

# 高程梯度性质和算法不确定性的初步分析

朱奇峰, 杨勤科, 龙永清, 王春梅, 庞国伟, 师 动

(西北大学 城市与环境学院, 西安 710127)

**摘 要:** 高程梯度指高程在平面坐标中  $x$  和  $y$  两个方向上的变化率。在坡度算法  $\theta = (p^2 + q^2)/2$  中分别用  $p$  和  $q$  表示, 因为高程梯度  $p$  和  $q$  算法的不同而衍生出多种坡度算法, 因此对  $p, q$  算法性质和精度的研究是坡度算法精度评价的基础。选取陕西省绥德县、秦岭、黑龙江省拜泉县的三个样区, 基于 7 种不同算法计算了高程梯度 ( $p$  和  $q$ ), 并对其统计分布特征、空间结构特征、 $p$  和  $q$  相互关系、算法稳定性与不确定性等进行了分析。结果表明: (1)  $p$  和  $q$  均值均为零, 相互独立, 方差相等, 呈正态分布; (2) 各种算法均能表现局地结构和样区之间的差异; (3) 从算法稳定性、精度等方面看, Evans-Young 算法, 特别是经过 Shary 改进的 Evans-Young 算法的效果最佳。

**关键词:** 高程梯度; 高程梯度算法; 统计分布; 算法稳定性

中图分类号: P208

文献标识码: A

文章编号: 1005-3409(2017)04-0019-08

## Preliminary Properties and Precision Analysis of Elevation Gradient Algorithm

ZHU Qifeng, YANG Qinke, LONG Yongqing, WANG Chunmei, PANG Guowei, SHI Dong

(College of Urban and Environment Science, Northwest University, Xi'an 710127, China)

**Abstract:** Elevation gradient is the ratio of the elevation in  $x$  &  $y$  direction of Cartesian coordinate system, usually they are denoted as  $p$  and  $q$  in the general algorithm of slope steepness, i. e.  $\theta = (p^2 + q^2)/2$ . The differences of varied slope algorithms are mainly due to the different algorithms of  $p$  and  $q$ , and therefore, the accuracy of slope steepness depends on the precision of  $p$  and  $q$  algorithms.  $p$  and  $q$  have been calculated and analyzed in seven  $p/q$  algorithms in three sample sites, including Suide and Qinling of loess hills in Shaanxi Province, Baiquan of undulate hill in Heilongjiang Province. The analysis includes statistical distribution and spatial pattern (or structure) of  $p$  and  $q$ , the interrelation of  $p$  and  $q$ , the stability and uncertainty of algorithms. The results show that: (1) the mean values of  $p$  and  $q$  are almost zero,  $p$  and  $q$  are independent from each other, the variance of  $p$  and  $q$  are equal, and they follow the pattern of normal distribution; (2) all algorithms can represent the local structure and the differences among the sample sites; (3) with respect to the algorithm stability and uncertainty, Evans-Young algorithm, especially the improved Evans-Young algorithm by Shary is the best.

**Keywords:** elevation gradient; elevation gradient algorithm; statistical distribution; stability of algorithms

高程梯度  $p$  和  $q$  分别指高程  $z$  在平面坐标系中  $x$  和  $y$  两个方向上的变化率<sup>[1-2]</sup>。作为坡度算法的基本计算项<sup>[1,3]</sup>, 因为高程梯度  $p$  和  $q$  算法的不同而衍生出多种坡度算法<sup>[1,4-6]</sup>。已有关于坡度算法的分析探讨<sup>[4,7-9]</sup>, 主要关注了坡度算法精度和不确定性特征。Skidmore 等<sup>[4-5]</sup> 比较了七种计算坡度和坡向的方法, 最终得出的结论是线性拟合算法或三阶反距离平方权差分算法精度最高。Zhou 等<sup>[6,10]</sup> 从高程梯度算法的误差来源来分析算法精度, 区分了误差的来源和性质, 认为三阶不带权差分算法 (即 Evans-Young 算法) 精度最

高。Florinsky 等<sup>[11-12]</sup> 比较了常见算法后认为 Evans-Young 算法较优。从总体上讲, 已有研究虽已注意到了  $p$  和  $q$  作为坡度坡向算法基本项的重要性, 但对于高程梯度 ( $p$  和  $q$ ) 的性质和统计分布特征未予关注。

本文选取陕西省绥德县、秦岭和黑龙江省拜泉县的三个样区, 对以上样区高程梯度 ( $p$  和  $q$ ) 进行初步分析, 重点研究各种算法下  $p$  和  $q$  的空间格局、统计分布、稳定性和不确定性等。通过研究, 得到高程梯度 ( $p$  和  $q$ ) 的性质和统计分布特征, 从而得到最优的  $p$  和  $q$  算法, 为更好地研究坡度特征提供支撑。

# 1 数据与方法

## 1.1 基础数据

1.1.1 基础数据说明 研究在三个样区展开,包括地势比较平缓的东北漫岗丘陵区,地势比较陡峭的黄土丘陵区 and 地势陡峭的秦岭山区。此三个样区均为侵蚀强烈地区,地形起伏由平缓到剧烈,对描述地形复杂程度的变化具有代表性,可以对比不同地形特征对于  $p$  和  $q$  算法的影响度。图 1 显示了三个样区 10 m 分辨率的 DEM,其中拜泉样区面积约为 549 km<sup>2</sup>。该样区位于小兴安岭与松嫩平原过渡带的漫岗丘陵地区,地势比较平坦。绥德样区面积约为 141 km<sup>2</sup>,位于陕北黄土丘陵区,地表破碎。秦岭样区面积约为 83 km<sup>2</sup>,该样区位于陕西南部秦岭山地区,地势陡峭,地形起伏剧烈。

