

格网 DEM 上径流长度计算误差的定量分析

刘学军^{1,2} 晋 蓓^{1,2} 胡加佩^{1,2} 马锦绢^{1,2}

(1 南京师范大学虚拟地理环境教育部重点实验室,南京市栖霞区文苑路 1 号,210046)
(2 南京师范大学地理科学学院,南京市栖霞区文苑路 1 号,210046)

摘 要:基于格网 DEM 的径流长度的计算误差主要来源于 DEM 数据和流径提取算法。通过数据独立的误差分析方法,定量分析流径提取算法模型产生的误差。首先选取 5 种数学曲面作为模拟 DEM 数据源,计算出径流长度,以由曲面公式推导的径流长度公式计算的径流长度作为理论真值,通过两者的对比,分析误差及其空间分布。结果表明,汇水区的径流长度计算精度明显高于分水区;分水区的径流长度含有一定的系统误差。

关键词:格网 DEM; 径流长度; 误差; 定量分析

中图法分类号:P208

径流长度(flow length)是描述地形形态和流域特征信息的重要参数,在水文模型、水土保持和土壤侵蚀等领域有广泛的应用^[1-2]。当其他外在条件相同时,物质沉积量、水力侵蚀和冲刷的强度由坡面的长度来决定,径流长度愈长,径流速度就愈大,汇集的径流也愈大,侵蚀力也愈强^[3]。

传统径流长度获取方式是野外实地测量或基于地形图、遥感影像进行的。前者在实际应用中有时很困难或者不现实;后者取决于地形图、遥感影像的质量^[4],同时费时费力,精度难以保证。随着地理信息系统(GIS)、数字高程模型(DEM)以及数字地形分析等的发展,直接在 DEM 上获取径流长度已成为现实。目前,DEM 径流长度计算模型从原理上有基于非流量累计的直接计算方式^[5-6]和基于累计流量的单位汇面积法^[7-8];从计算方式角度,有直接径流长度计算^[5-6]和间接径流长度计算^[9-10]。直接径流长度的计算是实际应用的前提,其结果不仅可用来进行土壤侵蚀研究,还能作为水文分析及其他研究的重要参数,参与相关地学过程的模拟和模型构建^[11]。

在直接径流长度计算中,目前以 Hickey^[6]所提出的算法应用较为广泛。与传统实地测量、手工方式相比,尽管大幅度提高了计算效率,但由于

地表本身的非解析性、地表物质运动的复杂性以及 DEM 的规则格网结构特征,径流长度算法在设计上都不同程度地存在假设,导致径流长度计算与实际不相符并进而影响地学分析和模拟的准确性^[12],而通过对实际 DEM 的计算和比较所得出的结论并不能排除 DEM 数据对地形表面的模拟误差。本文采用数据独立分析方法^[13],设计了 5 个模拟曲面并推导了各个模拟曲面上的径流长度计算公式。在模拟曲面 DEM 上,通过理论值和计算值的对比,详细讨论了 DEM 径流长度的误差性质和空间分布特征。

1 径流长度计算模型

1.1 径流长度数学模型

径流长度(FL),即水流路径长度,定义为从给定点逆流而上到水流起点的地面距离在水平面上的投影长度^[3]。显然,径流长度计算依赖于流向(flow direction,FD)和流线(flow line)。流向是水流坡度的最大方向,即坡向或其近似值;流线是沿流向从水流起点到给定点 P 的水流路径(flow path,FP)。流向和流线是由地表自然属性控制的。如图 1 所示,设区域 Ω 的地形曲面为

$z=F(x,y)$,其在平面上表示一等高线族 $F(x,y)=c$ (c 为任意高程)。由数学分析知,对于平面上高程为 H_P 的任意点 $P(x,y)$,其梯度为:

$$\text{Grad}f(x,y)=\frac{\partial F}{\partial x}\mathbf{i}+\frac{\partial F}{\partial y}\mathbf{j}\tag{1}$$

式中, \mathbf{i},\mathbf{j} 为单位方向。沿着 P 的梯度反方向, $F(x,y)$ 取得其下降最快值。该值即为 P 的坡度 S ,其下降方向即为流向:

$$\begin{cases} S=\sqrt{\left(\frac{\partial F}{\partial x}\right)^2+\left(\frac{\partial F}{\partial y}\right)^2} \\ \text{FD}=\arctan\left(\frac{\partial F}{\partial x}/\frac{\partial F}{\partial y}\right),\frac{\partial F}{\partial y}\neq 0 \end{cases}\tag{2}$$

