

泥石流固体堆积物粒度分布特征研究

徐新川, 陈剑平, 单 博

(吉林大学 建设工程学院, 吉林 长春 130026)

摘要:为客观准确研究泥石流固体堆积物粒度分布特征与泥石流影响因素之间的关系,通过对某水库库区的 23 条泥石流沟进行深入的调查研究分析,选取与泥石流流体性质有密切关联的 16 个泥石流影响因素,利用核主成分分析法(KPCA)对影响因素进行降维,形成线性无关的主成分。然后利用神经网络,预测了泥石流固体堆积物的分形维数。结果表明,网络预测误差最大仅为 2.6%,由此说明泥石流各影响因素间存在着复杂映射关系,具有良好的非线性相关性,能决定固体堆积物的粒度分布特征。

关 键 词:核主成分分析;神经网络;泥石流;粒度分布;分形维数

中图法分类号: P642

文献标志码: A

DOI:10.16232/j.cnki.1001-4179.2015.07.014

泥石流是产生于山坡与沟谷中的一种由液体和固体组成的两相流体,固相物质主要是由黏粒、粉粒、砂粒等不同粒径区段的土粒组成,液相物质主要是水。在泥石流运动中,浆体作为悬浮质将大小不等的块石及黏粒搬运到堆积区,形成具有一定粒度分布的松散堆积体。

目前,对泥石流堆积物粒度分析的方法大多都是基于统计学的规律和方法。但是运用传统的线性理论很难直观地描述泥石流固体堆积物的粒度分布特征。20 世纪 70 年代,美国数学家 B. B. Mandelbrot 首次提出的分形理论^[1],为研究泥石流堆积物的粒度分布特征提供了一种强有力的工具。李泳等通过研究发现,泥石流堆积物颗粒组成具有很好的分形特征^[2],颗粒的分维值随密度而变化,密度越大,分维值越大。倪华勇、刘希林等将泥石流堆积物的分形特征与堆积物的粒度特征相联系^[3],发现分维值可很好地反映泥石流堆积物颗粒组成及其粒度分布特征。易顺民分析了泥石流堆积物粒度分形结构特征,指出用分维值可以反映泥石流活动的形成演化特征^[4]。

然而将以粒径分形特性为主要内容的系统组构特征与泥石流沟谷特征、泥石流爆发特性及物源特征之

间关系联系起来进行研究的尚不多见。因此以颗粒流理论为基础,利用人工神经网络开展泥石流堆积物粒度分布特征与泥石流沟谷特征及泥石流爆发特性之间内在规律的研究具有一定的实际意义。

1 核主成分分析法

基于核函数的主元分析法(KPCA)通过将影响因素映射到一个高维空间后,对高维空间中的映射因素进行线性主元分析^[5-7],从而可以得到因素非线性主元。该方法抓住了主要关系对因素进行简化,实现了高维数据的降维,降维后的因素又能综合反映原始因素的综合信息。核主成分分析方法概述如下。

(1) 将泥石流样本数据映射到高维空间。设有 m 维的 n 个泥石流样本数据为 $x_1, x_2, \dots, x_n \in R^n$, 首先用一非线性映射 $\psi(\cdot)$ 把数据从原空间映射到高维空间,有 $\psi(x) = (\varphi(x_1), \varphi(x_2), \dots, \varphi(x_n))$, 然后在此高维空间进行线性主元分析。

(2) 计算泥石流各因素的协方差。特征空间中映射数据的协方差为

$$C = \frac{1}{n} \sum_{j=1}^n \varphi(x_j) \varphi(x_j)^T \quad (1)$$

收稿日期:2014-08-10

基金项目:国家自然科学基金资助项目(41330636)