本研究中三个样区使用的 10 m 分辨率 DEM 是基于 1:5 万地形图、使用 ANUDEM 软件生成的水文关系正确的 DEM (Hydrologically-correct DEMs)<sup>[13-14]</sup>, Hc-DEM 能更准确的表现侵蚀地形特征(流水线、分水线)和水文地貌关系特征<sup>[15]</sup>。另外,还使用了三个样区的原始数字线划地图数据(DLG)。

1.1.2 数据预处理 主要是目视检查是否存在高程异常值,然后做滤波除噪和填洼处理。(1)滤波:在 DEM 的建立过程中,会产生一些随机误差,将导致 DEM 表面有噪音存在,可以通过低通滤波来处理<sup>[16]</sup>。本研究中采用应用最广泛的高斯低通滤波对 DEM 进行预处理。(2)填洼:填洼的基本出发点是认为地表是一个连续光滑的表面,真实的洼地很少。但是由于各种原因(垂直精度限制等),在 DEM 表面依然存在一些伪洼地需要消除<sup>[16]</sup>。本研究中使用了 ArcGIS 中水文分析模块下的 Fill 工具进行填洼。

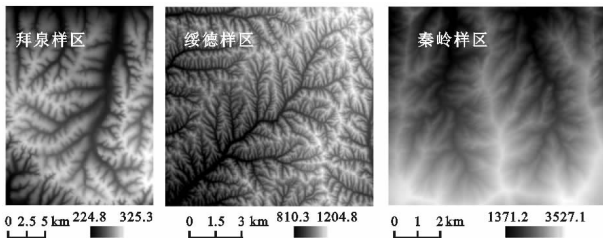


图 1 研究区 DEM 图

## 1.2 研究方法

1.2.1 高程梯度算法 高程梯度的常用算法主要包括 Evans-Young 算法<sup>[17]</sup>、Zevenbergen 算法<sup>[18]</sup>、Sobel 算法<sup>[19]</sup> 和 Florinsky 算法<sup>[12]</sup> 等。总结文献报道的算法,可归纳为 7 种,见表 1;式中  $3 \times 3$  矩阵  $z_i$  到  $z_9$  见图 2,  $5 \times 5$  矩阵情形类似。

$Z_1$	$Z_2$	$Z_3$
$Z_4$	$Z_5$	$Z_6$
$Z_7$	$Z_8$	$Z_9$

图 2 矩阵示意图

1.2.2 统计分析 本研究提取  $p$  和  $q$  后,计算  $p$  和  $q$  的基本统计参数,比较不同地形类样区和不同  $p$  和  $q$  算法的统计分布特征。基本统计参数有:最小值、最大值、平均值、标准差、众数、中位数、偏度、峰度等。平均值反映数据变化的整体趋势;标准差则反映数据的离散程度;众数是在  $p$  或  $q$  频率曲线的峰值对应的  $p$  或  $q$  的值;中位数是一组数据中间的数;偏度是指非对称分布的偏斜状态,用来测定参数分布的对称程度;峰度用来反映频数分布曲线顶端尖峭或扁平程度。

信息熵:本研究中使用 shannon<sup>[27]</sup> 提出的信息熵概念。高程梯度  $p$  和  $q$  的信息熵值越小,表明样区  $p$  和  $q$  的信息含量越小,表面越平滑。信息熵的计算公式如下:

$$H = - \sum p_i(x) \ln p_i(x) \quad (8)$$

式中:  $H$  为信息熵;  $p_i(x)$  为图像  $p$  或  $q$  值的频率;  $\ln p_i(x)$  为频率的自然对数。

1.2.3 高程梯度结构分析 结构指图像是否反映出高程变化率、并且在地区之间有差异。其分析可通过目视观察、局部方差 (local variance, Woodcock, 1987)<sup>[28]</sup> 和地统计学指标来分析。局部方差用公式 (9) 计算,局部方差平均值越大,表明数据表面的局部变化越大、表面越粗糙<sup>[28]</sup>。

$$LV_{i,j} = \frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n} \quad (9)$$

式中:  $LV_{i,j}$  为局部方差;  $x_k$  为分析窗口内第  $k$  个栅格的地形属性,即为  $p$  或  $q$  的值;  $\bar{x}$  为分析窗口内所有栅格的属性均值;  $n$  为分析窗口内的栅格个数。本研究中的分析窗口为  $3 \times 3$ , 即  $n=9$ 。

变异函数能够反映区域化变量的结构特征<sup>[29]</sup>。变程(Range)大小反映区域化变量自相关范围的大小,基台值(Sill)大小反映了区域化变量变化幅度的大小。变程越小,表明  $p$  或  $q$  自相关范围越小,  $p$  或  $q$  表面纹理越破碎;基台值越小,表明  $p$  或  $q$  的差异越小。

1.2.4 算法稳定性 在 DEM 的建立过程中,会存在一些随机误差,导致 DEM 表面存在噪音<sup>[16]</sup>。由于原始 Hc-DEM 已经滤波处理过,所以在本研究中对原始 Hc-DEM 上加上取值范围为 0 到 1,均值为 0.5 的随机噪音后(用来模拟原始数据随机误差)<sup>[2]</sup>,再得到其  $p$  和  $q$  图像,与未加噪音的  $p$  和  $q$  图像进行对比。

