

坡谱信息熵尺度效应及空间分异

李发源^{1,2}, 汤国安¹, 贾旖旎¹, 曹志冬³

(1 南京师范大学虚拟地理环境教育部重点实验室, 南京 210046; 2 中国科学院成都山地灾害与环境研究所, 成都 610041;

3 中国科学院地理科学与资源研究所, 北京 100866)

摘要: 选取陕北黄土高原 48 个不同地貌类型区, 以其对应的 1 10 000 及 1 50 000 数字高程模型 (DEM) 为数据源, 运用比较分析与数理统计的方法, 分析了坡谱信息熵在陕北黄土高原的空间分异特征及尺度效应。结果显示, 坡谱的信息熵在一定程度上可以反映地表的复杂度, 坡谱信息熵和沟壑密度间有较好的幂函数关系, 随着坡谱信息熵的增大, 样区沟壑密度也增大。坡谱信息熵的空间分异和陕北黄土高原的黄土地貌形态在空间上的变异是相关的, 可将其作为地貌类型划分的判别因子之一。

关键词: 坡谱; 信息熵; DEM; 尺度; 空间分异; 黄土高原

1 引言

坡度是数字地形分析中最重要的地形因子之一。对于有连续起伏的实际地面, 坡度多指该地面的平均坡度。虽然平均坡度在一定程度上能够反映地面的倾斜程度与起伏特征, 但相同的平均坡度地面的实际坡度组合往往并不相同。因此, 对于一个特定的统计区域, 地面坡度的组合特征更能反映地面的起伏变化特征。为定量地描述地面坡度的组合, 汤国安提出了坡谱的概念, 即在一定区域内不同级别坡度组合关系的统计模型。但是, 还没有对坡谱的统计模型量化表达的标准, 这为坡谱进一步的应用带来了较大的困难。本文试图以坡谱信息熵来量化表达坡谱特征, 并探讨坡谱信息熵的空间分异特征。

2 坡谱信息熵及其实验

2.1 坡谱信息熵

信息熵的概念是由 Shannon 于 1948 年提出, 经过半个多世纪的发展, 其应用几乎遍及了自然科学领域诸学科^[1]。信息熵在地形地貌研究中的应用, 首先是从地形图开始的。Sukhov (1967, 1970) 将

Shannon 公式用于地图信息量的测度, 即将 P_i 作为地图上某种地图符号的频率。这种测度方法忽视了地图符号间的空间关系, 因此 Neumann (1994) 等提出了一种估算地图符号拓扑信息的测度方法, 即利用 Rashevsky (1955) 年提出的对偶图的概念来表达地图拓扑信息, 然后利用 Shannon 公式来测度信息熵。首先在地貌研究中引入熵概念的是 Leopold 和 Langbein, 他们提出了地貌学熵的定义, 但他们的定义难以量算, 没有得到进一步研究。艾南山等在 Davis 和 Strahler 地貌发育模型的基础上, 通过构造 Strahler 曲线的密度函数, 推导了流域地貌系统的地貌信息熵^[2,3]。马新中等提出流域系统的超熵并用于流域系统的定量研究^[4]。陈楠、王雷、张茜等人都曾利用信息熵研究不同比例尺 DEM 信息容量的变化^[5-8]。

结合前人研究经验, 本文定义坡谱信息熵为:

$$H = - \sum_{i=1}^m P_i \ln P_i \quad (1)$$

其中 m 为分级数, P_i 为每一级别坡度的频率, 定义 $h_i = -P_i \ln P_i$ 为第 i 级分组的信息量。坡谱的信息熵反映坡谱的均匀度或坡谱内各组数据频率的差异程度, 频率越接近, 其信息熵越大。就坡谱信息熵的地貌意义而言, 可以认为在地表起伏平缓的区域, 如淳化样区 (黄土塬地貌), 由于坡度组合也较

收稿日期: 2006-07-29; 修回日期: 2007-02-20.

资助项目: 国家自然科学基金 (40571120, 40671148), 教育部博士点基金 (20050319006).