作者简介:徐新川,男,硕士研究生,主要从事岩土工程地质方面的研究。E-mail:361977270@qq.com

通讯作者:陈剑平,男,教授,博士生导师,主要从事岩土地质方面的研究。E-mail:249531952@qq.com

(3) 计算泥石流因素主成分。一般通过转化为求解特征值问题来计算主成分,满足

$$\lambda V = CV \tag{2}$$

其中,特征值 $\lambda > 0$;特征向量为 $V \in \psi(\cdot)$ 。

由于特征值 $\gamma \neq 0$ 对应的特征向量 V 是由特征空间的向量组成的,所以存在

$$V = \sum_{i=1}^n \alpha_i \varphi(x_i) \tag{3}$$

其中, α_i 为系数。将(2)式两边左乘 $\varphi(x_k)$,可得

$$\lambda [\varphi(x_k) V] = [\varphi(x_k) C V] \tag{4}$$

其中, $k = 1, 2, \dots, n$ 。

定义 $n \times n$ 矩阵 K 为

$$K_{ij} = K(x_i, x_j) = [\varphi(x_i) \cdot \varphi(x_j)] \tag{5}$$

由式(3)~(5)可得(6)式,并求解核矩阵 K 的特征向量和特征值

$$n \lambda a = K a \tag{6}$$

其中, $a = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$ 的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, 其对应的特征向量为 $\alpha_1, \alpha_2, \dots, \alpha_n$ 。

归一化特征向量 V ,此时 $\varphi(x)$ 在 V 上的映射为

$$\begin{aligned} t_k &= [v_k \cdot \varphi(x)] = \sum_{i=1}^n \alpha_i^k [\varphi(x_i) \cdot \varphi(x_j)] \\ &= \sum_{i=1}^n \alpha_i^k K(x_i, x) \end{aligned} \tag{7}$$

其中, t_k 即为对应于 $\varphi(x)$ 的第 k 个非线性主成分分量。

(4) 特征值 λ_i 的主成分 t_i 是由噪声引起的,比值 $\lambda_i / \sum_{i=1}^n \lambda_i$ 反映了分量 t_i 对整体方差的贡献,较重要的分量对应较大的比值。主成分个数的选取原则为 $\sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i > E$, 其中, k 为选取的主成分个数; E 为累计方差贡献率,通常大于 85%。

2 BP 神经网络

常用的 BP 神经网络模型包括输入层、隐含层和输出层,每一层都由若干个神经元组成,每个神经元都有输入和输出,输入信号从输入层传到隐含层,然后传到输出层,每层神经元的输出只能影响下一层神经元的输出。

网络学习的步骤如下:① 对网络各层的权值和阈值赋予初值;② 输入学习样本;③ 从输入层开始经隐含层向输出层正向计算各层网络输出;④ 输出误差从输出层开始,将误差信号经隐含层传向输入层,反方向传播使网络误差最小;⑤ 正向传播计算与误差逆向传播计算反复交替;⑥ 考察全局误差是否达到精度要

求。

尽管 BP 神经网络得到广泛的应用,但在实际计算中,BP 算法收敛速度慢,并且目标函数存在局部极小值。因此,人们提出了各种改进算法。目前常用的算法有共轭梯度法、高斯牛顿法、动量改进 BP 算法、变速率算法、Levenberg - Marquardt 法等。

3 应用实例

3.1 粒度分布影响因子的选取

在泥石流的形成区或流通区,岩石由于风化作用形成了由磨圆和分选都很差及棱角分明的大量块石、细小颗粒及碎屑物质组成的松散体,这些松散体构成了泥石流的物质来源。松散体在大量流水的作用下,形成泥石流流体,随流水在泥石流沟内运动迁移,最终在堆积区堆积。泥石流流体的运移过程实际上就是其势能向动能转化的过程,期间颗粒发生摩擦、碰撞进行能量消耗,进而使堆积物颗粒的粒度分布发生改变。