表 1  $p$  和  $q$  算法

算法名称	$p$ 和 $q$ 算法	编号	出处
1 简单差分算法	$p=\frac{z_5-z_4}{w}$	(1a)	Horn <sup>[19]</sup>
	$q=\frac{z_5-z_8}{w}$	(1b)	Jones <sup>[20]</sup>
2 二阶差分算法 (Zevenbergen 算法)	$p=\frac{z_6-z_4}{2\times w}$	(2a)	Zevenbergen and Thorne <sup>[18]</sup>
	$q=\frac{z_2-z_8}{2\times w}$	(2b)	
3 三阶差分算法 (Evans-Young 算法)	$p=\frac{z_3+z_6+z_9-z_1-z_4-z_7}{6\times w}$	(3a)	Evans <sup>[17]</sup> Young <sup>[21]</sup>
	$q=\frac{z_1+z_2+z_3-z_7-z_8-z_9}{6\times w}$	(3b)	
4 改进三阶差分算法 (改进 Evans-Young 算法)	$p=\frac{z_3+z_6+z_9-z_1-z_4-z_7}{6\times w}$	(4a)	Pennock <sup>[22]</sup> Sharpnack <sup>[23]</sup>
	$q=\frac{z_1+z_2+z_3-z_7-z_8-z_9}{6\times w}$	(4b)	
	$Z_e=\frac{z_1+z_6+41\times z_5+z_8+z_4}{45}$	(4c)	
5 三阶反距离平方权差 分算法(Sobel 算法)	$p=\frac{z_3+2z_6+z_9-z_1-2z_4-z_7}{8\times w}$	(5a)	Sobel <sup>[25]</sup>
	$q=\frac{z_1+2z_2+z_3-z_7-2z_8-z_9}{8\times w}$	(5b)	
6 三阶反距权差分算法 (Sharpnack 算法)	$p=\frac{z_3-z_1+\sqrt{2}\times(z_6-z_4)+z_9-z_7}{(4+2\times\sqrt{2})\times w}$	(6a)	Sharpnack <sup>[23]</sup>
	$q=\frac{z_7-z_1+\sqrt{2}\times(z_8-z_2)+z_9-z_3}{(4+2\times\sqrt{2})\times w}$	(6b)	
7 Florinsky 5 窗口算法	$p=\frac{1}{420\times w}\{44\times(z_4+z_{24}-z_2-z_{22})$ $+31[z_1+z_{21}-z_5-z_{25}+2(z_9+z_{19}-z_7-z_{17})]$ $+17[z_{15}-z_{11}+4(z_{14}-z_{12})]+5(z_{10}+z_{20}-z_6-z_{16})\}$	(7a)	Florinsky <sup>[12]</sup>
	$q=\frac{1}{420\times w}\{44\times(z_6+z_{10}-z_{16}-z_{20})$ $+31[z_{21}+z_{25}-z_1-z_5+2(z_7+z_9-z_{17}-z_{19})]$ $+17[z_3-z_{23}+4(z_8-z_{18})]+5(z_2+z_4-z_{22}-z_{24})\}$	(7b)	

式中: $p$  和  $q$  为高程梯度; $z_1$  到  $z_9$  分别为  $3\times 3$  窗口中各个像元的高程值,单位为 m; $w$  为像元的大小,本研究中  $w$  为 10 m。

由于添加噪音对  $p$  和  $q$  图像结构影响不大,所以仅仅在直方图相似指数(Histogram Similarity, HS)和误差和两个方面进行比较。直方图相似指数可用 L1 距离、L2 距离、L3 距离、卡方距离和直方图的交(即两个直方图共同的部分)等四种指标来度量<sup>[30]</sup>。本研究采用直方图的交来表示两条频率曲线的相似程度,其取值范围  $0\leqslant\text{HS}(X,Y)\leqslant 1$ 。HS 值越大,表示两幅图越相似,抗噪音能力越强。

$$\text{HS}(X,Y)=\sum_i\min(x_i,y_i)$$

(10)

式中: $\text{HS}(X,Y)$  为直方图相似度指数; $x$  和  $y$  分别为叠加随机噪音前和叠加随机噪音后的  $p$  和  $q$  频率曲线; $x_i$  和  $y_i$  为第  $i$  个分级间隔上叠加随机噪音前和叠加随机噪音后的频率值。

添加噪音后的误差用均方根误差(Root mean square error, RMSE)表示。将叠加噪音前的  $p$  和  $q$

图像数据作为标准,与叠加噪音后的  $p$  和  $q$  图像数据进行计算。

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^N(z_i-z_t)^2}{N-1}}$$

(11)

式中:RMSE 为均方根误差的值; $z_t$  和  $z_i$  分别为叠加随机噪音前和叠加随机噪音后的  $p$  和  $q$  图像数据; $N$  为图像数据栅格数。

1.2.5 算法不确定性 高程梯度计算中,DEM 上的误差会传递到  $p,q$  表面,根据误差传递共识推导的各种  $p,q$  算法中误差公式如下文公式(12)—(19)所示。

$$m_{p1}^2=m_{q1}^2=\frac{2\times m_z^2}{w^2}$$

(12)

$$m_{p2}^2=m_{q2}^2=\frac{m_z^2}{2\times w^2}$$

(13)

$$m_{p3}^2=m_{q3}^2=\frac{m_z^2}{6\times w^2}$$

(14)

$$m_{z_e}^2 = \frac{337 \times m_z^2}{405} \quad (15)$$

$$m_{p4}^2 = m_{q4}^2 = \frac{3 \times m_{z_e}^2}{16 \times w^2} \quad (16)$$

$$m_{p5}^2 = m_{q5}^2 = \frac{3 \times m_z^2}{16 \times w^2} \quad (17)$$

$$m_{p6}^2 = m_{q6}^2 = \frac{2 \times m_z^2}{(6 + \sqrt{2}) w^2} \quad (18)$$

$$m_{p7}^2 = m_{q7}^2 = \frac{3929 \times m_z^2}{17640 \times w^2} \quad (19)$$

式中:  $p$  或  $q$  后面的数字表示  $p$  或  $q$  的算法标号;  $m_z$  为 DEM 数据的中误差;  $w$  为像元大小, 本研究中的  $w$  为 10 m;  $m_p^2$  或  $m_q^2$  为高程梯度中误差, 本研究中  $m_p^2 = m_q^2$ ; 算法 4 中的  $m_{z_e}^2$  为滤波产生的误差。