设过 P 点的等高线为 $y=f(x)$, P 点的流向与经过该点的等高线正交。即当任一点沿曲面由高向低运动并经过 P 点时,其轨迹在平面的投影即为流线。流线上每一点的法线方向为该点的等高线切线方向,即流线处处与等高线正交。设过点 P 的流线方程为 $g(x)$,则有:

$$f'(x)g'(x)=-1\tag{3}$$

由该微分方程可求得过 P 的流线方程 $g(x)$ 为:

$$g(x)=\int\frac{-1}{f'(x)}\text{d}x\tag{4}$$

如果 P 的水流起点为 $Q(x_Q,y_Q)$,则 P 点的径流长度 FL_P 为沿 $g(x)$ 的曲线积分:

$$\text{FL}_P=\left|\int_{x_Q}^{x_P}\sqrt{1+(g'(x))^2}\text{d}x\right|=\tag{5}$$

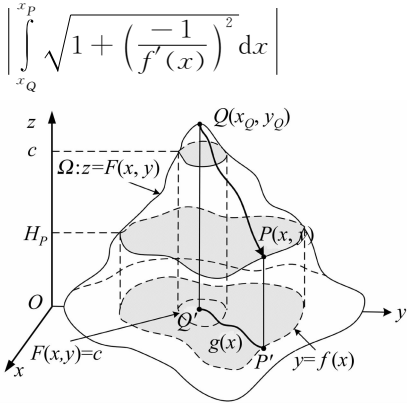


图 1 径流长度示意图
Fig. 1 Definition of Flow Length

综上所述,任意给定点的上游水流路径长度必然存在如下性质:① 水流路径与等高线垂直;② 水流路径必然为给定点和流经上其他点之间距离的最短者。

1.2 DEM 径流长度计算模型

目前通用的计算模型^[5]首先对 DEM 进行预处理,形成水流路径畅通的无洼地 DEM,在此基

础上,计算每个 DEM 栅格单元的径流长度^[3]。

2 实验设计

2.1 数学曲面设计

考虑到计算的简洁性并顾及曲面的地形意义,选择了均匀坡面、汇水区、分水区和混合区域等 5 个数学曲面。

1) 斜平面(plane):平面上坡度与坡向为常数,各点径流长度的起点为过该点的流线与所给平面边界的交点。

2) 下半球面(inversed sphere):模拟汇水区域,汇水点在球面中心,各象限的曲面为汇水山坡,坡面上任意点的径流长度起算点在球的边界。

3) 上半球面(sphere):描述分水区域,汇水区域在边界处,中心点的汇水面积为零,各象限的曲面为分水山坡,坡面上各点径流长度起点为球面中心。

4) 马鞍面(saddle surface):模拟混合区域。东西方向为分水线,南北方向为合水线。原点附近为地形鞍部,象限曲面为山坡,各点径流长度计算起点为过该点的流线与水平方向(x 轴)的交点,且位于水平方向(山脊线)上的各点径流长度为零。

2.2 数学曲面上径流长度真值计算

各模拟曲面定义及任意点径流长度计算公式如表 1。这里以马鞍面的计算为例进行推导:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=\frac{z}{c}\text{ 或 }z=f(x,y)=c\left(\frac{x^2}{a^2}-\frac{y^2}{b^2}\right)\tag{6}$$

其区域范围 $\Omega=[-A\leq x,y\leq A]$ 。设 $P(x_P,y_P,H_P)$ 为椭球面上任一点,则过 P 的等高线为:

$$H_P=c\left(\frac{x^2}{a^2}-\frac{y^2}{b^2}\right)\tag{7}$$

式(7)两边对 x 求导,当 $x\neq 0,y\neq 0$ 时,有:

$$0=c\left(\frac{2x}{a^2}-\frac{2yy'}{b^2}\right)\tag{8}$$

由式(8)有:

$$y'=\frac{b^2x}{a^2y}\tag{9}$$

式中,若 $y=0$ 或者 $x=0$,则相应切线与坐标轴平行或垂直。

由式(3)有:

$$y'=g'(x)=\frac{-1}{f'(x)}=-\frac{a^2y}{b^2x}\tag{10}$$

积分式(10),有过 P 的流线方程为:

$$y=kx^{-(a^2/b^2)}\tag{11}$$

式中, k 是积分常数, 由于 $g(x)$ 经过 P 点, 则有 $k = y_P x_P^{(a^2/b^2)}$ 。

在马鞍面中, 过 P 的等高线 $f(x)$ 和流经 $g(x)$ 形成一正交曲线网络, 其汇聚点在水平方向两侧的边界上, 即 $(-A, 0)$ 和 $(A, 0)$ 处。此两点即为 P 的水流起点 Q 。结合式(11)和式(5), 即

