	32	114	60	67	88	113	116	94
5	118	33	101	61	97	89	154	117	59
6	121	34	39	62	135	90	94	118	54
7	57	35	11	63	157	91	76	119	73
8	28	36	65	64	131	92	177	120	62
9	44	37	46	65	48	93	29	121	55
10	125	38	24	66	58	94	78	122	82
11	47	39	44	67	105	95	140	123	76
12	114	40	74	68	73	96	19	124	23
13	152	41	58	69	104	97	55	125	92
14	84	42	7	70	155	98	132	126	67
15	48	43	73	71	75	99	89	127	5
16	53	44	74	72	33	100	89	128	26
17	86	45	58	73	151	101	31	129	49
18	172	46	136	74	129	102	65	130	69
19	46	47	49	75	63	103	161	131	48
20	81	48	28	76	99	104	83	132	67
21	58	49	56	77	66	105	112	133	49
22	148	50	128	78	54	106	67	134	90
23	100	51	100	79	127	107	37	135	32
24	154	52	71	80	112	108	71	136	23
25	48	53	89	81	98	109	32	137	57
26	45	54	29	82	82	110	144	138	118
27	51	55	109	83	93	111	61	139	82
28	121	56	87	84	51	112	121	140	19

注:表中的间隔值由下一次阵性流龙头时间减去上一次龙尾时间得到。

所以只对原始数据作移动平均处理,在实际计算中发现当阶数取到 5 时效果最好。具体做法是对原始序列中 $2 \times 5 + 1$ 个相邻数据求其平均值,以得到的平均值作一新序列 $\{X'(t_n)\}$ (图 1)。如新序列的第一项

$$X'(t_1) = \frac{1}{2 \times 5 + 1} \sum_{i=1}^{11} X_i t$$

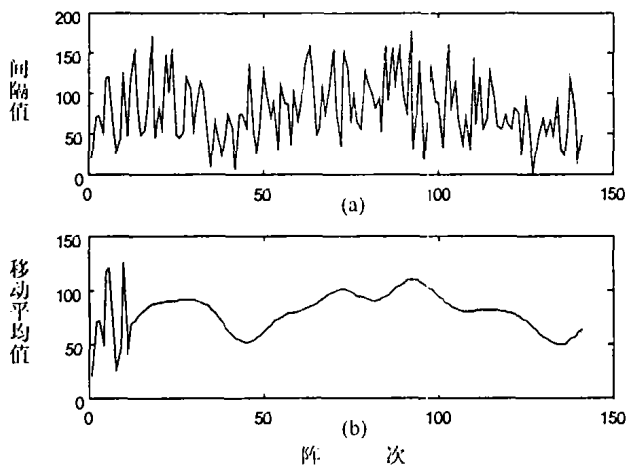


图 1 (a)原始序列;(b)5 阶移动平均后得到的新序列
Fig.1 (a)Original series, (b)new series by using moving average

第二项

$$X'(t_2) = \frac{1}{2 \times 5 + 1} \sum_{i=2}^{11} X_i t$$

如此类推。

1.2 去趋势项(主值项)

在一般的处理中,将时间序列为 $X(t)$ 分为三部分

$$X(t) = f(t) + p(t) + a(t)$$

其中 $f(t)$ 为主值项, $p(t)$ 为周期项, $a(t)$ 为随机干扰项。

$f(t)$ 代表了序列的递增或递减的趋势。在提取隐含周期前必须将趋势项分离出来。趋势项一般用多项式、指数线、双曲线等拟合,系数可用逐步线性回归分析获得。对上面所得的新序列 $\{X'(t_n)\}$ 分离主值项,经逐步回归计算,发现各项系数非常小,以至得到的 $f(t)$ 近于一条直线。所以把主值视为零,直接进行周期性分析。

1.3 周期性分析

周期性分析的主要思想是通过傅立叶变换把数

据从时域分析变为频域分析,再由傅立叶系计算周期图。然后进行统计检验,提取隐含周期。周期图的定义和计算公式说明如下:

对时间序列 $X(t)$ 进行傅立叶展开

$$X(t) = a_0 + \sum_{k=1}^M (a_k \cos(\frac{2\pi kt}{N}) + b_k \sin(\frac{2\pi kt}{N})),$$

$$M = [\frac{N}{2}], \quad ([\frac{N}{2}] < \frac{N}{2} \text{ 的最小整数}), t = 1, 2, \dots, N$$

(N 为时间序列的总长度)。

其中系数

$$a_0 = \frac{1}{N} \sum_{i=1}^N x_i, \quad a_k = \frac{2}{N} \sum_{i=1}^N x_i \cos(\frac{2\pi kt}{N}),$$

$$b_k = \frac{2}{N} \sum_{i=1}^N x_i \sin(\frac{2\pi kt}{N}), \quad k = 1, 2, \dots, [(N-1)/2]$$

当 N 为偶数时

$$a_{N/2} = \frac{1}{N} \sum_{i=1}^N x_i \cos(t\pi) = \frac{1}{N} \sum_{i=1}^N (x_{2i} - x_{2i-1})$$

$$b_{N/2} = 0$$

对每一频率 $f_k = \frac{k}{N}$ 定义周期图^[4]

$$I(f_k) = N/2(a_k^2 + b_k^2) \quad k = 1, 2, \dots, [(N-1)/2]$$

当 N 为偶数时

$$I(f_{N/2}) = Na_{N/2}^2$$

以下简记为 I_k 。周期图反映相应周期对序列贡献量的大小。由以上公式计算序列的周期图表2。

如果序列存在某个隐含周期,那么与该周期对应的周期图会出现峰值。但周期图上的峰值是否真正是序列的隐含周期,还必须进一步用统计检验的办法来判定。

参照 Fisher 统计检验方法^[4],先计算

$$g_k = I_k / \sum_{j=1}^M I_j, \quad I_{jk} \text{ 为 } \{I_1, I_2, I_3, \dots, I_k\}$$

中第 k 大的值。由 g_k 再计算

$$g' = g_1, g_2/(1-g_1) \dots g'_k = g_k/(1 - \sum_{j=1}^{k-1} g_j)。$$

然后计算相应的 Fisher 概率值

$$P_k = P\{g > g'_k\} = \sum_{j=0}^r (-1)^j C_k^{j+1} \times [1 - (j+k)g'_k]^{k-1},$$

其中 k 为周期图的个数(在本例中 $k=65$), r 是使 $1 - (r+k)g_k > 0$ 成立的最大正整数。计算的结果列在表3。

给定显著水平 α ,若有 $P_k < \alpha$ 成立,那么可以接受 $T_k = N/j_k$ 为序列的一个周期。

表2 序列的周期图 ($\times 10^4$)

Table 2 Periodogram of the new series ($\times 10^4$)

I_1	1.1171	I_{14}	0.0085	I_{27}	0.0043	I_{40}	0.0016	I_{53}	0.0016
I_2	0.9214	I_{15}	0.0102	I_{28}	0.0133	I_{41}	0.0013	I_{54}	0.0000
I_3	0.5842	I_{16}	0.0111	I_{29}	0.0071	I_{42}	0.0044	I_{55}	0.0007
I_4	0.3646	I_{17}	0.0010	I_{30}	0.008	I_{43}	0.0069	I_{56}	0.0034
I_5	0.0675	I_{18}	0.0029	I_{31}	0.0018	I_{44}	0.0009	I_{57}	0.0001
I_6	0.0713	I_{19}	0.0003	I_{32}	0.0004	I_{45}	0.0001	I_{58}	0.0003
I_7	0.0706	I_{20}	0.0043	I_{33}	0.0007	I_{46}	0.0003	I_{59}	0.0002
I_8	0.0019	I_{21}	0.0119	I_{34}	0.0005	I_{47}	0.0000	I_{60}	0.0004
I_9	0.0085	I_{22}	0.0025	I_{35}	0.0016	I_{48}	0.0000	I_{61}	0.0001
I_{10}	0.0125	I_{23}	0.0020	I_{36}	0.0008	I_{49}	0.0005	I_{62}	0.0014
I_{11}	0.0031	I_{24}	0.0001	I_{37}	0.0001	I_{50}	0.0010	I_{63}	0.0008
I_{12}	0.0019	I_{25}	0.0001	I_{38}	0.00303	I_{51}	0.0036	I_{64}	0.0029
I_{13}	0.0114	I_{26}	0.0022	I_{39}	0.0033	I_{52}	0.0025	I_{65}	0.0001

