

文章编号: 1000-0585(2002)05-0578-07

# 分布式侵蚀预报模型中网格面积的选定 ——以黄土高原丘陵沟壑区为例

卫海燕, 张科利, 王敬义

(北京师范大学资源与环境科学系, 北京 100875)

**摘要:** 在分布式土壤侵蚀过程模型中, 承载数据以及进行运算的最小单元, 即基本地块的选取是非常关键的, 它直接关系到模型的模拟精度和运算数据量。目前大多数分布式模型都采用平均布设矩形网格的方法, 这种方法在基本网格的大小选取上存在着盲目性和不统一性。本文以黄土高原丘陵沟壑区为例, 利用 GIS 和 SPSS 分析了黄土高原丘陵沟壑区属性均一的基本地块在面积上的统计规律, 给出了基本地块选取的合理依据, 以促进建立更好的分布式模型。

**关键词:** 基本地块; 分布式模型; 黄土高原丘陵沟壑区

**中图分类号:** S157.1      **文献标识码:** A

## 1 前言

土壤侵蚀是人们普遍关注的生态环境问题之一。开展土壤侵蚀预报是有效监测水土流失和评估水保措施的手段, 侵蚀模型则是进行土壤流失预报的重要工具。自 20 世纪 60 年代以来, 国内外已经开发出许多实用的侵蚀预报模型<sup>[1~3]</sup>。根据目的及适用范围, 土壤侵蚀模型又可分为坡面模型和流域模型。流域土壤侵蚀预报模型分为集总模型和分布式模型两种。集总模型反映流域的总体或者平均行为。分布式模型则将流域划分为若干网格, 通过对每个网格的赋值来反映影响土壤侵蚀的各种因素在流域内的差异, 然后根据一系列反映侵蚀过程的运算程序计算各个网格的产流产沙, 最后进行合理合并, 从而达到比较准确地预报整个流域产流产沙的目的。分布式模型比集总式模型能更准确地反映自然小流域的侵蚀状况。

GIS 和 RS 技术的进步使分布式模型得到了长足的发展。许多研究者试图结合 GIS 和 RS 在采集、存储、管理空间数据上的有效性, 研制出能够分析预报各种尺度流域产沙产流过程的模型<sup>[4~6]</sup>。但如何使分布式模型的计算方法更科学、更真实地反映侵蚀的自然过程, 同时能使计算结果的误差最小是目前人们普遍关注的问题。作为采集和承载数据, 以及模型运算最小单元的网格(地块)的选取是解决这一问题的关键。

合理的网格应该具有以下特征: 首先应能尽量准确地反映区域地表要素的空间差异。当流域被赋予各种属性要素后, 它的每个网格所表现的综合特征应该均一, 并且与相邻单

收稿日期: 2002-03-08; 修订日期: 2002-06-22

基金项目: 国家重点基础研究发展规划项目(G1999043605)资助

作者简介: 卫海燕(1978-), 女, 西藏昌都人, 博士生。专业方向为土壤侵蚀与水土保持。Email: hywei@mail.bnu.edu.cn

元格不同。其次, 在保证真实反映地表状况精度的同时, 网格数目应尽可能的少, 以减少数据冗余, 缩减工作量。

在目前大多数模型中, 多是通过在流域内平均分布等大的矩形栅格来提取处理数据<sup>[7-9]</sup>。这种平均分布网格的处理方法存在着一定问题: ①单元格大小的选择具有随意性, 不统一; ②单元格的大小直接决定了是否准确反映了地表特征差异, 以及模型的精度, 同时也决定了数据量。但精度和数据量之间却存在着矛盾, 如果单元格小, 那么精度就高, 但数据量就大; 反之, 若单元格偏大, 数据量会减少, 但也导致精度下降。如何选取合理的单元格大小以保证精度又保持合适的工作量, 在很大程度上应该从研究区的地形条件和模型需要达到的精度出发。但在实际工作中, 基本网格大小的选择往往根据经验而定, 缺少理论根据。黄土高原丘陵沟壑区地貌类型复杂, 地形特征区域变异很大, 怎样布设网格才能解决精度和数据量之间的矛盾还需要作深入研究。