作者简介: 李发源 (1978-), 男, 博士研究生, 主要从事数字地形分析、黄土高原自然地理的研究。E-mail: li_fayuan@sina.com

简单, $0 \sim 9^\circ$ 坡度级别对应的地面占据了样区的大部分面积, 所以坡谱内各组数据频数的差异较大, 坡谱的信息熵也小, 淳化样区坡谱信息熵 $H_{30} = 2.149 \text{ nat}$ (30 表示分级数); 而在地表起伏较大的区域, 如延川样区 (黄土梁峁状丘陵沟壑地貌), 坡谱内各组数据频数的差异较黄土塬区小, 坡谱的信息熵也随之增大, 延川样区坡谱信息熵 $H_{30} = 2.809 \text{ nat}$ 。

2.2 实验数据及方法

黄土高原是我国具有独特自然地理景观的区域之一, 一直以来都是国内外地学研究的热点。具有不同地面复杂度、粗糙度的多种地貌类型在黄土高原呈现有规律的空间分异, 是地貌研究最佳的实验场。同时, 黄土高原地区 DEM 建设, 覆盖全区的 1:100 000、1:250 000、1:500 000、1:100 000 比例尺数据都已完成, 为研究提供了极大的方便。本文以陕北黄土高原作为研究区域, 由南到北选取了 48 个样区 (图 1), 以其对应的 1:100 000 及 1:500 000 DEM 为数据源, DEM 是采用国家测绘局编制的 1:100 000、1:500 000 地形图 (等高距分别为 10 m、20 m), 通过扫描、数字化、构建 TIN、高程内差建立的, 水平分辨率分别为 5 m 和 25 m。

为了深入剖析不同比例尺 DEM 信息量的尺度效应, 在以上 48 个样区中又选取了神木、绥德、延川、甘泉、宜君和淳化 6 个典型样区作为研究重点, 代表为风沙-黄土丘陵过渡地貌、黄土峁状丘陵沟壑地貌、黄土梁峁状丘陵沟壑地貌、黄土梁状丘陵地貌、黄土残塬丘陵沟壑地貌、黄土塬地貌。坡谱信息熵的提取流程如图 2 所示。