## 2 结果与分析

### 2.1 高程梯度的结构特征

在 ArcGIS 中通过编程实现每种算法, 分别提取  $p$  和  $q$ , 得到拜泉样区、绥德样区和秦岭样区 7 种算法下的  $p$  和  $q$  的图像共 42 幅。图 3 以算法 4 得到的结果为例展示了  $p$  和  $q$  在不同样区的纹理, 其他算法得到的  $p$  和  $q$  纹理与其类似。可以看出: (1) 在不同样区之间,  $p$  和  $q$  取值范围随着地形起伏情况的变化而变化, 地形起伏越大, 其取值范围越大, 绝对值也越大, 图像的纹理也越复杂。因而各种算法均能表现局地结构和样区之间的差异; (2) 地统计学计算结果见表 2, 拜泉、绥德和秦岭三个样区的变程 (Range) 为 478.6~531.7 m, 113.4~115.2 m 和 230.4~407.7 m。这是由于绥德地处黄土丘陵区, 地表土质疏松, 现代侵蚀强烈; 秦岭虽然地壳抬升强烈, 但是地表多坚硬基岩, 因而切割小于黄土丘陵; 拜泉地势相对平缓, 侵蚀微弱。但是基台值 (Sill) 方面, 则是秦岭最大、拜泉最小, 绥德居中; 这一顺序与高程梯度值域相一致。

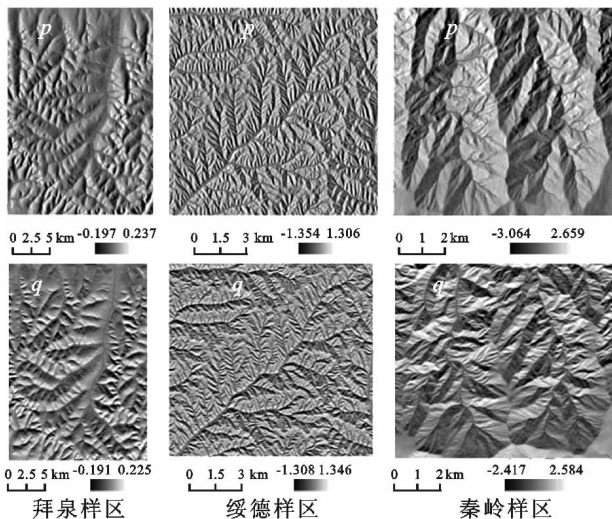


图 3  $p$  和  $q$  图像

表 2 三个样区高程梯度  $p$  和  $q$  变异函数

算法	$p$ 和 $q$	拜泉		绥德		秦岭	
		基台值	变程/m	基台值	变程/m	基台值	变程/m
算法 1	$p$	0.00085	496.3	0.162	111.7	0.494	389.9
	$q$	0.00073	531.7	0.117	111.7	0.342	230.4
算法 2	$p$	0.00084	478.6	0.156	111.7	0.490	389.9
	$q$	0.00087	496.3	0.114	113.4	0.338	248.1
算法 3	$p$	0.00084	478.6	0.155	115.2	0.487	407.7
	$q$	0.00072	496.3	0.111	115.2	0.333	230.4
算法 4	$p$	0.00084	496.3	0.153	113.4	0.484	407.7
	$q$	0.00072	496.3	0.110	115.2	0.332	230.4
算法 5	$p$	0.00084	478.6	0.155	113.4	0.488	407.7
	$q$	0.00072	496.3	0.112	115.2	0.333	230.4
算法 6	$p$	0.00084	478.6	0.155	115.2	0.487	407.7
	$q$	0.00072	496.3	0.111	115.2	0.334	230.4
算法 7	$p$	0.00084	478.6	0.161	111.7	0.492	389.9
	$q$	0.00073	496.3	0.116	111.7	0.340	230.4

在各种算法均能表现局地结构和样区之间的差异的基础上, 各样区、各种算法的局地方差计算统计表明 (表 3), 算法 3 和算法 4 (Evans-Young 算法和改进的 Evans-Young 算法) 局部方差最小, 说明其表面比较光滑, 从高通滤波器的基本特征 (计算局部变异的同时, 清除过高频率的噪音) 出发考虑, 初步认为 Evans-Young 算法和改进的 Evans-Young 算法表现最优。

### 2.2 高程梯度的统计分布特征

拜泉、绥德、秦岭三个样区  $p$  和  $q$  的统计值 (表 4, 表 5, 表 6) 表明: (1) 三个研究区的  $p$  和  $q$  的均值十分接近零。拜泉样区标准差最小, 绥德样区标准差较大, 秦岭样区标准差最大。这种特征的地学原理在于地表经过一个抬升运动后, 高起的地方将被夷平、低洼地方将被填充, 这也就是戴维斯侵蚀循环的基本思想。(2)  $p$  和  $q$  的标准差均近似相等。三个样区无论地势平缓还是陡峻, 都具有这种特征。相同样区  $p$  和  $q$  标准差的不同, 是由于地势的趋势所致。三个样区, 均是东西方向上变异较大, 这与三个样区的流域大多呈南北向 (其中绥德呈北北东—南南西方向) 有关。(3) 三个样区的众数为零, 中位数近似为零, 故三个样区均接近正态分布。拜泉样区  $p$  和  $q$  为正偏态, 绥德样区  $p$  和  $q$  为负偏态, 秦岭样区的  $p$  为负偏态,  $q$  为正偏态。(4)  $p$  和  $q$  相互独立: 从数学意义上讲, 因为  $p$  和  $q$  是坡度  $\theta$  在  $x$  和  $y$  方向上的分量, 故  $p$  和  $q$  是不相关的。相关系数计算结果表明 (表 7), 三个研究区的相关系数均接近于零, 可以认为  $p$  和  $q$  相互独立。