影响泥石流沟的因素有地质构造、地层岩性、气象水文、植被及人类活动等。常用的影响因素有:流域面积 S_1 ,主沟长度 S_2 ,流域最大相对高差 S_3 ,主沟平均比降 S_4 ,形成区山坡平均坡度 S_5 ,流域切割密度 S_6 ,主沟床弯曲系数 S_7 ,松散固体物源储量 S_8 ,泥砂补给段长度比 S_9 ,植被覆盖率 S_{10} ,一次泥石流(可能)最大冲出量 S_{11} 和泥石流发生频率 S_{12} 。由于在不同的发展阶段,沟谷的平面形态亦不同,以致泥石流堆积物的粒度分布特征也不同,故本文引用了文献[8]中提到的沟谷形态要素 R_f 、圆度 R_c 和狭长度 R_e 。

3.2 粒度分布模型及分形维数计算

泥石流物质在机械搬运的过程中满足标度不变性,说明泥石流堆积物颗粒具有一定的分形特性^[9]。关于其分形值的大小,前人已经做了一些研究。据岩石破碎模型可知,泥石流堆积物中颗粒的尺寸和频度之间满足经验关系 Weibull 分布^[10]

$$\frac{N(< r)}{N_0} = 1 - \exp[-(\frac{r}{r_0})^b] \tag{8}$$

式中, r 为泥石流堆积物颗粒粒径; $N(r)$ 为所有尺寸小于或者大于 r 的颗粒质量之和; N_0 为整个颗粒集合的质量; r_0 为颗粒的平均尺寸; b 为常数。

倪华勇、易顺民等人采用岩石破碎模型,通过粒度分析法求解得到泥石流堆积体颗粒粒度分维值,计算式为

$$D = 3 - b \tag{9}$$

式中, D 为所求分形维数; b 为泥石流堆积物颗粒粒径和颗粒累积百分含量 $N(r)/N_0$ 在双对数坐标系中拟

合直线的斜率。

根据各个粒径颗粒的累积百分含量,在双对数坐标系中求出泥石流固体堆积物颗粒分布的分形维数。表 1 中的样本实测值即为通过试验求得的分形维数,绘制粒度分布曲线,其中一个如图 1 所示。统计所有粒度分布数曲线可知,相关性系数均值在 0.95 以上。强相关性说明,堆积物颗粒粒径与大于该粒径的累积百分含量在双对数曲线中有良好的线性关系,呈现出良好的分形特征。

表 1 学习样本及训练结果

沟名	网络输入值									样本	网络
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	期望值	预测值
下白滩	-11.512	4.081	-3.565	-1.897	-2.521	3.196	-0.909	-7.987	-1.274	2.403	2.409
上白滩	-15.790	-1.772	-7.249	-18.835	9.215	-11.253	-14.365	-7.442	6.672	2.458	2.467
猪拱地	-11.342	1.560	-4.878	-6.034	-8.029	0.070	-2.190	-1.917	-1.798	2.436	2.432
鲮鱼河	17.690	16.897	-8.793	24.211	6.558	-6.408	-10.005	9.927	2.125	2.605	2.600
秀水河	-10.871	26.399	57.096	-1.802	0.808	3.944	-5.454	0.201	-0.476	2.427	2.432
濠沽沟	-3.598	-0.872	0.784	2.941	-11.138	3.214	5.781	6.249	-4.109	2.392	2.407
架车河	2.664	-37.466	16.508	12.611	12.395	-0.049	11.185	-3.776	13.769	2.615	2.611
傅家沟	-9.191	3.694	-1.808	-1.064	-11.541	3.874	3.133	-1.398	-0.742	2.494	2.452
矮坝沟	-6.862	-17.818	3.642	-9.050	-20.989	-10.357	-2.257	11.117	7.419	2.475	2.479
盐水井	7.618	12.580	-12.351	14.261	3.202	13.453	-9.371	0.448	8.109	2.543	2.550
糯扎村	-4.732	6.426	-7.339	-1.013	1.281	6.261	5.613	-5.376	0.586	2.563	2.559
河漂沟	4.027	-48.890	7.128	13.058	-2.283	1.910	-11.294	-4.385	-12.624	2.403	2.439
红门厂	-10.359	2.689	-8.676	-6.558	4.593	2.524	2.039	-1.980	1.518	2.207	2.218
志力沟	-10.017	5.528	-6.754	-2.101	4.969	1.313	3.890	-1.464	-0.146	2.344	2.355
雅街德	1.787	5.084	-2.231	1.021	-13.410	-0.505	5.863	1.218	2.324	2.608	2.601
大乾沟	-8.490	-6.383	-1.701	-8.504	2.466	-2.203	0.993	4.202	-0.502	2.379	2.382
勐果河	7.269	-0.311	-2.331	5.084	7.105	6.310	4.283	6.643	-2.770	2.322	2.534
法帕沟	-4.850	17.766	0.434	18.805	4.747	-21.400	8.043	-5.128	-6.060	2.312	2.321