可求得马鞍面上任意点处的径流长度。

为公式推导和讨论方便, 平面的位置设置在第一象限, 左下角为坐标原点; 其他曲面的坐标原点均设置在其几何中心, Y 轴为南北方向(北正南负), X 轴为东西方向(东正西负), Z 轴为高程。依据上述的推导方法得出 5 种曲面的径流长度计算式如表 1 所示。

表 1 数学曲面及其径流长度表达式/m
Tab. 1 Formulas of Mathematical Surfaces and Their Flow Length

曲面类型	曲面示意图	曲面定义及径流长度公式
平面		$z = ax + by + c$ 参数: $a = 2, b = 1, c = 100$; 范围: $0 \leq x, y \leq 500$ $FL_P = \left \int_{x_Q}^{x_P} \sqrt{1 + (g'(x))^2} dx \right $ 式中, $g'(x) = \frac{b}{a}, x_Q = A$ 或 $x_Q = \frac{a}{b} [A - (y_P - \frac{b}{a} x_P)]$ $x^2 + y^2 + z^2 = a^2, z \leq 0$ 参数: $a = 400$; 范围: $-400 \leq x, y \leq 400$ $FL_P = \left \int_{x_Q}^{x_P} \sqrt{1 + (g'(x))^2} dx \right $
下半球面		式中, $g'(x) = \frac{y_P}{x_P}, x_Q = \begin{cases} \sqrt{\frac{R^2 x_P^2}{y_P^2 + x_P^2}}, P \in I, VI \\ -\sqrt{\frac{R^2 x_P^2}{y_P^2 + x_P^2}}, P \in II, III \end{cases}$ $x^2 + y^2 + z^2 = a^2, z \geq 0$ 参数: $a = 400$; 范围: $-400 \leq x, y \leq 400$ $FL_P = \left \int_{x_Q}^{x_P} \sqrt{1 + (g'(x))^2} dx \right $
上半球面		式中, $g'(x) = \frac{y_P}{x_P}$ $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, z \geq 0$ 参数: $a = 400, b = 300, c = 300$; 范围: $-400 \leq x \leq 400, -300 \leq y \leq 300$ $FL_P = \left \int_{x_Q}^{x_P} \sqrt{1 + (g'(x))^2} dx \right $
上半椭球面		式中, $g'(x) = \frac{a^2}{b^2} \times y_P \times x_P \times \frac{-a^2}{b^2} \times x \frac{a^2}{b^2} - 1$ $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c}, (a \neq b)$ 参数: $a = 2, b = 1, c = 0.001$; 范围: $-400 \leq x, y \leq 400$ $FL_P = \left \int_{x_Q}^{x_P} \sqrt{1 + (g'(x))^2} dx \right $
马鞍面		式中, $g'(x) = -\frac{a^2}{b^2} \times y_P \times x_P \times \frac{a^2}{b^2} \times x^{-(\frac{a^2}{b^2} + 1)}$ $x_Q = \begin{cases} A, P \in I, VI \\ -A, P \in II, III \end{cases}$

3 结果与分析

对各数学曲面按照确定的分辨率离散(平面和马鞍面为 10 m, 上下球面、椭球面为 2 m), 建

立数学曲面 DEM。根据表 1 中的公式计算每个格网点上的径流长度理论值, 同时在 ArcGIS 环境中计算径流长度值(观测值), 并计算误差(即理论值-观测值)。表 2 是径流长度误差的统计结果, 其中统计指标为中误差(root mean square er-

ror, RMSE), 标准差 (standard error, SE) 以及最大、最小和平均误差。图 2 是各个曲面的误差频率分布图。图 3 为径流长度误差的空间分布图。

表 2 不同曲面的径流长度计算误差统计数据/m

Tab. 2 Statistics of Error in Extracting Flow Length from Different Surfaces

曲面类型	最小值 min	最大值 max	平均值 (ME)	标准差 (SD)	中误差 (RMSE)
平面	-148.09	205.47	-2.78	70.64	70.69
下半球面	-32.62	400	-0.36	8.33	8.34
上半球面	-25.63	398.90	132.96	94.19	162.94
上半椭圆面	-28.42	396.02	124.58	87.9	152.47
马鞍面	-388.28	394.98	97.29	132.66	164.51