本文以陕西省安塞县纸房沟流域为例, 利用 GIS 技术将土地利用图和坡度图叠加, 得到流域内属性均一的基本地块, 然后用 SPSS 分析了地块面积的统计规律和空间分布情况。其结果可为分布式模型的网格面积大小的确定提供依据。

## 2 资料与方法

陕西省安塞县纸房沟流域面积  $8.27\text{km}^2$ , 土壤类型为黄土母质上发育的黄棉土。流域内地形破碎, 沟壑纵横。35° 以上的坡耕地占土地总面积的 31.19%, 是典型的黄土高原梁峁丘陵地貌, 该流域在黄土高原丘陵沟壑区具有一定的代表性。影响土壤侵蚀的因素有降雨、土壤、地形、土地利用类型等。对于小流域来说, 可以认为流域内的土壤及降雨分布是均一的, 分析计算土壤侵蚀只需考虑地形和土地利用类型。因此, 分布式模型的网格内只要地形和土地利用类型一致, 就可以提取及承载具有代表性的数据。所以本文选择了纸房沟小流域坡度等级图和土地利用图进行叠加来产生地块, 可以认为叠加得到的网格就是流域产流产沙的实际分异单元。文中采用的坡度等级图和土地利用图选自国家“七五”攻关课题“黄土高原综合治理定位试验示范综合研究”的研究图集中纸房沟流域的坡度分级和土地利用图<sup>[10]</sup>。坡度等级图将流域内土地分为 8 个等级, 土地利用图按照土地利用类型的不同分共为 17 种 (表 1)。

表 1 坡度与土地利用分级表  
Tab 1 Slope gradient grades and land use types

级别	a	b	c	d	e	f	g	h
坡度等级	< 3°	3~ 5°	5~ 8°	8°~ 15°	15°~ 25°	25°~ 35°	> 35°	水平梯田
代号	1	2	3	4	5	6		
土地利用类型	沟台旱耕地	沟台水浇地	灌木林	果园	居民点	裸岩		
代号	7	8	9	10	11	12		
土地利用类型	明沙地	旱耕地	乔木林	人工草地	疏林地	水库水面		
代号	13	14	15	16	17			
土地利用类型	塌地旱地	梯田旱地	天然草地	未成林	坟地			

具体的操作步骤是先将纸房沟流域的地面坡度等级图和土地利用图扫描，在 Mapinfo 中进行屏幕数字化（见图 1，图 2），然后在 PC Arc/info 中将两幅图件转化为数字矢量图，进行叠加，叠加后用 Arc/info 的 Eliminate 功能将图斑面积小于  $20\text{m}^2$  的细小多边形合并到相邻的大多边形中（消除细碎多边形），将纸房沟流域分割为属性均一的基本地块（网格）。最后计算这些基本地块的面积，用统计软件 SPSS 进行统计分析。