这里我们给定显著水平 $\alpha = 0.005$,从表3中可以看出 P_1 到 P_7 的值都 $< \alpha$ (前面标有星号者)。所以 $130/1, 130/2, 130/3, 130/4, 130/5, 130/6, 130/7$ 都应是序列的隐含周期。各周期分量所对应的系数分别为

$$a_0 = 81.1427,$$

$$a_1 = -11.8247,$$

$$a_2 = 3.7909,$$

$$a_3 = -4.2235,$$

$$a_4 = -7.1887,$$

$$a_5 = 0.9145,$$

$$a_6 = 3.2290,$$

$$a_7 = -1.0305$$

$$b_1 = -5.6608,$$

$$b_2 = 11.2862,$$

$$b_3 = 8.4876,$$

$$b_4 = 1.2734,$$

$$b_5 = 3.0808,$$

$$b_6 = -0.7354,$$

$$b_7 = 3.1295$$

从而得出周期项

$$P(t) = a_0 + \sum_{k=1}^7 (a_k \cos \frac{2\pi kt}{130}) + b_k \sin \frac{2\pi kt}{130}。$$

最后得到的随机干扰项 $a(t)$ 可作进一步处理。 $a(t)$ 可以认为是由影响泥石流运动的偶然因素产生的。假如将 $a(t)$ 视作平稳过程(即偶然因素影响的均值在每一阵次都相等)的话,那么可以建立 $a(t)$ 的自回归模型^[2,4],给出序列更精确的结果。但是序列的周期性分析做到这一步就足够了。

2 结论

通过对 8909 号和 8911 号等其他几场泥石流作类似的分析,发现序列的周期分量都集中在开头几个周期(130/1, 130/2, 130/3, 130/4, 130/5, 130/6, 130/7)。一个影响因素对应一个周期分量。这表明阵性流间歇期的周期性不是很突出。间歇期的周期性不明显,也可能暗示着松散物质和水的补给随机

性较大。至于这种周期特征与泥石流的动力机制是否有内在的关联,还需作进一步的研究。当然,只分析同一场泥石流多阵次的周期性意义不是很大。但是假如大尺度时空范围内发生的泥石流自相似于一多阵次泥石流,那么就有可能从一次泥石流运动特征推断出整个流域内长时间泥石流暴发的规律。

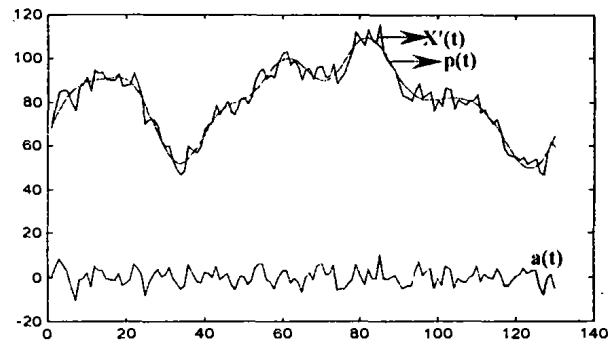


图2 序列 $X'(t)$ 、周期项 $p(t)$ 和随机干扰项 $a(t)$
Fig.2 New series $X'(t)$, periodic section $p(t)$ and random section $a(t)$