3 样区实验与结果分析

3.1 坡谱信息熵空间分异

首先计算了陕北黄土高原 48 个样区 1:100 000 及 1:500 000 比例尺 DEM 的坡谱信息熵值。结果显

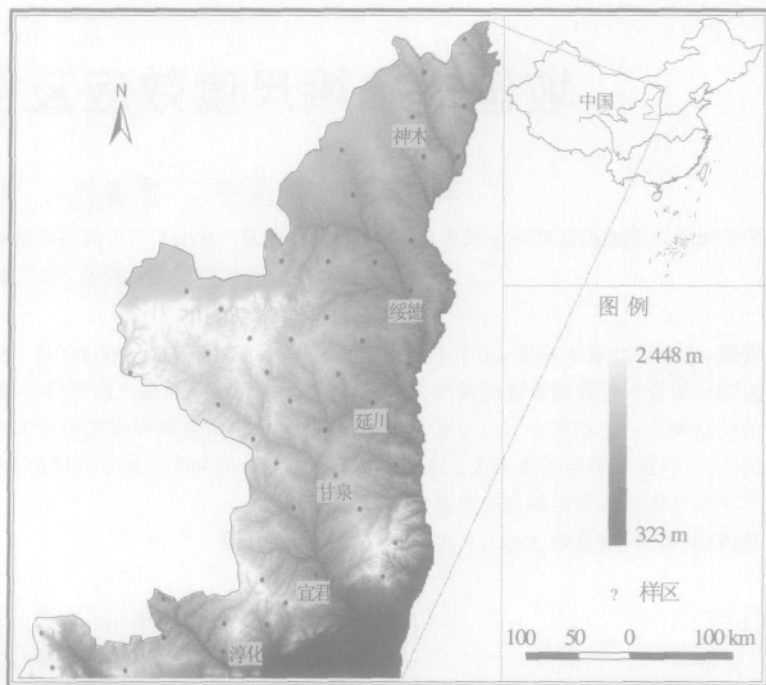


图 1 实验样区分布示意图

Fig.1 Map of the distribution of the test areas

示各样区前者的值均高于后者, 1:100 000 DEM 的坡谱信息熵值变化范围为 0.893~2.945 nat, 平均 2.609 nat, 1:500 000 DEM 的坡谱信息熵值变化范围为 0.460~2.660 nat, 平均 2.323 nat。然后, 运用 ArcGIS 地统计分析工具对陕北黄土高原坡谱信息熵的空间分异进行了预测 (图 3), 根据平均误差 (Mean Error, ME) 绝对值接近 0, 标准化平均误差 (Mean Standardized Error, MSE) 接近 0, 平均标准误差 (Average Standardized Error, ASE) 与均方根误差 (Root-Mean-Square Error, RMSE) 最接近以及标准化均方根 (Root-Mean-Square Standardized Error, RMSSE) 接近 1 的原则选择最佳半方差理论模型, 各项评价指标地统计分析过程的参数见表 1。结果显示, 1:100 000 和 1:500 000 DEM 坡谱的信息熵在空间表现相似的变异特征。坡谱信息熵最小值出现在北部风沙-丘陵过渡区, 大

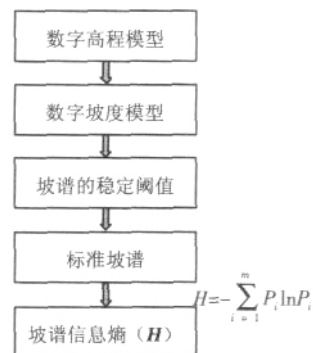


图 2 坡谱信息熵提取流程

Fig.2 Extracting procedure of SSIE

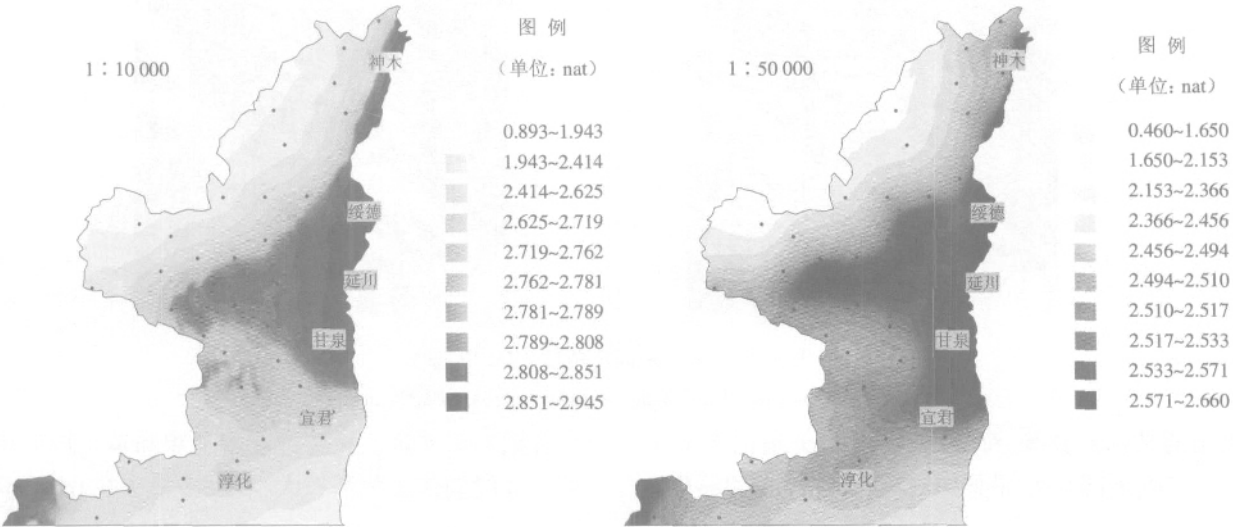


图 3 坡谱信息熵空间分异图
Fig.3 Spatial distribution of SSIE

值集中在丘陵沟壑区。对于黄龙等突起的石质山地(宜君东北方向),由于靠近黄河峡谷,侵蚀相对强烈,地面沟壑发育,切割深度较大,地形变化相对复杂,坡谱信息熵也较大。由南向北,随着地貌形态从塬、长梁残塬、梁状丘陵沟壑、梁峁状丘陵沟壑、峁梁状丘陵沟壑、峁状丘陵沟壑,再到风沙地貌区,坡谱信息熵由小变大再变小,反映了地形起伏由北向南缓和-剧烈-缓和的变化趋势,体现了地貌形态在空间的变化特征。