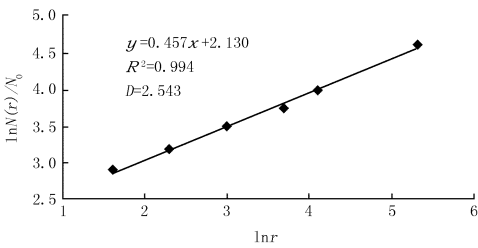


图 1 盐水井沟泥石流堆积物粒度分布曲线及分形维数

3.3 核主成分分析法计算

根据第 2 节所述步骤,输入 23 条泥石流沟的 16 个影响因素,限于篇幅,本文在此不列举。计算得相关

系数矩阵特征值如表 2 所示。可见前 9 个主成分的累计方差贡献值为 92.2%,超过了一般所要求的 85%,所以本文计算所得 F_1, F_2, \dots, F_9 可作为主成分,这 9 个主成分基本代替了原来的 16 个影响因素所反映的信息。表 2 为前 9 个主成分的特征向量,其代表了每个影响因素对主成分的贡献率大小,即荷载值。

表 2 各主成分特征值的贡献率

主成分	特征值	累计贡献率	主成分	特征值	累计贡献率
F_1	30.03	0.300	F_6	3.70	0.842
F_2	19.99	0.500	F_7	3.14	0.874
F_3	14.53	0.645	F_8	2.61	0.900
F_4	9.38	0.739	F_9	2.21	0.922
F_5	6.60	0.805			

3.4 BP 神经网络预测及结果讨论

由于 3.3 中的主成分很好地综合反映了各泥石流沟的信息,主成分取值如表 1 所示,采用前述的步骤对各泥石流沟固体堆积物粒度分维数进行预测。文中采用的 BP 神经网络为 3 层 BP 网络模型,并遵循 Kolmogorov 定理,设置 n 个输入单元、 $2n + 1$ 个中间单元及 m 个输出单元,这样的网络模型基本可以准确表达任何映射,并可确保网络学习及预测效率^[10-13]。

选择其中的 18 条沟作为样本进行神经网络测试,其它 5 条沟验证学习结果的收敛性。经过 2 000 次迭代后网络学习趋于成功,即神经网络中存储了样本所包含的内在复杂非线性映射关系^[14]。预测结果见表 3。

由网络预测结果可知,网络预测值与样本期望值之间的最大误差为 2.6%。在泥石流堆积物研究中,可通过核主元分析优化影响因素,然后借助训练好的神经网络,预测堆积区堆积物的分形维数,研究其粒度分布特征。

4 结论

(1) 本文通过核主元分析方法对泥石流影响因素降维,选取了对泥石流流体性质影响较大的 9 个主成分,消除了泥石流各影响因素间复杂线性相关性,并对各影响因素对泥石流堆积物粒度分布特征的贡献程度

表 3 网络预测结果及误差

沟名	网络输入值									样本	网络	误差/
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	期望值	预测值	%
竹鲜河	8.277	1.220	-7.243	1.982	-5.498	3.219	-2.571	0.434	-0.291	2.122	2.114	0.4
樟木河	-10.876	-5.987	-2.593	-18.112	22.388	1.146	2.778	12.893	-6.404	2.500	2.489	1.1
田房河	83.541	3.405	4.338	-17.244	-0.332	-2.378	1.729	-3.875	-1.118	2.549	2.482	2.6
方山果	-5.505	6.319	-7.858	1.873	-5.685	1.393	1.085	-3.169	-0.769	2.515	2.527	0.5
大庆沟	-8.879	5.850	-4.558	-3.633	1.699	2.724	1.999	-5.435	-3.439	2.618	2.594	0.9