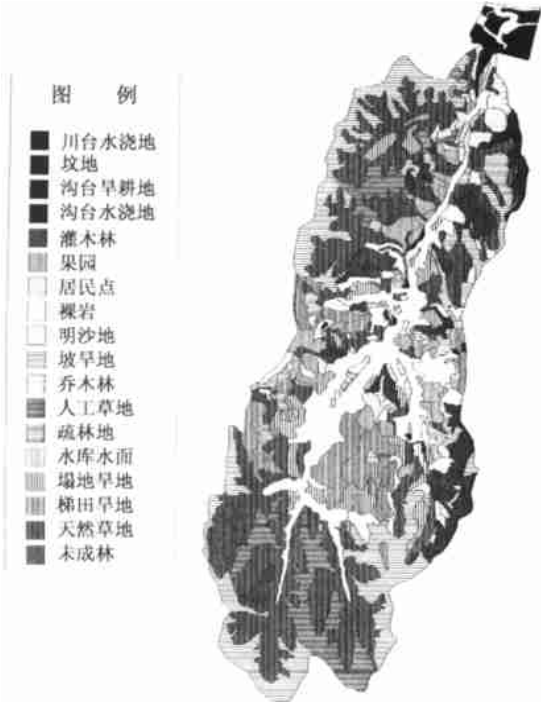


图 1 纸房沟流域土地利用类型图

Fig. 1 Land use map of the Zhifanggou catchment

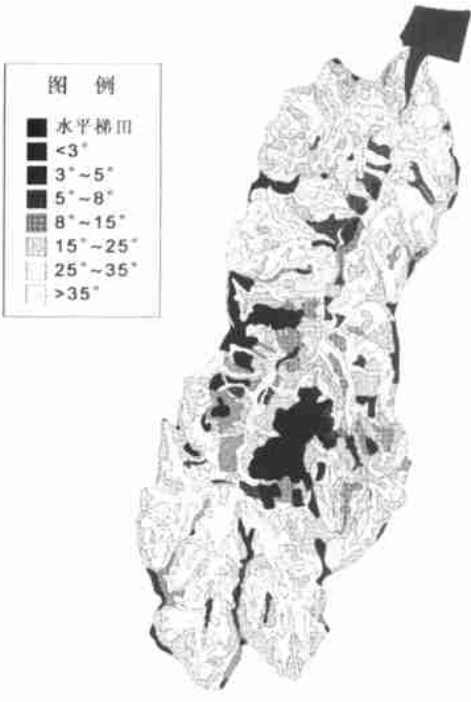


图 2 纸房沟流域坡度分级图

Fig. 2 Slope gradient map of the Zhifanggou catchment

### 3 结果与讨论

#### 3.1 基本地块面积的统计规律

按照上述处理方法，纸房沟流域坡度等级图和土地利用图矢量化后产生的图斑数分别为 383 个和 293 个，叠加后产生的网格数为 2293 个，消除细碎多边形后图斑数为 2083 个。也即，在只考虑坡度和土地利用变化时，面积为  $8.27\text{km}^2$  的小流域可以分割为 2083 个属性相对均一的地块。对这些地块的面积大小进行统计分析，其统计特征值如表 2。

表 2 网格面积统计分析结果

Tab 2   Statistic values of the patches' area								
特征值	平均值	中位数	标准差	最大值	最小值	跨度	Skewness ( 偏度)	Kurtosis ( 峰度)
面积 ( m <sup>2</sup> )	4368. 16	1462. 42	10974. 33	196029. 7	0. 37	196029. 3	8. 447	98. 857
百分位数	5		25		50		75	95
面积 ( m <sup>2</sup> )	39. 2096		332. 988		1462. 42		4227. 65	16304

表 2 中数据表明, 在黄土高原丘陵沟壑区, 坡度和土地利用相对均一的地块面积变化跨度很大, 也就是说, 基本地块最大面积和最小面积相差较大, 流域内既有大块的产流产沙性质均一的地块, 同时也有切割细碎的属性相同的小面积地块。其中, 面积超过 10 万平方米的 7 个大网格总面积占流域总面积的 10.34%。另一方面, 尽管基本地块面积变化较大, 但集中程度较好。Skewness= 8.447 说明偏正态分布, Kurtosis= 98.857 表明峰度高, 集中趋势明显。百分位数的统计结果还显示, 有 50% 的地块面积集中在  $1462.42\text{m}^2$  以下, 有 75% 的地块面积集中在  $4227.65\text{m}^2$  以下。