地形的起伏变化(图 5),在 1 10 000 DEM 晕渲图中,能清楚地看到沿坡面发育的一系列小冲沟,而 1 50 000 DEM 确因栅格分辨率增大而平滑掉了大部分冲沟的信息。实际上,是由于小比例尺地形图在生产过程中因制图综合对地物进行了一定程度的概括,舍去了大比例尺地形图包含的部分细节,带来了信息的损失。

表 1 预测精度评价
Tab.1 Evaluation of prediction accuracy

	1 : 10 000	1 : 50 000
ME	0.00576	0.01032
RMSE	0.3844	0.3594
ASE	0.2789	0.2781
MSE	0.01496	0.02522
RMSSE	1.323	1.328

3.2 典型样区分析

比较 6 个典型样区 1 10 000、1 50 000DEM 的坡谱信息熵值(图 4)可以看出总的趋势是丘陵沟壑区的坡谱信息熵要大于塬区的熵值,反映了丘陵沟壑区地形起伏较后者强烈,侵蚀相对较强,地表复杂程度较大。其次,1 10 000DEM 的坡谱信息熵高于 1 50 000DEM 的信息熵,这是由于相对于 1 50000 DEM 而言,1 10 000 DEM 能反应更细微的

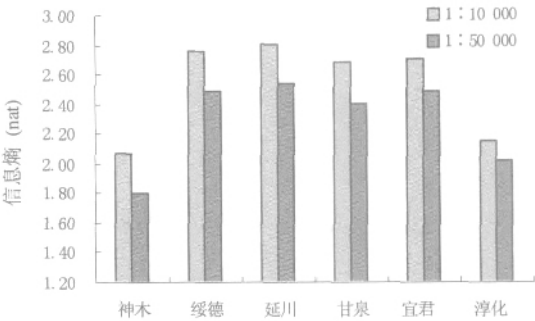


图 4 典型样区不同比例尺 DEM 坡谱信息熵对比
Fig.4 Comparison of SSIE of different scale DEMs for typical test areas

3.3 信息量变化分析

坡谱信息熵在一定程度上反映了地形起伏的强弱和地表复杂度的大小,但在实际研究中,还需要知道不同级别坡度的尺度变异特征,进而探讨不同尺度 DEM 提取坡度的不确定性及误差校正。因此对 6 个典型样区 1 10 000、1 50 000 DEM 提取的各个级别坡度的信息量进行分析。这里的坡度分级

