

# 基于地形剖面线的陕北黄土高原 地貌信息熵差分析

殷宗敏<sup>1,2</sup>, 刘万青<sup>1</sup>

(1. 西北大学 城市与环境学院, 陕西 西安 710127; 2. 兰州大学 资源环境学院, 甘肃 兰州 730000)

**摘要:** [目的] 计算黄土高原地貌评价指标和剖面上高程数据信息熵, 用以拟合地形起伏, 评价地形, 为地貌研究提供新理论依据。[方法] 通过对研究区域进行剖面分析, 获取地形剖面线, 利用高程数据计算不同地貌类型中剖面线信息熵, 绘制地形剖面线图和地形剖面线信息熵图, 用来检验和评价地貌信息熵对地形的表达效果, 统计两者离散程度指标, 通过线性拟合找出两者相关性指标。[结果] 将两图进行对比分析, 发现两图走势具有一定的相似性, 采用统计指标作为指示器, 结果表明高程曲线与信息熵曲线具有较强的相关性, 采用侵蚀模数验证, 表明信息熵在表示地貌起伏或侵蚀方面具有可靠性。[结论] 地貌信息熵可以用来模拟地形, 且模拟效果良好, 可以用于评价地貌类型。

**关键词:** 黄土高原; 地形剖面线; 信息熵; 对比分析

文献标识码: B

文章编号: 1000-288X(2017)01-0166-07

中图分类号: P931.6

**文献参数:** 殷宗敏, 刘万青. 基于地形剖面线的陕北黄土高原地貌信息熵差分析[J]. 水土保持通报, 2017, 37(1): 166-172. DOI: 10.13961/j.cnki.stbctb.2017.01.030; Yin Zongmin, Liu Wanqing. Information entropy difference analysis based on terrain profile line of Loess Plateau in Northern Shaanxi Province[J]. Bulletin of Soil and Water Conservation, 2017, 37(1): 166-172. DOI: 10.13961/j.cnki.stbctb.2017.01.030

## Information Entropy Difference Analysis Based on Terrain Profile Line of Loess Plateau in Northern Shaanxi Province

YIN Zongmin<sup>1,2</sup>, LIU Wanqing<sup>1</sup>

(1. College of Urban and Environmental Sciences, Northwest University, Xi'an, Shaanxi 710127, China;  
2. College of Earth Environmental Sciences, Lanzhou University, Lanzhou, Gansu 730000, China)

**Abstract:** [Objective] This paper studies the geomorphic types from the perspective of information entropy and adds new index for landscape evaluation in Loess Plateau. The terrain can be fitted and evaluated by calculating the information entropy on the elevation data, which can provide a new theoretical basis for studying geomorphology. [Methods] Through analyzing terrain profile of the study area, we extracted the terrain profile line, then used elevation data to calculate the section line information entropy and elevation for different physiognomy types. We rendered terrain profile section line map and terrain information entropy diagram, and used them to test and evaluate the effect of geomorphic information entropy in expressing terrain information, and to calculate the index of dispersion degree, and to find the correlation between them through linear fitting. [Results] With a comparison of two charts, the two figures showed similar trends. The statistical indicators that there was a strong correlation between the elevation curve and the information entropy curve. The validation by erosion modulus showed that the information entropy was reliable in representing the topographic changes or erosion. [Conclusion] Terrain information entropy can be used to simulate the terrain, and the simulation results are favorable, it can be used to evaluate the landscape types.

收稿日期: 2016-06-03

修回日期: 2016-06-27

资助项目: 国家自然科学基金项目“地表坡度的统计分布理论模型研究”(41371274)

第一作者: 殷宗敏(1992—), 男(汉族), 江西省九江市人, 硕士研究生, 主要研究方向为地貌信息熵及分形分维。E-mail: 2267662868@qq.com。

通讯作者: 刘万青(1965—), 男(汉族), 陕西省渭南市人, 硕士, 副教授, 主要从事土壤侵蚀与水土保持等方面的研究。E-mail: liuwqing@nwu.edu.cn。