地块面积的集中程度还可以从表 3 更明显地看出。小于  $500\text{m}^2$  的地块个数占总地块个数的 30.44%, 面积在  $1000\text{m}^2$  到  $5000\text{m}^2$  之间的地块最多, 占总地块个数的 36.1%, 大于  $5000\text{m}^2$  的地块个数只占 21.56%。可以看出, 叠加后产生的地块面积集中在小于  $500\text{m}^2$  以下以及  $1000\text{m}^2 \sim 5000\text{m}^2$  之间的范围。需要指出的是, 地图扫描和数字化的误差可能增加了小地块的数量。

表 3 图斑面积频度分布表  
Tab 3 Frequency distribution of the patches' area

面积范围 ( $\text{m}^2$ )	0. 3~ 500	500~ 1000	1000~ 5000	5000~ 40000	40000~ 200000
个数	634	248	752	422	27
百分比 (%)	30. 44	11. 91	36. 1	20. 26	1. 3

表 4 在叠加结果统计分析的基础上, 列出了如果以平均值、中位数、最大值、最小值以及 5 个百分位数所代表的网格为基本网格, 相应的平均分布需要采用的正方形网格大小和网格数量。举例说明, 如果平均分布的网格精度至少要保证叠加结果中百分位数为 25 的面积  $333\text{m}^2$  (也就是说, 属性均一的地块在数量上有 25% 的地块面积小于基本网格的面积), 那么平均分布方式所选用的单元网格至少应该是  $18\text{m} \times 18\text{m}$ , 这样最后得到的网格数目为 28083 个。可以看出, 平均分布式的网格数量在这里远远超过了本文中叠加产生的网格数量, 这是因为平均分割的方法在大面积属性相同的地块上存在着数据冗余。而从前面的分析我们已经看到, 在黄土高原丘陵沟壑区, 属性均一的大面积地块确实存在。从表 4 中还可以看到在采用平均分布方法时, 模拟精度和数据量之间存在着不可避免的矛盾。例如当精度至少保证百分位数为 50 的面积  $1462\text{m}^2$  时, 网格数目比采用百分位数 25 相应的面积  $333\text{m}^2$ , 减少了 3 倍。

尽管平均分布式存在着数据冗余以及工作量和精度之间的矛盾, 目前仍是人们在分布式模型中最常使用的一种方法。建议在使用平均分布式时, 基本网格的选取从区域的地表状况和模型模拟的需要出发考虑。当网格过大, 使得精度不足以表现模型所需要的地表差异时, 模拟的结果就可能因之误差很大; 另一方面, 不顾忌数据量而一味地追求精度, 采用过小的网格, 不仅是一种浪费, 而且也不能达到好的模拟效果。但是在选择了一种大小的网格后, 究竟这种基本网格所能保证的精度如何, 往往是不可知的。这便增加了基本网格选取的困难和盲目性。本文的结果将对黄土高原丘陵沟壑区分布式模型中基本网格大小的选取有一定的参考价值。例如上文所述, 如果采用  $18\text{m} \times 18\text{m}$  的网格, 那么有 75% 的基本地块面积是大于基本网格面积的, 这样便能使模型设计者和使用者对模型的精度有一个大概的估计。

表 4 图斑面积统计分析结果与平均分布网格的比较

Tab 4 Comparison between the overlay method and the rectangle- distributed method

叠加产生的图斑面积统计分析结果 (m <sup>2</sup> )		相应的平均分布结果	
		网格大小 (m× m)	网格个数
平均值	4368. 1589	66× 66	2089
中位数	1462. 419	38× 38	6301
最大值	196029. 7	443× 443	46
最小值	0. 37307	0. 61× 0. 61	24457768
百分位数 5	39. 2096	6× 6	252747
百分位数 25	332. 988	18× 18	28083
百分位数 50	1462. 42	38× 38	6301
百分位数 75	4227. 65	65× 65	2154
百分位数 95	16303. 996	128× 128	555

3. 2 基本地块的属性差异

表 5 坡度分级和土地利用交叉统计结果表 (地块数)