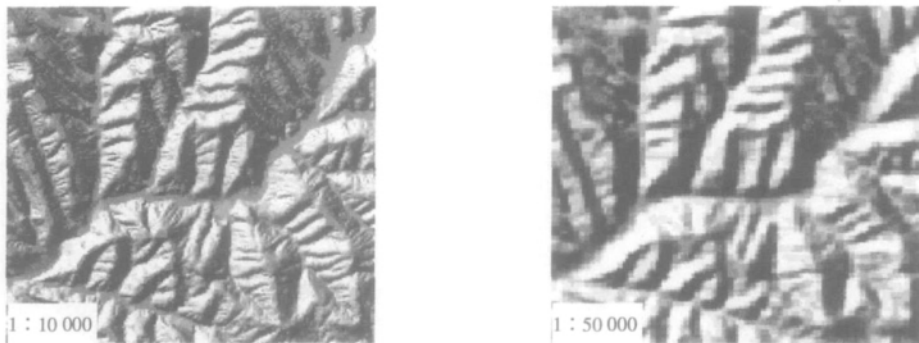


图 5 绥德样区不同比例尺 DEM 晕渲图

Fig.5 Comparison of hillshades of different scale DEMs in Suide

采用的是等差分级, 每 3 分为一级, 分别记作 1、2、3、...、30(下同)。结果如图 6 所示, 在神木和淳化样

区分别为风沙黄土丘陵过渡型地貌和黄土塬地貌, 第 1~4 级的信息量都较大, 坡度多集中在 0°~12° 范

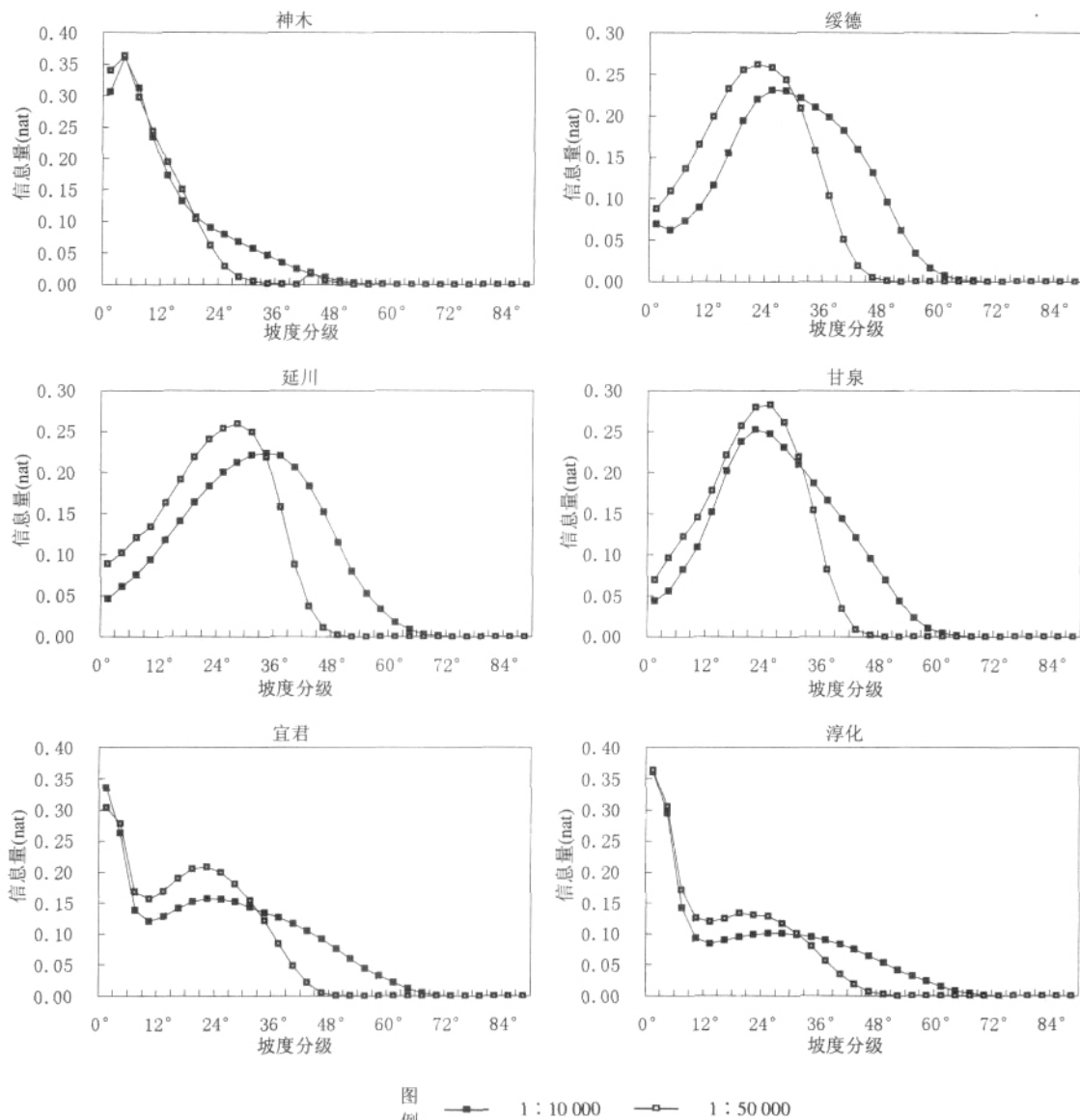


图 例 —■— 1 : 10 000 —□— 1 : 50 000

图 6 不同坡度分级信息量对比

Fig.6 Comparison of information content of different slope classes

围内, 地形起伏平缓; 绥德、延川、甘泉样区为黄土丘陵沟壑地貌, 信息量大值集中在第 5~17 级, 其对应的坡度为 15°~51°; 实际地形起伏较前者强烈; 宜君样区处于由丘陵沟壑向塬过渡的地貌类型区内, 所以, 它既具有塬区坡度组合的特征, 又具有丘陵沟壑区坡度组合的特征。另外, 几乎所有样区第 1~5 级 1 50 000 坡谱的信息量均大于 1 10 000 坡谱的信息量, 即 0°~15° 坡度所占整个样区坡度的比例

