

文章编号:1671-8860(2008)12-1259-04

文献标志码:A

# DEM 误差的空间自相关特征分析

王耀革<sup>1</sup> 王志伟<sup>1</sup> 朱长青<sup>1,2</sup>

(1 信息工程大学理学院, 郑州市嵩山南路 73 号, 450001)

(2 南京师范大学虚拟地理环境教育部重点实验室, 南京市文苑路 1 号, 210046)

**摘要:**采用空间自相关分析方法,从空间角度对数字高程数据误差的空间分布特征进行了研究。实验表明,利用双线性曲面表示地形表面时,产生的数字高程数据误差的全局 Moran's I 指数趋近于 0,在整个区域单元上的分布不存在显著的全局空间自相关,邻近区域单元上高程数据误差之间的关系在整体上既不综合表现为趋同,也不综合表现为趋异,高程数据误差的整体空间格局为随机格局;而且数字高程数据误差在空间上的分布与地形坡度和地表粗糙度有一定的联系,一般情况下,平均坡度、地表粗糙度越大,高程数据的全局 Moran's I 指数偏离 0 稍远一些;否则,距离 0 近一些,但全局空间自相关仍不显著,在整体上表现为随机格局。

**关键词:**数字高程数据误差;空间自相关;空间格局

**中图法分类号:**P208; P231.5

数字高程数据误差作为数字高程模型(DEM)生产、质量控制、精度评定和分析应用的重要组成部分,一直是国内外学术界广泛关注的热点问题之一。以往的高程数据误差模型都是通过统计模型对指标进行加权评判,缺乏空间视角,难以真正反映误差空间差异的变化与机制,DEM 误差的空间分布是否存在着特定的规律,是否随着地形部位的不同而变化,这是学术界普遍关心的问题<sup>[1-6]</sup>。空间自相关分析是研究邻近位置的属性(或现象)相关性的空间统计学方法,本文采用该方法对数字高程数据误差进行了研究。

## 1 数字高程数据误差的空间自相关分析

### 1.1 空间关系矩阵

建立空间关系矩阵,以明确各误差在空间位置上的关系。建立空间关系矩阵的关键在于空间邻接的定义。针对规则格网 DEM 模型误差的空间邻接的定义可以是基于格网邻接和基于距离邻接两种。本文建立的空间关系矩阵中权值的确定采用八方向邻近原则,它是基于距离定义的,若用  $d_{ij}$  表示格网点  $i$  与  $j$  之间的距离;  $d$  为设定的空

间距离,这里取值为格网分辨率,若  $0 < d_{ij} \leq d$ , 则定义为相邻, 权重用 1 表示; 若  $d < d_{ij} \leq \sqrt{2}d$ , 则定义为次相邻, 权重用  $1/\sqrt{2}$  表示; 其余格网点的权重用 0 表示,因此,所得的空间关系矩阵  $\omega_{ij}$  是一个由 0, 1 和  $1/\sqrt{2}$  组成的三维矩阵。

### 1.2 全局空间自相关分析

全局空间自相关分析是整个区域上各格网点高程数据对应误差与相邻格网点高程数据误差的空间自相关性质与水平的综合反映,是对数字高程数据误差在整个区域的空间特征的描述,主要通过对全局 Moran's I、全局 Geary's C 和 Join Count 等<sup>[7]</sup>全局空间自相关统计量的估计,分析区域空间内误差的空间关联和空间差异程度。其中,最常用的全局 Moran 指数  $I$  为:

$$I = \frac{\sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n w_{ij} (\epsilon_i - \bar{\epsilon})(\epsilon_j - \bar{\epsilon}) / S_1^2}{\sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n w_{ij}} \quad (1)$$

式中,  $n$  表示整个区域上除边界上的格网点外所有格网点的个数;  $i, j (i, j \in n)$  代表不同的格网点;  $\epsilon_i (i=1, \dots, n)$  为上述  $n$  个格网点高程数据对应的误差值;  $\bar{\epsilon}$  为  $n$  个格网点高程数据的误差值的平均值,  $\bar{\epsilon} = \frac{1}{n} \sum_{i=1}^n \epsilon_i$ ;  $S_1^2 = \frac{1}{n} \sum_{i=1}^n (\epsilon_i - \bar{\epsilon})^2$ ;  $w_{ij}$  为空