**Keywords:** Loess Plateau; terrain profiles; information entropy; comparison analysis

地形起伏作为地理学研究的一个方向,对于区域规划、生态修复、水土保持等具有重要价值。黄土高原作为一个独特的区域,长期受到专家学者的重视,对于其地形起伏特征的研究已有多种方法<sup>[1]</sup>。主要从遥感影像表面纹理或者单纯、离散的高程数据来进行分析;如果地形起伏研究从多组内部平行剖面线开始,将提供一种崭新的研究思路。

香农的信息熵理论,是用于度量信息量的概念,用来评价信息量的多少。一个系统越是有序,信息熵就越低;反之,一个系统越是混乱,信息熵就越高<sup>[2]</sup>。对于黄土地貌而言,若侵蚀状态轻微,则表现为表面较平坦,无明显沟壑,为一种有序的状态,故计算所得信息熵较小;反之,若侵蚀严重,则表面崎岖,沟壑纵横,表现为一种无序的状态,故计算所得信息熵较大。由此可见,信息熵在评价地貌类型方面具有价值。若区域面内侵蚀严重,地面起伏则明显,即高程变化剧烈,因此信息熵与高程变化有关;对于评价高程变化,可引入剖面起伏度的概念。定义为剖面线上波峰(波谷)的高程和位于其两侧的波谷(波峰)高程均值之差<sup>[3]</sup>。

对于剖面线而言,其有序的高程数据按照剖面线走向依次排列,根据地理学第一定律,相邻事物间具有强相关性,可知:有序的高程信息蕴含一定信息量,而信息量可以通过信息熵来度量。因此,对地形剖面线上高程数据计算信息熵理应反映地形起伏变化,即地形起伏。通过对剖面上高程数据计算信息熵,能够拟合地形起伏,评价地形<sup>[4]</sup>。因此本文从该角度出发,做一些有意义的尝试。

对于地形特征的评价指标,现有的评价参数如下。① 坡面地形因子:坡面姿态因子、坡形因子、坡长因子、坡位因子、坡面复杂度因子。② 特征地形要素:地形特征点、山脊线、山谷线、沟沿线、水系、流域<sup>[5]</sup>。③ DEM 地形统计分析:相关分析、回归分析、趋势面分析。④ DEM 地学模型分析:以 DEM 为对象构建模型<sup>[6]</sup>。本文创造性的引入了信息熵的概念,对于信息熵的研究,有 3 种常用方法:① 概率统计方法,直接来自香农公式,忽略了地物间的强相关性;② 组合的方法,这是建立在地物表面构成元素多元的基础上的方法,不具有具体计算的功能<sup>[7-8]</sup>。③ 综合特征值量测法,公式为:

$$i = \lg(1 + \omega) \quad (1)$$

式中: $\omega$ ——制图特征<sup>[7]</sup>;  $i$ ——信息熵(hart,以 10 为底)。

本文拟以该公式为基础,对其进行适当修改后,

用于计算地形剖面线的信息熵,并借此反映地形起伏状态<sup>[8]</sup>。

## 1 数据与方法

### 1.1 数据来源及预处理

通过数据共享平台(STRM 下载地址:<http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp>),获得黄土高原地区的数字高程模型,数据格式为 GRID,坐标系为 WGS\_1984,高程最小值为 93 m,最大值为 4 981 m,平均高程为 1 408 m。矢量化陕西省内黄土高原地区的不同地貌类型分区图。鉴于地貌分区图图斑数量较多,且大小不一,如果统一采样间距时兼顾各图斑,将明显限制采样间距长度,影响结果表达;因此采用取舍的思路,选取面积较大、长宽比接近 1 的区域,符合该条件的区域有 10 个,分别为:残塬沟壑、高塬沟壑、黄龙山低山丘陵沟壑、宽梁残塬沟壑、梁峁丘陵沟壑、梁状丘陵沟壑、崩梁丘陵沟壑、崩状丘陵、破碎塬沟壑、沙丘草滩沙地。对这 10 类区域的 DEM 数据做高程 3 D 分析,绘制剖面图,剖面线采用平行线采样法绘制,选取合适位置的 6 组相同长度的剖面线所得高程的平均值作为最终采样线,获得该采样线的高程数据<sup>[9]</sup>。