表 3 三个样区高程梯度  $p$  和  $q$  统计值(信息熵和局部方差均值)

算法	$p$ 和 $q$	拜泉		绥德		秦岭	
		信息熵	局部方差均值	信息熵	局部方差均值	信息熵	局部方差均值
算法 1	$p$	7.0888	0.0000230	7.3916	0.03081	7.9233	0.03352
	$q$	6.9658	0.0000198	7.2427	0.02425	7.7526	0.03539
算法 2	$p$	7.0879	0.0000220	7.3781	0.02870	7.9188	0.03180
	$q$	6.9648	0.0000190	7.2281	0.02246	7.7462	0.03351
算法 3	$p$	7.0871	0.0000214	7.3693	0.02743	7.9136	0.03047
	$q$	6.9637	0.0000184	7.2159	0.02118	7.7410	0.03219
算法 4	$p$	7.0866	0.0000212	7.3664	0.02707	7.9125	0.03014
	$q$	6.9633	0.0000183	7.2131	0.02087	7.7393	0.03181
算法 5	$p$	7.0873	0.0000215	7.3714	0.02772	7.9149	0.03078
	$q$	6.9639	0.0000186	7.2188	0.02147	7.7423	0.03249
算法 6	$p$	7.0872	0.0000215	7.3703	0.02757	7.9141	0.03062
	$q$	6.9638	0.0000185	7.2172	0.02132	7.7415	0.03233
算法 7	$p$	7.0891	0.0000226	7.3903	0.02989	7.9228	0.03284
	$q$	6.9662	0.0000195	7.2400	0.02338	7.7526	0.03472

表 4 拜泉样区高程梯度  $p$  和  $q$  统计值

算法	$p$ 和 $q$	最小值	最大值	平均值	标准差	偏度	峰度	众数	中位数
算法 1	$p$	-0.2035	0.2490	0.0011	0.0296	0.3243	4.1934	0	-0.0002
	$q$	-0.1939	0.2292	-0.0012	0.0266	0.3629	4.1805	0	-0.0031
算法 2	$p$	-0.1997	0.2420	0.0011	0.0296	0.3214	4.1617	0	-0.0002
	$q$	-0.1923	0.2271	-0.0013	0.0265	0.3621	4.1617	0	-0.0031
算法 3	$p$	-0.1980	0.2390	0.0011	0.0295	0.3201	4.1460	0	-0.0003
	$q$	-0.1911	0.2260	-0.0013	0.0265	0.3617	4.1489	0	-0.0031
算法 4	$p$	-0.1974	0.2374	0.0011	0.0295	0.3190	4.1361	0	-0.0003
	$q$	-0.1906	0.2248	-0.0013	0.0265	0.3618	4.1454	0	-0.0031
算法 5	$p$	-0.1985	0.2393	0.0011	0.0295	0.3205	4.1493	0	-0.0003
	$q$	-0.1914	0.2262	-0.0013	0.0266	0.3618	4.1521	0	-0.0031
算法 6	$p$	-0.1982	0.2391	0.0011	0.0295	0.3203	4.1476	0	-0.0003
	$q$	-0.1913	0.2261	-0.0012	0.0265	0.3618	4.1504	0	-0.0031
算法 7	$p$	-0.2031	0.2496	0.0011	0.0296	0.3240	4.1896	0	-0.0002
	$q$	-0.1950	0.2334	-0.0013	0.0266	0.3635	4.1841	0	-0.0031

2.3 高程梯度算法稳定性和不确定性

稳定性指算法对于噪音的抗干扰能力,不同算法的抗噪性不同,抗噪性越强,说明此算法稳定性越好。

不确定性指数数据中的误差会叠加算法误差并转递到最终计算结果中,使结果出现误差。误差越小,说明此算法抗噪能力强。

表 5 绥德样区高程梯度  $p$  和  $q$  统计值

算法	$p, q$	最小值	最大值	平均值	标准差	偏度	峰度	众数	中位数
算法 1	$p$	-1.4551	1.3230	0.0074	0.4021	-0.0812	2.3144	0	0.0230
	$q$	-1.3574	1.4408	0.0052	0.3420	-0.1361	2.6142	0	0.0210
算法 2	$p$	-1.3738	1.3207	0.0074	0.3968	-0.0828	2.3119	0	0.0220
	$q$	-1.3273	1.3733	0.0051	0.3372	-0.1431	2.6105	0	0.0210
算法 3	$p$	-1.3616	1.3098	0.0074	0.3933	-0.0855	2.3129	0	0.0230
	$q$	-1.3125	1.3536	0.0051	0.3332	-0.1536	2.6140	0	0.0230
算法 4	$p$	-1.3535	1.3064	0.0073	0.3923	-0.0857	2.3120	0	0.0230
	$q$	-1.3083	1.3458	0.0050	0.3324	-0.1549	2.6125	0	0.0230
算法 5	$p$	-1.3646	1.3125	0.0074	0.3941	-0.0849	2.3126	0	0.0210
	$q$	-1.3143	1.3585	0.0051	0.3342	-0.1510	2.6130	0	0.0230
算法 6	$p$	-1.3631	1.3111	0.0074	0.3937	-0.0852	2.3128	0	0.0230
	$q$	-1.3133	1.3560	0.0051	0.3337	-0.1524	2.6135	0	0.0230
算法 7	$p$	-1.4303	1.3327	0.0073	0.4014	-0.0823	2.3216	0	0.0230
	$q$	-1.3403	1.4118	0.0050	0.3411	-0.1395	2.6199	0	0.0220