注:样本期望值为通过试验所求出的分形维数,网络预测值即为 BP 神经网络经过误差还原后得到所求解的分形维数。

做出明确评价。

(2) BP 神经网络预测结果与试验所测值相比,最大误差值为 2.6%,误差较小,说明泥石流影响因素与分形维数间具有一定的非线性相关性,泥石流固体堆积物的粒度分布特征决定于泥石流的影响因素。

(3) 本次研究中未考虑地层岩性对堆积物粒度分布的影响,需作进一步的深入研究。

参考文献:

[1] Mabdekhrot B B. The fractal geometry of nature[M]. SanFrancisco: [s. n.],1983:1-3.

[2] 李泳,陈晓清,胡凯衡. 泥石流颗粒组成的分形特征[J]. 地理学报,2005,60(3):495-502.

[3] 倪华勇,刘希林. 泥石流堆积物的粒度分布及其分形结构[J]. 沉积与特提斯地质,2008,28(3):35-40.

[4] 易顺民. 泥石流堆积物的分形结构特征[J]. 自然灾害学报,1994,3(2):91-96.

[5] 蒋少华,桂卫华,阳春华,等. 基于核主元分析与支持向量机的监控诊断方法及其应用[J]. 中南大学学报:自然科学版,2009,40(5):204-208.

[6] Bernhard Scholkopf, Alexander Smola, Klaus - Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem[J]. Neural Computation,1996:1-18.

[7] 张吉萍,陈虬. BP 网络在边坡稳定性分析中的应用[J]. 西南交通大学学报,2001,36(6):649-650.

[8] 张文,陈剑平,秦胜武,等. 基于主成份分析的 FCM 法在泥石流分类中的应用[J]. 吉林大学学报:地球科学版,2010,40(2):367-372.

[9] 高召宁. 颗粒物质的分形特性与其物理力学性质的关系探讨[J]. 实验力学,2011,26(3):285-290.

[10] 陈颢,陈凌. 分形几何学[M]. 北京:地震出版社,1998.

[11] 张晨,王清,张文,等. 基于神经网络对泥石流危险范围的研究[J]. 哈尔滨工业大学学报,2010,42(10):1642-1645.

[12] 张晨,陈剑平,肖云华. 基于神经网络对有限元强度折减法分析[J]. 吉林大学学报:地球科学版,2009,39(1):114-118.

[13] 李发斌,崔鹏,周万村,等. 用遗传神经网络分析泥石流活动性[J]. 中国地质灾害与防治学报,2003,14(3):16-20.

[14] 夏元友,熊海丰. 边坡稳定性影响因素敏感性人工神经网络分析[J]. 岩石力学与工程学报,2004,23(16):2703-2707.

(编辑:郑 毅)

Study of grain distribution characteristics of solid accumulation in debris flow

XU Xinchuan, CHEN Jianping, SHAN Bo

(College of Construction Engineering, Jilin University, Changchun 130026, China)

Abstract: To objectively and accurately study the relationship between the grain size distribution characteristic of solid in debris flow and the influential factors of debris flow, 23 debris flow gullies in Wudongde Reservoir region are further investigated and analyzed. 16 debris flow influential factors that have closely related to the debris flow's fluid characteristics are chosen, and the kernel principal component analysis (KPCA) is used to reduce dimension of these influencing factors, and a linearly independent principal component is formed. The results show that largest prediction error is 2.6%, which demonstrates that the influential factors have complicated mapping relation with each other, also possess well nonlinear correlation.

Key words: kernel principal component analysis; neural network; debris flow; size distribution; fractal dimension