表 2 和图 2、图 3 揭示了径流长度计算误差的性质、大小和空间分布。就误差性质而言,由表 2 和图 2 可知,在斜坡、汇水区域(平面、下半球面),标准差 SD 和中误差 RMSE 比较接近,误差平均值接近于零,径流长度没有明显的系统性误差;而在分水区域(上半球面、上半椭圆面),两者相差较大,表明分水区域径流长度误差含有明显的系统成分。从误差大小看,汇水区域的误差较小,误差平均值接近于零(-0.36),小误差频率比较高,表明误差比较集中。随着曲面从汇水区域、斜坡到分水区域的逐步过渡,误差平均值虽在减少,但中误差随之增大,误差分布呈离散趋势。

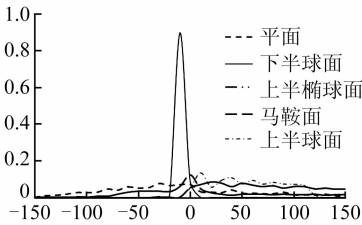


图 2 误差频率图

Fig. 2 Frequency Map of Error

- 径流长度误差的空间分布有如下特征。
- 1) 就单一区域而言,分水区域的误差大于合水区域,坡面计算误差介于两者之间;
 - 2) 混合区域上(图 3(e)),分水区域误差大于合水区域,这主要是马鞍面上,先分水再合水,分水区域径流长度误差逐步累积到合水区域的缘故;
 - 3) 误差极值一般出现在边界上,如区域边缘或者出水口;
 - 4) 误差分布呈八方向特征,同时汇水区域和分水区域刚好相反。
- 之所以如此,根源在于径流长度计算模型中的流向计算方法和径流长度累计思想,分析如下:

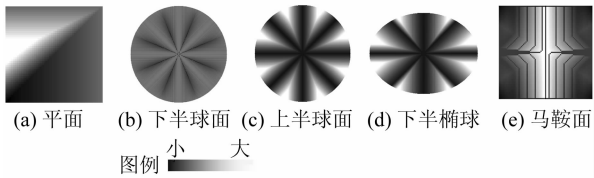


图 3 径流长度误差空间分布

Fig. 3 Spatial Distribution of Error

1) 流向计算。目前算法中流向计算采用的是 D8 算法,其核心是通过计算中心格网与其周围 8 个格网点的最大落差来确定当前格网单元的流向,即流向是在间隔为 45° 的 8 个格网方向进行选择,流向不具备连续性(连续流向是在 $0^\circ \sim 360^\circ$ 之间的任意值)。流向的不连续导致当前栅格单元失去上游来水,原本连续水流路径也变得不连续,从而导致径流长度误差。如图 4 所示,图 4(a)为一局部 DEM,图 4(b)、4(d)为按 § 2.2 方法所获取的流向矩阵和径流长度矩阵,图 4(c)为相应的径流长度的真值。该局部 DEM 为马鞍面中心的左上部分,水流方向以北和东北方向为主,格网单元处于坡面位置。每一格网单元都应有其流径,然而 e_6 、 f_5 等坡面栅格单元本应有上游来水,由于上游单元流向的改变失去上游来水而导致径流长度为零,这同时也影响到这些单元的下游单元径流长度值。

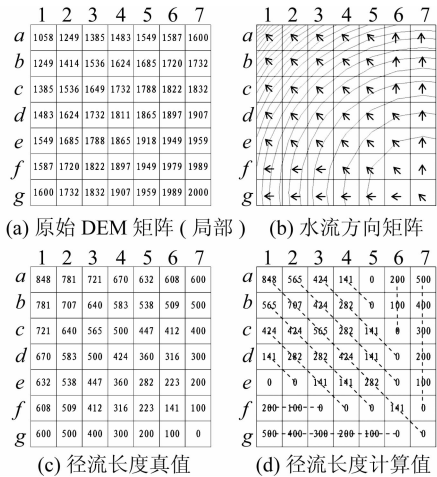


图 4 流向对径流长度计算影响分析
(上半球面,DEM 分辨率:100 m)