### 1.2 分析方法

信息熵是一个数学上颇为抽象的概念,理解为某种特定信息的出现概率<sup>[10]</sup>。将信息熵应用于地理学,其蕴含了地物所具有的信息,反映为地物的评价指标。由于地物在空间格局上表现为极强的相关性,该相关性反映了制图特征;所以在应用公式(1)时, $\omega$ 应反映出地物间的相关性,文中地物表现为高程数据,因此  $\omega$  中应蕴含相邻采样点间高程的对应关系,故引入剖面起伏度的概念,用  $A$  表示,其为剖面线上波峰(波谷)的高程和位于其两侧的波谷(波峰)高程均值之差<sup>[2]</sup>。剖面线表现为线状要素,其信息量取决于构成线要素的节点在表达要素整体形态上的重要性程度。若重要性程度越高,则信息量越大<sup>[11]</sup>。如何评价重要性程度的高低,根据邓敏等<sup>[11]</sup>的研究成果,可以对节点拟定一个半径绘制圆,与节点相邻两边有交点,将节点到两交点连线的垂距除以两交点间距离,所得比值为重要性程度,比值越大则重要性程度越高。

该重要性程度值能够反映出制图特征,因此用重要性程度值充当制图特征值  $\omega$ 。对于剖面线,由于拟定合适的半径值在大量数据处理时不方便,因此对其进行改进,将剖面起伏度的绝对值  $|A|$  除以 2 个采样

间距(2d)所得的值替代重要性程度,即  $w$  取值为  $\frac{|A|}{2d}$ 。因此,信息熵  $i$  的计算公式为:

$$i = \lg(1 + \frac{|A|}{2d}) \quad (2)$$

式中: $A$ ——起伏度; $d$ ——采样间距(m); $i$ ——地貌信息熵(Hart)。

信息熵反映的是地形的变化程度,不具有区分正负地形的作用。广义的正地形指的黄土高原区域相对于周边地区明显高出的区域,负地形指的是相对凹陷的区域<sup>[12]</sup>。信息熵计算的结果始终为正值,对于正地形部位或者负地形部位变化程度相同的地貌类型,所得的信息熵值可能相等,为了更好区分这两类地貌,引入系数  $k$ :

$$k = \frac{\bar{h} - h}{|\bar{h} - h|} \quad (\bar{h} \neq h) \quad (3)$$

式中: $h$ ——高程(m); $\bar{h}$ ——平均高程(m); $k$ ——正负地形判别系数(当  $h = \bar{h}$ , 定义  $k=1$ )。

将预处理得到的陕北不同地貌类型的 DEM 数据与陕北黄土地貌分区图叠加显示。使陕北黄土地貌分区图在 DEM 数据的上层,并在该层上绘制剖面线<sup>[9]</sup>,剖面线的绘制方法采用创建渔网方式,对渔网绘制赋予高程信息;同时利用线裁剪方式获得 6 组平行线,顾及到地貌走向,选取的平行线位置大致居中分布,平行线线长 36 km,线间间距为 2 km,利用平行线,绘制剖面图;出于对比考虑,选取横纵走向 2 种剖面线。利用 ArcGIS 软件,获得高程数据,通过 Excel 的函数功能,利用上文的计算公式计算剖面线的信息熵。

利用获得的数据,分别绘制地形剖面线图和地形剖面线信息熵图,并且将两图进行对比分析,获得结论,并对分析结果进行可靠性验证。

至于区域图斑中采样间距的确定,综合考虑 2 个因素:一是研究区的空间尺度大小(一般尺度越大,则间距越大;反之亦然)。二是地貌单元的空间规模(规模越大,则间距越大;反之亦然)<sup>[13]</sup>。最终确定采样线采样长度为 36 km,平行采样线间间距为 2 km,采用该尺度,能够使得在指定区域图斑中横向、纵向都能采样,保证了数据的完整性,可以研究地形走向对计算所得信息熵的影响。图 1 为分析方法的技术流程。