Tab. 5 Land use vs slope gradient crosstabulation

土地利用	坡度分级								合计
	水平梯田	< 3°	3~ 5°	5~ 8°	8~ 15°	15~ 25°	25~ 35°	> 35°	
果园	18	2		8	9	11	38	18	104
乔木林	22	7		8	20	61	83	48	249
疏林地	11		4	1	11	17	30	19	93
未成林	2	6		2	4	13	35	16	78
灌木林	2	4	4	4	11	30	56	35	146
天然草地	34	18	11	34	47	143	235	105	627
人工草地	4	3	5	6	4	26	54	48	150
居民点	3	3			1	4	9	8	28
明沙地		1							1
裸岩	2	3		1	1	2	11	5	25
坟地	2					4			
川台水浇地		6							6
沟台水浇地		3							3
沟台旱浇地	1		1	1	2	3	2		10
梯田旱地	35	1	1	5	11	31	31	14	139
塌地旱地	12			7	15	27	38	31	130
坡旱地	17	9	8	4	14	47	114	77	290
水库水面		1				6	1		8
合计	165	67	34	81	150	425	737	424	2083

在 SPSS 中，对叠加产生的地块进行交叉统计，得到了所有的地块在坡度和土地利用类型上的交叉分布情况，反映了安塞纸房沟流域所有的地块在属性分配上的差异，结果见

表 5。从表 5 中可以看到：①交叉结果中网格数目最多的是坡度为 25°~ 35°的天然草地，其次是 15°~ 25°的天然草地，25°~ 35°的坡旱地和 35°以上的天然草地。这四种网格的面积分别占总面积的 12.67%，4.31%，6.28%，14.23%。说明流域内陡坡草地的比例非常大。②叠加后产生的网格中，存在着属性不协调的情况。例如梯田和裸岩叠加产生了 2 个网格，这与实际情况很可能是不相符合的。但是这部分地块在面积上和数量上所占的比例都不多。这些属性不协调的地块的产生是由于原始图件的不精确性、数字化的误差以及叠加运算的系统误差等引起的。刘文宝、戴洪磊等<sup>[11, 12]</sup>研究表明在 GIS 的叠加过程中会产生属性不协调，原因主要来自数据源的不确定性、数据建模的不确定性和分析过程中引入的不确定性等。

4 结论

- (1) 可以用坡度图和土地利用图叠加得到流域产流产沙计算的基本地块。每地块内部都是属性均一的，并且有别与任何相邻的地块。每一个基本地块都可以作为承载各种属性的基本单元，并且不存在有相邻的地块需要合并的情况。
- (2) 在黄土高原丘陵沟壑区，坡度和土地利用相对均一的地块面积跨度很大，可以从零点几平方米到几十万平方米。
- (3) 产流产沙基本地块的集中程度较好，有近 67% 的基本地块的面积集中在 500m<sup>2</sup> 以下和 1000m<sup>2</sup>~ 5000m<sup>2</sup> 之间。这个结论对黄土高原丘陵沟壑区分布式模型中基本网格大小的选取提供了依据。分布式模型的基本网格的选取必须从模型设计精度和区域地表差异状况出发考虑。参考本文的分析结果，将使黄土高原丘陵沟壑区侵蚀模型的设计者和使用者对模型的精度有近似的估计。

参考文献：

[ 1 ] Wischmeier W H, Smith D D. Predicting rainfall erosion losses. USDA Agricultural Handbook No. 537, 1978.

[ 2 ] Renard K G, *et al.* RUSLE- A guide to conservation planning with the revised universal soil loss equation. USDA Agricultural Handbook No. 703, 1997.

[ 3 ] Beasley D B, *et al.* ANSWERS: A model for watershed planning. Transactions of the ASAE, 1980, 23: 938~ 944.