表 6 秦岭样区高程梯度  $p$  和  $q$  统计值

算法	$p$ 和 $q$	最小值	最大值	平均值	标准差	偏度	峰度	众数	中位数
算法 1	$p$	-3.1250	2.7177	0.0250	0.7057	-0.1861	2.4857	0	0.0900
	$q$	-2.4706	2.6122	-0.1556	0.5796	0.3725	3.0236	0	-0.2190
算法 2	$p$	-3.0999	2.6910	0.0250	0.7029	-0.1862	2.4766	0	0.0910
	$q$	-2.4375	2.5980	-0.1560	0.5757	0.3743	3.0223	0	-0.2180
算法 3	$p$	-3.0684	2.6618	0.0251	0.7006	-0.1873	2.4653	0	0.0920
	$q$	-2.4230	2.5868	-0.1559	0.5729	0.3777	3.0149	0	-0.2180
算法 4	$p$	-3.0635	2.6588	0.0252	0.7002	-0.1876	2.4607	0	0.0930
	$q$	-2.4168	2.5839	-0.1566	0.5720	0.3795	3.0170	0	-0.2190
算法 5	$p$	-3.0762	2.6691	0.0251	0.7012	-0.1871	2.4679	0	0.0920
	$q$	-2.4267	2.5896	-0.1559	0.5736	0.3769	3.0167	0	-0.2180
算法 6	$p$	-3.0722	2.6654	0.0251	0.7009	-0.1872	2.4665	0	0.0920
	$q$	-2.4248	2.5881	-0.1559	0.5732	0.3774	3.0158	0	-0.2180
算法 7	$p$	-3.1136	2.6975	0.0252	0.7059	-0.1872	2.4822	0	0.0930
	$q$	-2.4686	2.6143	-0.1566	0.5793	0.3761	3.0304	0	-0.2200

表 7 三个样区高程梯度  $p$  和  $q$  的相关系数

算法	$p$ 和 $q$	拜泉	绥德	秦岭
算法 1	$p_{-1}$ 和 $q_{-1}$	-0.0764	0.1002	-0.1655
算法 2	$p_{-2}$ 和 $q_{-2}$	-0.0755	0.1118	-0.1644
算法 3	$p_{-3}$ 和 $q_{-3}$	-0.0756	0.1117	-0.1627
算法 4	$p_{-4}$ 和 $q_{-4}$	-0.0757	0.1121	-0.1635
算法 5	$p_{-5}$ 和 $q_{-5}$	-0.0756	0.1118	-0.1632
算法 6	$p_{-6}$ 和 $q_{-6}$	-0.0756	0.1117	-0.1630
算法 7	$p_{-7}$ 和 $q_{-7}$	-0.0755	0.1105	-0.1650

2.3.1 高程梯度算法稳定性 算法的稳定性主要通过添加噪音前后图像的直方图变化、添加噪音引起的误差来分析。直方图相似度指数(HS)计算结果(表 8)表明,拜泉样区的 HS 在 0.49~0.94,其中算法 4 的 HS 值最大;绥德样区的 HS 在 0.97~0.99,其中  $p$  的算法 4 的 HS 值最大, $q$  的算法 3 的 HS 值最大;秦岭样区的 HS 为 0.96 左右。其中  $p$  的算法 4 的 HS 值最大, $q$  的算法 3 的 HS 值最大。

添加噪音引起的误差计算(表 8)表明,三个样区

的 RMSE 在 0.01~0.04,其中算法 4 的 RMSE 值最小,算法 3 的 RMSE 较小。从总体上看,Evans-Young 特别是经过 Shary 改进后的 Evans-Young 算法抗噪效果最好,表现最为稳定。

2.3.2 高程梯度算法的不确定性 误差传递理论认为,数据中的误差会叠加算法误差并传递到最终计算结果中。数据误差(高程中误差)主要受到数据和插值算法的影响,本文不做重点讨论,利用原始数据(DLG)和拟合 DEM 计算的中误差(表 9) $m_z$  代入公式 12—19 计算  $p, q$  的中误差,用来说明算法的综合误差。计算结果(表 9)表明,拜泉样区在 0.002~0.025,绥德样区在 0.035~0.457,秦岭样区在 0.034~0.439。就地区而言,拜泉样区地势相对平缓,误差比较小,绥德和秦岭两样区地势比较陡,因而误差较大。从各种算法看,改进的 Evans-Young 算法的误差最小,说明此算法最好。

表 8 三个样区高程梯度  $p$  和  $q$  均方根误差与直方图相似度指数

算法	$p$ 和 $q$	拜泉		绥德		秦岭	
		均方根误差	直方图相似度指数	均方根误差	直方图相似度指数	均方根误差	直方图相似度指数
算法 1	$p$	0.04077	0.5746	0.04087	0.9724	0.04087	0.9634
	$q$	0.04086	0.4969	0.04080	0.9741	0.04082	0.9654
算法 2	$p$	0.02041	0.8315	0.02038	0.9781	0.02041	0.9624
	$q$	0.02043	0.7645	0.02041	0.9728	0.02039	0.9659
算法 3	$p$	0.01178	0.9228	0.01177	0.9784	0.01180	0.9637
	$q$	0.01182	0.8820	0.01178	0.9999	0.01176	0.9679
算法 4	$p$	0.01108	0.9438	0.01108	0.9805	0.01111	0.9652
	$q$	0.01113	0.9057	0.01109	0.9803	0.01107	0.9647
算法 5	$p$	0.01249	0.9246	0.01248	0.9791	0.01251	0.9641
	$q$	0.01253	0.8629	0.01248	0.9773	0.01247	0.9673
算法 6	$p$	0.01195	0.9153	0.01194	0.9796	0.01197	0.9643
	$q$	0.01199	0.8743	0.01195	0.9801	0.01193	0.9662
算法 7	$p$	0.01318	0.9140	0.01318	0.9786	0.01321	0.9581
	$q$	0.01325	0.8778	0.01321	0.9777	0.01319	0.9663