间关系矩阵,反映格网点*i*与*j*的空间关系。

Moran指数能从理论上测度数字高程数据误差的空间自相关性,发现数字高程数据误差在空间分布的差异性和相关性,反映了误差平均值和方差的变化。全局 Moran's *I* 的期望值为:随着样本数 *n* 的增大,期望值  $1/(n-1)$  将逐渐趋于 0。*I* 的值介于 -1~1 之间,当 *I* 大于期望值时,表示数字高程数据误差在整个区域内存在空间正相关,即较大(或较小)的误差在空间上显著集聚;当 *I* 小于期望值时,表示数字高程数据误差在整个区域内存在空间负相关,说明相邻空间单元的误差值呈此长彼消的状态;*I* 的值越接近 1 或者 -1,表示空间自相关的程度越强烈;当 *I* 接近期望值时,则表明数字高程数据误差不存在空间自相关现象。Moran's *I* 的结果在很大程度上取决于所选择的空间关系矩阵。

### 1.3 局部空间自相关分析

全局 Moran's *I* 是研究区域上数字高程数据误差的空间自相关的综合度量,仅说明整个区域上各格网点高程数据误差与相邻格网点高程数据误差之间空间差异的平均程度。在整个区域单元上高程数据误差的空间差异较小的情况下,局部区域单元上的高程数据误差的空间差异有可能较大。为了全面反映区域单元上误差空间差异的变化趋势,还需采用 ESDA 局部分析方法。1995 年,Anselin 提出了空间关联的局部指标 LISA (local indicators of spatial association)<sup>[8]</sup>,可以揭示局部每个单元上误差的空间自相关性质。LISA 本质上是将 Moran's *I* 分解到各个区域单元,对某个格网点 *i* 可表示为:

$$I_i = (\epsilon_i - \bar{\epsilon}) \sum_{j=1, j \neq i}^n w_{ij} (\epsilon_j - \bar{\epsilon}) / S_2^2 \sum_{j=1, j \neq i}^n w_{ij} \quad (2)$$

式中,  $S_2^2 = \frac{1}{n} \sum_{j=1, j \neq i}^n (\epsilon_j - \bar{\epsilon})^2$ 。

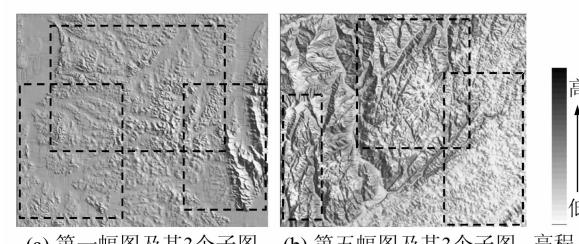
局部 Moran's *I* 可以看作是全局 Moran's *I* 的各区域分量,反映了某格网点周围误差的空间集聚程度。*I<sub>i</sub>* 的显著性可以采用 Bonferroni 标准加以判断,当总的显著性水平设定为 *a* 时,每一个区域的显著性要根据 *a/n* 的原则进行判断。随着样本数 *n* 的增大,期望值逐渐趋于 0,若 *I<sub>i</sub>* 显著大于 0,表明存在空间正相关现象,即说明格网点 *i* 的高程数据误差与周围相邻格网点的高程数据误差呈空间集聚现象,高程数据误差的空间差异较小;若 *I<sub>i</sub>* 显著小于 0,表明存在空间负相关现象,即说明格网点 *i* 的高程数据误差与周围相邻格网点的高程数据误差的空间差异较大。

可以用 Moran 散点图来描述,其横坐标为各格网点对应高程数据误差标准化处理后的值,纵坐标为其空间连接矩阵所决定的相邻格网点的高程数据误差值的平均值(经过标准化处理)。Moran 散点图划分为 4 个象限,表达了某一格网点对应的高程数据误差和其周围格网点对应的高程数据误差四种类型的局域空间联系,即“高高”(第一象限)、“低高”(第二象限)、“低低”(第三象限)和“高低”(第四象限)。“高高”和“低低”表明误差具有较高的空间正相关,提示区域的集聚和相似性;“高低”和“低高”则表示误差存在较强的空间负相关,区域具有异质性。

## 2 数字高程数据误差的来源与分析

### 2.1 研究区域与数据来源

本文选用 5 幅 1:5 万的 DEM 数据(12 km × 18 km)进行实验,地形因子的变化范围如下:坡度变化范围为 3°~45°,高程变化范围为 10~1 400 m,地表粗糙度为 0.198 558~3.006 29。取 5 幅图及每幅图的 3 个大小不等的子图作为实验区,共 20 个实验区,实验区的编号以地形的平均坡度为依据从小到大排列,如第一幅图为实验区 4,其 3 个子图分别为实验区 1、2、5;第五幅图为实验区 15,其 3 个子图分别为实验区 11、17、19,它们的 DEM 示意图如图 1 所示。对各实验区的数字高程数据误差的全局自相关性和局部自相关性进行研究,每个格网点的高程数据误差是利用次相邻的 4 个格网数据采用内插所得的数据与原始数据相减得到的,分别用双线性内插、样条函数内插等几种内插模型作实验,所得的结论是一致的。文中数据采用常用的双线性内插模型表示地形表面时产生的高程数据误差。