[ 4 ] 刘卫国, 吕鸣伦. 地理信息系统和遥感技术支持下的山地环境梯度分析方法研究. 地理研究, 1997, 16(3): 63~ 69.

[ 5 ] 李清河, 李昌哲, 齐实, 等. 流域降雨径流路径的数字模拟技术. 地理研究, 2000, 19(2): 209~ 216.

[ 6 ] 阎国年, 钱亚东, 陈钟明. 黄土丘陵沟壑区沟谷网络自动制图技术研究. 测绘学报, 1998, 27(2): 131~ 137.

[ 7 ] 陈永宝, 陈志伟, 郭伟民. 应用 GIS 实现水土流失定量遥感监测. 水土保持研究, 1998, 5(2): 152~ 158.

[ 8 ] 高志强, 刘纪远. 基于遥感与 GIS 的中国土地资源生态环境及全国人口分布的关系研究. 遥感学报, 1999, 3(2): 67~ 70.

[ 9 ] 林辉, 彭长辉. 地理信息系统中栅格单元大小和形状的选择. 遥感信息, 2001, (1).

[ 10 ] 中国科学院、水利部西北水土保持研究所. 黄土高原综合治理试验示范区专题地图集. 北京: 测绘出版社, 1991.

[ 11 ] 刘文宝, 邓敏, 夏宗国. 矢量 GIS 中属性数据的不确定性分析. 测绘学报, 2000, 29(1): 76~ 81.

[ 12 ] 戴洪磊, 徐萍林, 杜道生. 矢量 GIS 环境下多边形叠置中位置误差的传播分析. 测绘学报, 1998, 28(2): 148~ 152.

# A study on the chosen land patch in the distributed erosion model: taking the hilly loess region as an example

WEI Hai-yan, ZHANG Ke-li, WANG Jing-yi

(Department of Resources and Environmental Science, Beijing Normal University, Beijing 100875, China)

**Abstract:** Distributed erosion model is one of the effective methods to evaluate soil loss. The land patch is a key component in establishing a distributed erosion model. Now many researchers prefer rectangle as cells in the model. However, they had not made an agreement about the cell size. Furthermore, when the rectangle cells are employed, it is somehow difficult to keep a balance between the precision and the load of the model, because both of them are decided from the size of the cell. The larger the cell is, the more laborious the work is, but it is not always the case that the model will be more precise.

Actually, the cell size should be determined according to the uniformity of the land in the region where the model is to be applied and the accuracy of the model needed. However, there is a chance that the user of the model, and even the designer of the model, have no actual image of the region, and then have no idea about what size the cell should be. Similarly, when a cell is chosen, the accuracy is not clearly presented.

Zhifanggou catchment is a typical area of the hilly loess region in the Loess Plateau. Supported by GIS and SPSS, the authors got the patches of the catchment which have uniform characters respectively, and analyzed the statistic result of the patches' areas. It was indicated that the areas ranged from less than  $1\text{ m}^2$  to almost  $200,000\text{ m}^2$ . On the other hand, nearly 67% of the patches congregate into two groups: either below  $500\text{ m}^2$  or between  $1,000\text{ m}^2$  to  $5,000\text{ m}^2$ . According to the results, one can get an approximate image of the hilly loess region and then be more reasonable to establish or use a distributed erosion model for that region.

**Key words:** land patch; distributed erosion model; hilly loess region in the Loess Plateau

---

致歉: 2002年(第21卷)第3期《地理研究》发表的《近期中国省际经济社会要素流动的空间特征》一文,在撰写过程中未经作者同意,使用了北京大学匿名评审阶段的博士学位论文《开放条件下中国城市经济区研究》中人口迁移流、铁路客流、航空客流、信件流部分的数据资料;虽然在初稿参考文献中注明了该篇博士论文,但是由于作者及编校工作的疏漏,在文章发表时参考文献中却未能注明,在此一并向论文作者张莉博士表示诚挚的歉意。 张敏