因为三阶差分算法(算法 3)和改进的三阶差分算法(算法 4)这两种算法是三阶差分形式,使用了  $3\times 3$  局部窗口中的多数像元,对 DEM 数据都有过滤和平滑作用,避免各种极端数据的出现,因此能够提高 DEM 地形表示精度,故其在计算  $p$  和  $q$  时抗噪性最好。简单差分算法(算法 1)、二阶差分算法(算法 2)只使用了局部窗口的少部分像元,这两种方法对原始数据误差(噪音)比较敏感,故而精度较低,抗噪性不佳。

表 9 三个样区高程梯度  $p$  和  $q$  中误差

中误差		拜泉	绥德	秦岭
高程中误差( $m_z$ )		0.359362	4.781516	4.686012
	算法 1	0.002583	0.457258	0.439174
高程梯度中误差 ( $m_p$ 或 $m_q$ )	算法 2	0.000646	0.114314	0.109794
	算法 3	0.000215	0.038105	0.036598
	算法 4	0.000201	0.035670	0.034260
	算法 5	0.000242	0.042868	0.041173
	算法 6	0.000348	0.061673	0.059234
	算法 7	0.000288	0.050923	0.048909

3 结论与讨论

3.1 讨论

在对  $p$  和  $q$  算法不确定性方面的认识,本研究的认识与 Zhou 等<sup>[7]</sup> 和刘学军等<sup>[6]</sup> 的认识基本相同,即 Evans-Young 算法精度最高。但是与一些学者的认识有所不同,如 Florinsky<sup>[11]</sup> 和 Skidmore<sup>[4]</sup> 均认为 Sobel 算法较优,而我们的分析中,Sobel 算法的表现排在 Evan-Young 算法之后。本研究的一点进步是从滤波器的基本性状去分析每种算法的优劣,并解释 Evan-Young 算法( $p$  和  $q$  与坡度算法)较优的原因。

本研究中,数据的计算可初步认识  $p$  和  $q$  的性质(均值为零、正态分布、方差相等、相互独立),这是本研究的特色所在。虽然我们已尝试从戴维斯侵蚀循环的基本学说去解释  $p$  和  $q$  均值为零的原因,但是从总体上看  $p$  和  $q$  的统计分布特征还有待从数学、物理学(重力势能、零势能面)方面的论证和计算分析。理论上讲  $p$  和  $q$  应该相互独立,但是数据计算结果却不完全相互独立,不同的算法其结果也稍有不同; $p$  和  $q$  的频率分布接近正态分布,但是还不是完全的正态分布,不同的算法之间也有差异。这些都有待于进一步的分析论证。

3.2 结论

(1)  $p$  和  $q$  的基本性质是,均值近似为零,相互独立,方差近似相等,十分接近正态分布。这一性质对坡度的统计分布十分重要,因为坡度算法(公式 2)中的主体部分将符合卡方分布,这一数学特征使坡度理论分布模型的推导成为可能。

(2) 各种算法均能表现局地结构和样区之间的差异,地形越复杂, $p$  和  $q$  表面起伏越大,信息含量也越大。拜泉样区最小,秦岭样区较大,绥德样区最大,与实际地形相符。

(3) 从算法稳定性、不确定性(误差结构)、信息含量等方面看,Evans-Young 算法,特别是经过 Shary 改进的 Evans-Young 算法的效果最佳。

参考文献:

[1] Olaya V. Basic land-surface parameters[J]. Developments in Soil Science, 2009,33(8):141-169.

[2] Gonzalez R C, Woods R E. Digital Image Processing [M]. Beijing:Publishing House of Electronics Industry, 2010.

[3] Pike R J, Evans I S, Hengl T. Geomorphometry: a brief guide[J]. Developments in Soil Science, 2009,33(8):3-30.

[4] Skidmore A K. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model[J]. International Journal of Geographical Information System, 1989,3(4):323-334.

[5] Skidmore A K. Evolution of methods for estimating slope gradient and aspect from digital elevation models [C]. Classics from IJGIS:Twenty Years of the International Journal of Geographical Information Science and Systems, 2006:111-116.

[6] 刘学军,龚健雅,周启鸣,等. 基于 DEM 坡度坡向算法精度的分析研究[J]. 测绘学报,2004,33(3):258-263.

[7] Zhou Q, Liu X. Error analysis on grid-based slope and aspect algorithms[J]. Photogrammetric Engineering & Remote Sensing, 2004,70(8):957-962.

[8] Warren S D, Hohmann M G, Auerswald K, et al. An evaluation of methods to determine slope using digital elevation data[J]. Catena, 2004,58(3):215-233.

[9] 汤国安,赵牡丹,李天文,等. DEM 提取黄土高原地面坡度的不确定性[J]. 地理学报,2002,58(6):824-830.

[10] 刘学军,卞璐,卢华兴,等. 顾及 DEM 误差自相关的坡度计算模型精度分析[J]. 测绘学报,2008,37(2):200-206.

[11] Florinsky I V. Accuracy of local topographic variables derived from digital elevation models[J]. International Journal of Geographical Information Science, 1998,12(1):47-62.

[12] Florinsky I V. Computation of the third-order partial derivatives from a digital elevation model[J]. International Journal of Geographical Information Science, 2009,23(2):213-231.

[13] Hutchinson M F. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits[J]. Journal of Hydrology, 1989,106(3/4):211-232.