表3 Fisher 概率值
Table 3 Fisher probability value

K	g_k	P_k	K	g_k	P_k	K	g_k	P_k
1	0.3331	* 0.0000	23	0.0691	0.9527	45	0.1376	0.7821
2	0.4120	* 0.0000	24	0.0716	0.9412	46	0.1426	0.7811
3	0.4443	* 0.0000	25	0.0708	0.9578	47	0.1574	0.6913
4	0.4742	* 0.0000	26	0.0753	0.9303	48	0.1447	0.8527
5	0.1856	* 0.0000	27	0.0793	0.9029	49	0.1645	0.7419
6	0.2255	* 0.0000	28	0.0847	0.8514	50	0.1650	0.7944
7	0.2784	* 0.0000	29	0.0786	0.9401	51	0.1626	0.8628
8	0.0758	0.5199	30	0.0853	0.8893	52	0.1699	0.8661
9	0.0770	0.5152	31	0.0851	0.9104	53	0.2018	0.7158
10	0.0800	0.4730	32	0.0813	0.9561	54	0.2286	0.6199
11	0.0833	0.4280	33	0.0848	0.9444	55	0.2138	0.7906
12	0.0855	0.3524	34	0.0922	0.8964	56	0.1817	0.9708
13	0.0891	0.3646	35	0.0975	0.8622	57	0.2071	0.9423
14	0.0819	0.5433	36	0.0983	0.8788	58	0.2566	0.8242
15	0.0888	0.4201	37	0.1096	0.7831	59	0.3065	0.7081
16	0.0811	0.6042	38	0.1027	0.8845	60	0.3685	0.5836
17	0.0860	0.5270	39	0.01126	0.8096	61	0.3690	0.7455
18	0.0939	0.4048	40	0.0973	0.9611	62	0.5487	0.3677
19	0.0656	0.9450	41	0.1036	0.9421	63	0.5918	0.4998
20	0.0694	0.9145	42	0.1105	0.9170	64	0.5680	0.8640
21	0.0743	0.8611	43	0.1087	0.9482	65	1.0000	1.0000
22	0.0674	0.9571	44	0.1217	0.8862			

参考文献:

[1] 张军,熊刚. 云南蒋家沟泥石流运动观测资料集[M]. 北京:科学出版社,1997.
[2] 项静恬,史九恩,孔楠,等. 动态和静态数据处理[M]. 时间序列和数理统计分析,北京:气象出版社 1991.

[3] 张建中,周琴芳. 常用时间序列分析软件包[M]. 北京:气象出版社,1986.
[4] 中科院计算中心概率统计组. 概率统计计算[M]. 北京:科学出版社,1979.

Periodic Analysis of Intermittent Debris Flow

HU Kai-heng and LI Yong

(*Institute of Mountain Hazards and Environment, Chinese Academy of Sciences & Ministry of Water Conservancy, Chengdu 610041 PRC*)

Abstract: The article intends to disclose the periodic character of intermittent debris flow. We selected a series of observed data that is long enough to be analyzed. The value of element is the interval between two intermittences and the subscript of series is the number of the intermittence, which is different from traditional time series. Random disturbance was weakened and long time tendency was emphasized by using moving average process. In order to analysis periodic character, discrete Fourier transform is used to convert the series from time domain to frequency - domain. Then the value of periodogram was calculated by Fourier coefficient based on their definition. By Fisher statistical test with conspicuous level $\alpha = 0.005$ We obtain seven periods: 130/1, 130/2, 130/3, 130/4, 130/5, 130/6, 130/7 (130 is the length of the series). The same analysis to other observed data obtains similar results which implicit periods gather in the ahead location of periodogram. Some reasonable geographical explanations are given in the end of article: 1. Nature condition triggering debris flow is of periodicity, 2. The movement of debris flow is longly periodic, 3. The supply of loose material forming debris flow recovers slowly. It is possible that reason implies in the results as dynamical mechanics.

Key words: intermittent debris flow; time series; implicit period; periodogram