(a) 第一幅图及其3个子图 (b) 第五幅图及其3个子图 高程

图 1 部分实验区 DEM 示意图

Fig. 1 DEM Sketch Maps of Several Empirical Areas

### 2.2 计算结果与分析

#### 2.2.1 全局空间自相关分析

利用全局 Moran's *I* 指数计算上述 20 个实

验区的地形线性表达产生的高程数据误差的全局自相关指数,数据见表1。

表1 数字高程数据误差的全局自相关指数

Tab. 1 Global Moran's  $I$  of Digital Elevation Error

实验区	平均坡度/(°)	平均高程/m	地表粗糙度	误差的全局相关性指数	局部相关性指数符号相同的数值所占的比例
1	1.686 35	57.999 2	0.198 558	-0.130 950	0.995 084
2	3.914 88	72.651 5	0.368 830	-0.014 648	0.991 873
3	5.370 08	82.009 0	0.164 070	0.003 109	0.993 963
4	7.612 58	104.479	0.653 724	0.076 941	0.991 526
5	10.653 4	120.751	1.091 900	0.085 194	0.982 014
6	12.094 2	127.621	1.050 310	0.134 632	0.985 242
7	14.107 7	148.649	1.099 090	0.169 082	0.984 816
8	15.848 4	167.461	1.136 74	0.190 289	0.984 265
9	20.329 5	249.903	1.556 54	0.199 143	0.986 883
10	23.865 3	623.104	1.590 93	0.236 381	0.987 812
11	26.625 4	250.857	1.615 99	0.269 976	0.986 196
12	27.187 0	580.385	1.673 35	0.256 384	0.988 144
13	29.992 1	502.705	1.947 23	0.296 152	0.983 624
14	30.572 2	575.067	1.862 06	0.283 287	0.984 732
15	31.175 8	512.766	2.005 72	0.270 956	0.983 425
16	33.171 1	570.873	2.007 76	0.305 316	0.982 382
17	35.195 8	659.960	2.121 02	0.317 445	0.982 808
18	38.959 7	714.081	2.629 30	0.294 310	0.978 598
19	42.273 8	712.789	1.102 76	0.270 758	0.973 858
20	44.030 9	1130.16	3.258 6	0.338 188	0.980 053

根据表1的实验结果可以得出,数字高程数据误差的全局 Moran's  $I$  指数趋近于0。由前面的分析知,利用双线性曲面表示地形表面时产生的高程数据误差在整个区域单元上的分布不存在显著的全局空间自相关,表明邻近区域单元上的高程数据误差之间的关系在整体上既不综合表现为趋同,也不综合表现为趋异,高程数据误差的整体空间格局为随机格局。

将实验数据用散点图表示,如图2所示。由图2也可以看出,利用双线性曲面表示地形表面时产生的数字高程数据误差,当地势比较平坦时,不存在全局相关性,随地形坡度和地表粗糙度的增大,全局相关性指数也有增大的趋势,但全局空间自相关仍不显著,数字高程数据误差在整体上的分布格局为随机格局。

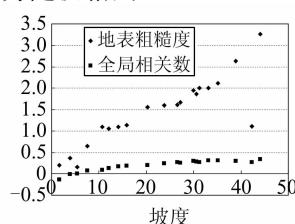


图2 数字高程数据误差的全局自相关指数与坡度、地表粗糙度关系图

Fig. 2 Connection Between Global Moran's  $I$  of Digital Elevation Error, Slope and Surface Roughness

### 2.2.2 基于格网单元的局部空间自相关分析

利用局部 Moran's  $I$  指数计算上述20个实验区的数字高程数据误差的局部自相关指数,由于数据量太大,绘制的表格无法放入文中,只绘制出数字高程数据误差的局部相关性指数符号相同的数值所占比例的 Moran 散点图。因篇幅有限,任意取表1中的实验区4、7、12、15这4幅图的局部相关性分布图,如图3所示。