Fig. 4 Effect of Flow Direction on Flow Length

2) 径流长度累计。D8 算法的流量分配特征是“多来水单去水”,即可接受多个上游栅格单元的来水但只能从一个方向流出,这对汇水区域的计算较为有利,与汇水区域的汇流特征相符,如在下半球面的径流长度计算结果(表 2)。然而分水区域的水流具有发散特征,“多来水单去水”的流

量分配模式使得当前栅格单元上游水流路径产生多义性,当前栅格单元具有多条汇流路径,那么选择哪一个为当前栅格单元的水流路径? 非流量累积算法(参看 § 2.2)认为,应取上游单元中具有最大径流长度的单元与当前单元内部径流长度累加作为当前栅格单元径流长度。如图 5 中的 g_7 单元,其上游有 g_6 和 f_6 两个来水单元,其中 g_6 具有最大径流长度,而 g_6 又有两个上游来水单元,以 f_5 为最长,其余类推,则形成如图 5(d)中粗实线所示的水流路径,导致计算单元径流长度过大。这不仅违反了流径应有的性质,即两点之间的径流长度为两点之间的最短路径,同时也导致沿弯曲路径进行径流长度计算,引起较大的计算误差。



图 5 径流长度累计对计算影响分析
(下半球面,DEM 分辨率:100 m)

Fig. 5 Effect of Flow Accumulation on Flow Length

4 结 语

- 1) 径流长度误差主要分布在分水区域,合水区域误差较小;分水区域的径流长度误差有一定的系统性;
- 2) 导致上述问题的基本原因是目前径流长度计算模型在原理上存在较大缺陷,特别是流向计算和径流长度累计方面;
- 3) 提高径流长度计算的准确性,需要从改善流向计算方面入手,并确定合理的水流路径。这将是本文后续的主要研究方向。研究结论对改善已有算法和设计新算法,分析地形分析、水文模拟结果的可信度提供了理论支持。

参 考 文 献

[1] 唐从国,刘丛强. 基于 GIS 的乌江流域地表径流模

拟研究[J]. 武汉大学学报·信息科学版,2007,32(5):410-413

[2] 刘耀林,罗志军. 基于 GIS 的小流域水土流失遥感定量监测研究[J]. 武汉大学学报·信息科学版,2006,31(1):35-38

[3] 周启鸣,刘学军. 数字地形分析[M]. 北京:科学出版社,2006

[4] Paz A R, Collischonn W, Risso A, et al. Errors in River Lengths Derived from Raster Digital Elevation Models[J]. Computers and Geosciences, 2008, 34(11): 1 584-1 596

[5] Hickey R, Smith P J. Slope Length Calculations from a DEM with Arc/Info Grid[J]. Computers, Environment and Urban Systems, 1994, 18(5): 365-380

[6] Hickey R. Slope Angle and Slope Length Solutions for GIS[J]. Cartography, 2000, 29(1): 1-8

[7] Desmet P J J, Govers G. Comparison of Routing Algorithms for Digital Elevation Models and Their Implication for Predicting Ephemeral Gullies[J]. International Journal of Geographical Information Science, 1996, 10(3): 311-331

[8] 李俊. 基于 DEM 的黄土高原坡长的自动提取和分析[D]. 西安:西北大学,2007

[9] Moore I D, Wilson J P. Length-slope Factor for the Revised Universal Soil Loss Equation; Simplified Method of Estimation[J]. Journal of Soil and Water Conservation, 1992, 49: 174-180

[10] Mitasova H, Hofierka J, Zlocha M, et al. Modeling Topographic Potential for Erosion and Deposition Using GIS[J]. International Journal of Geographical Information Science, 1996, 10(5): 629-641

[11] 曹龙熹,符素华. 基于 DEM 的坡长计算方法比较分析[J]. 水土保持通报,2007,27(5): 58-62

[12] 晋蓓,刘学军,甄艳,等. ArcGIS 环境下 DEM 的坡长计算与误差分析[J]. 地球信息科学学报,2010,12(5):700-706

[13] Zhou Q. A Data-independent Method for Quantative Accuracy Assessment of Morphological Parameters Extracted from Grid-based DTM[J]. International Archives of Photogrammetry and Remote Sensing, 2000, 33(Part B4): 1 235-1 242

第一作者简介:刘学军,博士,教授,博士生导师。主要研究方向为 DEM 及其地形分析、GIS 空间分析、空间数据不确定性等。
E-mail:liuxuejun@njnu.edu.cn

(下转第 779 页)