大, 表现在地形上则是小比例尺 DEM 描述的地形较大比例尺 DEM 描述的地形平缓。

从 1 10 000~1 50 000 坡谱信息损失量(h_i)的计算结果(图 7)可看出, 在 0°~25° 范围内, 几乎所有样区各个级别信息损失量都为负, 说明 1 50 000 DEM 经过制图综合和平滑, 所反映的地形起伏较 1 10 000 DEM 缓和, 体现了制图综合使坡度变小、地形起伏变缓的效果。

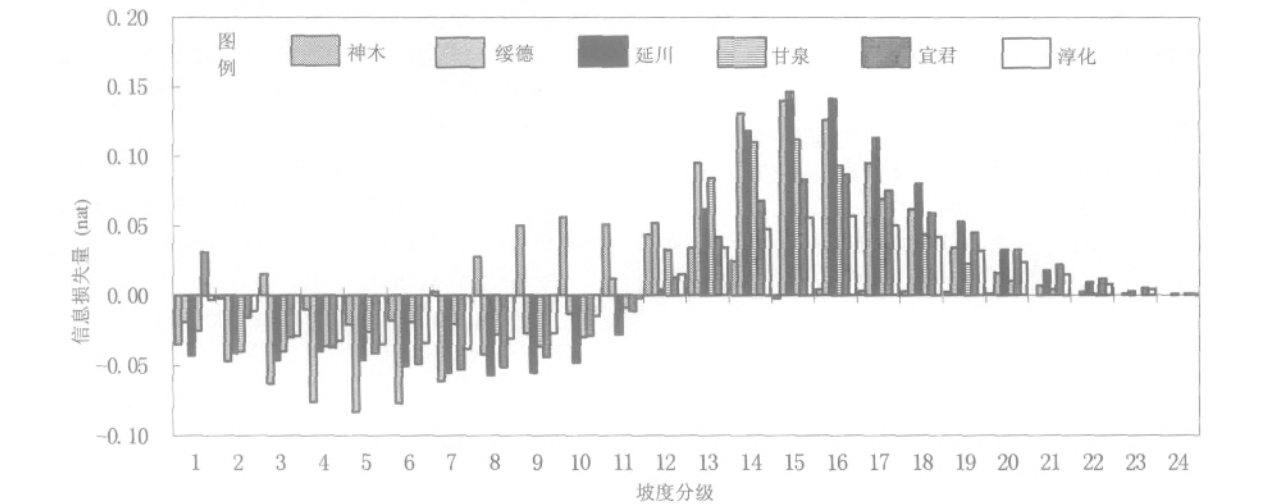


图 7 不同坡度分级信息损失量对比
Fig.7 Comparison of information loss of different slope classes

3.4 坡谱信息熵与地貌变化关系

深入研究坡谱和黄土地貌演化的规律, 通过坡谱的变化揭示黄土地貌演化的时间序列是坡谱研究重要内容之一。因此, 首先需要研究坡谱和其他可反映地貌演化特征的因子之间的关系。以沟壑密度为例, 沟壑密度是一个综合性很强的地貌指标, 它既能反映土壤侵蚀的严重程度, 又能反映地貌的演化阶段, 张丽萍、马新中等都曾对此做过深入的研究^[9]。笔者计算了 48 个样区的沟壑密度, 并将其和坡谱信息熵做了相关分析, 由图 8 可见坡谱信息熵和沟壑密度表现出较好的幂函数关系, 随着坡谱信息熵的增大, 样区沟壑密度也增大。结合前面的研究可以认为, 在陕北南部黄土塬区, 坡谱信息熵小, 沟壑密度也小, 由南向北, 随着土壤侵蚀加剧, 地形起伏变大, 到丘陵沟壑区, 坡谱信息熵逐渐增大, 沟壑密度也增大。据马乃喜的研究, 黄土高原地貌演化的后期演化方式就是指侵蚀作用使先成的黄土塬、梁、峁逐步分化解体为新的地貌类型^[10]。那么, 深入研究坡谱、沟壑密度、土壤侵蚀等的相互关系, 有望从坡谱的角度揭示地貌演化的规律。