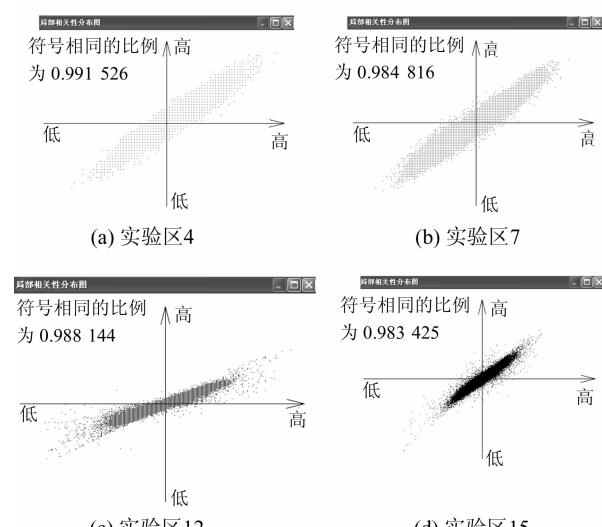


图3 几个实验区的局部自相关指数分布图

Fig. 3 Local Moran's  $I$  Diagrams of Several Empirical Areas

图3中几幅图颜色上的差异与实验结果中数据的重复量和选择的格网数据量的多少有关。由图3可知,数字高程数据误差在局部区域单元上,局部相关性指数符号相同的数值所占的比例都在97%以上,表明实验区中一点的高程数据误差较高,则其周围格网点的高程数据误差也较高;反之,该点的高程数据误差较低,则其周围格网点的高程数据误差也较低,利用双线性曲面表示地形表面时产生的高程数据误差在局部区域单元上表现出非常明显的空间集聚现象,误差的空间差异很小,在局部区域具有很高的空间正相关性。

### 3 结语

通过本文的研究,认识了数字高程数据误差的空间分布特征,揭示了数字高程数据误差的差异在空间上的相互关系,为数字高程数据误差的微观模拟问题提供了重要的理论基础,对于解决实际工作中估算数字高程数据误差在具体栅格点位的大小提供了客观依据,可用于数字高程模型的精度评定和分析应用等方面的研究。

### 参考文献

- [1] Oksanen J, Sarjakoski T. Uncovering the Statistical and Spatial Characteristics of Fine Toposcale DEM Error[J]. International Journal of Geographical Information Science, 2006, 20(4): 345-369

- [2] 史文中. 空间数据误差处理的理论与方法[M]. 北京:科学出版社,1998
- [3] Monckton C. An Investigation into the Spatial Structure of Error in Digital Elevation Data[M]. London: Taylor and Francis, 1994
- [4] 汤国安,龚键雅,陈正江,等. 数字高程模型地形描述精度量化模拟研究[J]. 测绘学报,2001,30(4): 361-365
- [5] 周启鸣,刘学军. 数字地形分析[M]. 北京:科学出版社,2006:217-221
- [6] 刘学军,王叶飞,曹志东,等. 基于DEM的坡度坡向误差空间分布特征研究[J]. 测绘通报,2004(12): 11-13,37
- [7] Arthur G, Keith O J. The Analysis of Spatial Association by the Use of Distance Statistics[J]. Geographical Analysis, 1992, 24(24): 189-206
- [8] Anselin L. Local Indicators of Spatial Association—LISA[J]. Geographical Analysis, 1995, 27(2): 93-116
- [9] 朱长青,史文中. 空间分析建模与原理[M]. 北京:科学出版社,2006
- [10] 张磊,伍吉仓,陈艳玲. InSAR高程模型及其精度分析[J]. 武汉大学学报·信息科学版,2007,32(2): 108-111,119

**第一作者简介:**王耀革,博士生。主要从事数学应用和空间数据不确定性的研究。

E-mail:wyg711218@163.com

## Analysis of Digital Terrain Error Based on Spatial Autocorrelation

WANG Yaoge<sup>1</sup> WANG Zhiwei<sup>1</sup> ZHU Changqing<sup>1,2</sup>

(1) School of Sciences, Information Engineering University, 73 Songshan South Road, Zhengzhou 450001, China)

(2) Key Laboratory of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University,  
1 Wenyuan Road, Nanjing 210046, China)

**Abstract:** Using spatial autocorrelation analysis, the digital elevation error is studied; and the results show that there is no significant global spatial autocorrelation of digital elevation error in all cells, the spatial pattern of digital elevation error is random on the whole. There is significant local spatial autocorrelation between one cell and its neighbors, and the difference of digital elevation error between one cell and another is small, the spatial pattern of digital elevation error converges strongly in local areas.

**Key words:** digital elevation error; spatial autocorrelation; spatial pattern