[14] 杨勤科,李领涛,张彩霞. ANUDEM:专业化数字高程

- 模型插值算法及其特点[J]. 干旱地区农业研究, 2006, 24(3):36-41.
- [15] 杨勤科, 师维娟. 水文地貌关系正确 DEM 的建立方法[J]. 中国水土保持科学, 2007, 5(4):1-6.
- [16] Reuter H I, Hengl T, Gessler P, et al. Chapter 4 Preparation of DEMs for Geomorphometric Analysis[J]. Developments in Soil Science, 2009, 33(8):87-120.
- [17] Evans I S. An integrated system of terrain analysis and slope mapping. final report (report 6) on grant da-ero-591-73-g0040[J]. Physica B: condensed Matter, 1979, 210(1):59-66.
- [18] Zevenbergen L W, Thorne C R. Quantitative analysis of land surface topography[J]. Earth Surface Processes and Landforms, 1987, 12(1):47-56.
- [19] Horn B K P. Hill shading and the reflectance map[J]. Proceedings of the IEEE, 1981, 69(1):14-47.
- [20] Jones K H. A comparison of algorithms used to compute hill slope as a property of the DEM[J]. Computers and Geosciences, 1998, 24(4):315-323.
- [21] Young M, Evans I S. Statistical Characterization of Altitude Matrices by Computer. Report5. Terrain Analysis; Program Documentation[R]. Durham Univ. (United Kingdom) Dept. of Geography, 1978.
- [22] Pennock D J, Zebarth B J, De Jong E. Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada[J]. Geoderma, 1987, 40(3/4):297-315.
- [23] Sharpnack D A, Akin G. An algorithm for computing slope and aspect from elevations[J]. Photogrammetric Engineering, 1969, 35(3):247-248.
- [24] Shary P A, Sharaya L S, Mitusov A V. Fundamental quantitative methods of land surface analysis[J]. Geoderma, 2002, 107(1):1-32.
- [25] Sobel I, Feldman G. A  $3 \times 3$  isotropic gradient operator for image processing[R]. Stanford Artificial Project, 1968:271-272.
- [26] Unwin D J. Introductory Spatial Analysis[M]. Taylor & Francis, 1981.
- [27] Shannon C E. A mathematical theory of communication[J]. Acm Sigmoble Mobile Computing And Communications Review, 2001, 5(1):3-55.
- [28] Woodcock C E, Strahler A H. The factor of scale in remote sensing[J]. Remote sensing of Environment, 1987, 21(3):311-332.
- [29] Webster R, Oliver M A. Geostatistics for environmental scientists[M]. USA: John Wiley & Sons, 2007.
- [30] 牛亮, 杨勤科. DEM 尺度变换中直方图相似度计算与应用[J]. 水土保持研究, 2010, 17(3):120-125.

~~~~~  
(上接第 18 页)

- [9] 卜耀军, 温仲明, 焦峰, 等. 黄土丘陵区人工与自然植物群落物种多样性研究: 以安塞县为例[J]. 水土保持研究, 2005, 12(1):4-6.
- [10] 何艺玲, 傅懋毅. 人工林林下植被的研究现状[J]. 林业科学研究, 2002, 15(6):727-733.
- [11] 程积民, 万惠娥, 胡相明. 黄土丘陵区植被恢复重建模式与演替过程研究[J]. 草地学报, 2005, 13(4):324-328.
- [12] 王国梁, 刘国彬, 刘芳, 等. 黄土沟壑区植被发育过程中植物群落组成及结构变化[J]. 生态学报, 2003, 23(12):2551-2557.
- [13] 师阳阳, 陈云明, 张光辉, 等. 不同退耕年限撂荒地植物多样性及生物量分析[J]. 水土保持研究, 2012, 19(6):36-44.
- [14] 吴旭, 陈云明, 唐亚坤. 黄土丘陵区刺槐和侧柏人工林树干液流特征及其对降水的响应[J]. 植物生态学报, 2015, 39(12):1176-1187.
- [15] 王军, 付博杰, 邱扬, 等. 黄土高原小流域土壤养分的空间异质性[J]. 生态学报, 2002, 22(8):1173-1178.
- [16] 宋永昌. 植被生态学[M]. 上海: 华东师范大学出版社, 2001.
- [17] 温远光, 赖家业, 梁宏温, 等. 大明山退化生态系统群落的外貌特征研究[J]. 广西农业大学学报, 1998, 17(2):154-159.
- [18] Raunkiaer C. The life forms of plants and statistical plant geography[M]. New York: Oxford University Press, 1932.
- [19] 中国植被编辑委员会. 中国植被[M]. 北京: 科学出版社, 1980.
- [20] 雷明德. 陕西植被[M]. 北京: 科学出版社, 1999.
- [21] 贺山峰, 蒋德明, 李晓兰, 等. 小叶锦鸡儿固沙群落草本种群重要值与生态位的研究[J]. 干旱区资源与环境, 2007, 21(10):150-155.
- [22] 谢春平, 王华辰, 陈林, 等. 木林子自然保护区河谷樱花群落特征分析[J]. 南京林业大学学报, 2014(S):81-88.
- [23] 曲红, 王百田, 王棣, 等. 黄土区不同配置人工林物质多样性研究[J]. 生态环境学报, 2010, 19(4):843-848.
- [24] 焦菊英, 张振国, 贾燕峰, 等. 陕北丘陵沟壑区撂荒地自然恢复植被的组成结构与数量分类[J]. 生态学报, 2008, 28(7):2981-2997.
- [25] 王军, 付博杰, 邱扬, 等. 黄土高原小流域土壤养分的空间异质性[J]. 生态学报, 2002, 22(8):1173-1178.
- [26] 李裕元, 邵明安. 子午岭植被自然恢复过程中植物多样性的变化[J]. 生态学报, 2004, 24(2):252-260.
- [27] 邵彬, 邓坤枚. 长白山北坡亚高山云冷杉林的植物种类组成及重要值[J]. 自然资源学报, 2000, 15(1):66-73.
- [28] 裴娟, 艾应伟, 刘浩, 等. 坡面和坡向对遂渝铁路岩石边坡创面人工土壤植被恢复的影响[J]. 水土保持通报, 2009, 29(2):197-201.