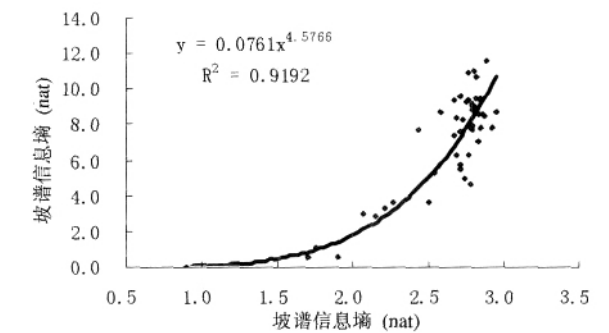


图 8 陕北黄土高原坡谱信息熵-沟壑密度图
Fig.8 Plots of SSIE vs. gully density in the Loess Plateau

4 结论

坡谱信息熵能从宏观上定量反映不同比例尺 DEM 地形信息变化情况, 小比例尺 DEM 熵值的减小, 体现了在制图综合过程中地形信息的损失。同时, 单个级别信息量的变化更精细地刻画了不同地貌类型区坡度组合及不同比例尺 DEM 相应级别坡度信息的变化规律。坡谱信息熵和沟壑密度表现出

较好的幂函数关系,随着坡谱信息熵的增大,样区沟壑密度也增大,通过这样的桥梁,可望获得坡谱和地貌演化之间的某种联系。另外,坡谱信息熵的空间分异规律与黄土高原地貌形态的空间变化规律是一致的,因此,可以将其作为地貌类型识别的因子之一。当然,能否将某一地貌类型区的坡谱信息熵界定在一个特定范围之内,是实现地貌类型识别的关键,需要进一步研究。

参考文献

- [1] C E Shannon. A mathematical theory of communication. The Bell System Technical Journal, 1948, 27: 379~423.
- [2] 艾南山,岳天祥.地貌系统的信息熵及其计算方法.熵与交叉科学,气象出版社,1988.
- [3] 岳天祥,艾南山,张英保.论流域系统稳定性的判别指标——超熵.水土保持学报,1989,3(2):20~28.
- [4] 马新中,陆中臣,金德生.流域地貌系统的侵蚀演化与耗散结构.地理学报,1993,48(4):367~376.
- [5] 陈楠,汤国安,刘咏梅等.基于不同比例尺的DEM地形信息比较.西北大学学报(自然科学版),2003,33(2):237~240.
- [6] 陈楠,林宗坚,李成名等.基于信息论的不同比例尺DEM地形信息比较分析.遥感信息,2004,(3):5~9.
- [7] 王雷.黄土高原数字高程模型的地形信息容量研究.西安:西北大学,2005.
- [8] 张茜,汤国安,王雷等.陕北黄土高原不同比例尺DEM地形信息容量转换图谱.西北大学学报(自然科学网络版),2006(已录用).
- [9] 张丽萍,马志正.流域地貌演化的不同阶段沟壑密度与切割深度关系研究.地理研究,1998,17(3):273~278.
- [10] 马乃喜.黄土地貌演化与土壤侵蚀关系的分析.水土保持通报,1998,16(2):6~10.

Scale Effect and Spatial Distribution of Slope Spectrum's Information Entropy

LI Fayuan^{1,2}, TANG Guoan¹, JIA Yini¹, CAO Zhidong³

(1 Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing 210097;

2 Institute of Mountain Hazards and Environment, CAS, Chengdu 610041, China;

3 Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

Abstract: 48 different landform areas of the Loess Plateau in northern Shaanxi province were selected as test areas with the objective of investigating spatial distribution and scale effect of the slope spectrum's information entropy (SSIE) based on DEMs. Experimental results show that the SSIE can reflect the terrain complexity to a certain extent. There exists a good power function relation between SSIE and gully density: gully density increases with the increase of SSIE. Spatial distribution of SSIE and loess landform are correlative and the SSIE can be taken as a discriminate factor to identify loess landform.

Key words: slope spectrum; information entropy; DEM; loess plateau; spatial distribution