## 2 结果与分析

### 2.1 试验结果

对陕北黄土地区的 10 个样区,选择横纵走向的剖面线,进行相应的处理,计算信息熵,由于信息熵计

算结果为较小的小数,因此对结果乘以系数 10 000,扩大其表现效果。利用获得的地形剖面线高程数据,绘制地形剖面线图;利用获得的信息熵数据,绘制出地形剖面线信息熵图。计算各研究地貌的高程、信息熵标准差,并将两者绘制于同一表中(表 1),方便进行对比分析。

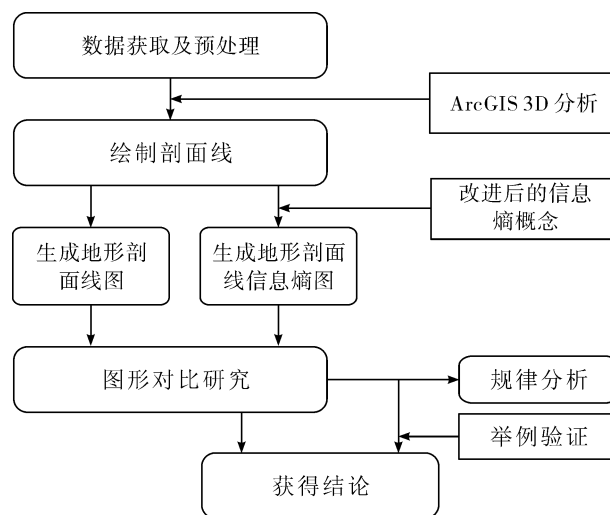


图 1 分析方法技术流程

表 1 陕北黄土地区信息统计

地貌类型	信息标准差			
	横向采样		纵向采样	
	高程/ m	信息熵/ hart	高程/ m	信息熵/ hart
沙丘草滩沙地	26.380	6.077	42.920	8.400
崩状丘陵	32.232	50.934	24.434	50.645
崩梁丘陵沟壑	63.612	70.664	32.042	63.482
残塬沟壑	41.382	52.717	43.860	54.762
梁状丘陵沟壑	41.568	72.588	28.537	54.169
梁崩丘陵沟壑	55.716	63.842	43.346	65.562
宽梁残塬沟壑	63.032	55.736	80.061	78.097
破碎塬沟壑	54.485	60.033	71.583	77.859
高塬沟壑	84.203	68.135	49.429	54.950
黄龙山低山丘陵沟壑	160.839	67.236	121.903	97.531

由于原始采样数据中采样点较多,在 Excel 中用图表绘制时,剖面线起伏过于剧烈,不适合用于规律分析。综合区域内数据特性,选取一定间距采样点数作为移动平均单位,使得图表中波形平滑易于分析,对比发现,采取 10 个采样点数作为移动平均的单位,能够达到不错的效果。

标准差是总体各单元标准值与其平均数离差平方的算数平均数的平方根。可以反映组内个体间的离散程度。分别统计,高程和信息熵数据中标准差,

观察表 1 中高程和信息熵的标准差,发现高程和信息熵具有对应关系,大致表现为,高程的标准差越大,对应信息熵的标准差也增大,反之亦然;因此推断,高程和信息熵存在明显相关,找出评价指标。在 Excel 中以高程为横轴,信息熵为纵轴,绘制散点图,并采用多项式拟合绘制趋势线,多项式系数选取 6,并显示  $R^2$ ,结果见表 2。

表 2 高程曲线与信息熵曲线的  $R^2$  值

地貌类型	不同采样方式 $R^2$ 值	
	横向	纵向
沙丘草滩沙地	0.891 5	0.888 1
崩状丘陵	0.916 0	0.886 8
崩梁丘陵沟壑	0.893 9	0.873 4
残塬沟壑	0.892 3	0.832 8
梁状丘陵沟壑	0.927 9	0.920 9
梁崩丘陵沟壑	0.910 8	0.901 4
宽梁残塬沟壑	0.872 5	0.871 4
破碎塬沟壑	0.905 5	0.895 1
高塬沟壑	0.911 2	0.872 2
黄龙山低山丘陵沟壑	0.871 2	0.916 6

对比研究表 2 中的  $R^2$  值,可得出如下分析结果:

- ① 无论横向、纵向采样,高程曲线与信息熵曲线的  $R^2$  值均大于 0.8,说明高程与信息熵具有强相关性;
- ② 对比表 2 中同种地貌类型的  $R^2$  值,值的大小说明信息熵曲线与高程曲线的拟合程度,拟合程度越高,  $R^2$  值越大,  $R^2$  值大的采样方向为合理的采样方向,即地形走向与该采样方向一致;
- ③ 除黄龙山低山丘陵沟壑应采取纵向采样外,其余地貌采样方式横向更为合适。

在分析过程中也发现了几个值得深入讨论的问题: ① 绘制剖面线时,不同区域内采样线长度相等,然而在不同区域中,采样线的长度占区域的比例不同,其是否影响剖面线对地貌的表达? ② 表 2 中  $R^2$  值太大,是否存在偏高的现象,偏高的原因需要探究; ③ 高程和信息熵的标准差的对应存在突变情况,即高程的标准差大时,信息熵的标准差反而小,其原因是什么?

## 2.2 结果分析

对比地形剖面线图 2—3,可得出如下分析结果:

(1) 地形剖面线波形与地形剖面线信息熵波形的走势具有相似性,波形波峰和波谷有对应关系。

(2) 单点高程高时,信息熵可能小,如:残塬沟壑和高塬沟壑;当然信息熵也可能大,如:高塬沟壑和梁状丘陵沟壑。从而反映出信息熵与单点高程值无直

接的对应关系,然而从图 2—3 可以看出信息熵与起伏度有关。

(3) 同种地貌类型其横纵剖面线形态存在差异,同一区域横纵剖面线上平均高程大致相同。

(4) 横向采样的标准差和平均绝对偏差大于纵向采样时,说明横向起伏程度大于纵向起伏程度。

(5) 高程的标准差越大,对应的信息熵的标准差也越大。

## 2.3 结果检验

为了验证信息熵,在表示地貌起伏或侵蚀方面的可靠性,可选择侵蚀模数来验证,查询资料得知,黄土高原沟壑区侵蚀模数(1955—2007 年平均值)为 5 076.7 t/(km<sup>2</sup>·a),表示强度侵蚀;残塬沟壑区侵蚀模数(1955—2007 年平均值)为 8 769.4 t/(km<sup>2</sup>·a),表示强烈侵蚀<sup>[13]</sup>。观察试验中残塬沟壑和高塬沟壑的绝对值平均信息熵,发现残塬沟壑的绝对值平均信息熵大于高塬沟壑的绝对值平均信息熵,运用引言中的理论,信息熵越大,表明地形起伏越明显,土壤侵蚀越强烈,该试验中的信息熵与资料中的侵蚀模数相贴合,说明该参数在判别地貌类型上具有一定价值。

## 3 讨论与结论

### 3.1 讨论

文中在采样区域内选择线性剖面数据来计算信息熵,如果利用长条带状的方式进行数据采样,计算每个小方块中的平均高程数据,代替剖面线中的点状数据,更能说明区域特征<sup>[14]</sup>。

对于一种地貌类型,其地形剖面线中高程越大,地形剖面线的信息熵反而越小,可能的解释是,对于正地形而言,更容易发生重力侵蚀,表现为整体性平移,如滑坡,对于负地形而言,更容易发生水力侵蚀,如沟蚀,使得大地被切割的支离破碎<sup>[15]</sup>。重力导致的侵蚀,造成的黄土地区破碎程度不如水力侵蚀,故信息熵小。

### 3.2 结论

由于地形复杂,如果单纯从高程出发,将掩盖地形的其他特征指标,而信息熵的运用,弱化了高程的影响,使得地形表达趋于全面。根据图 2—3 和表 2,并结合结果检验中的检验数据,得出信息熵指标在模拟和评价地形方面具有可靠性,与参考文献[16]一致,大多数情况下两折线的峰值、谷值分别对应,且计算的两者的  $R^2$  数大于 0.8,说明高程与信息熵具有强相关性,所以得出结论,信息熵可以用来模拟地形,且模拟效果良好。

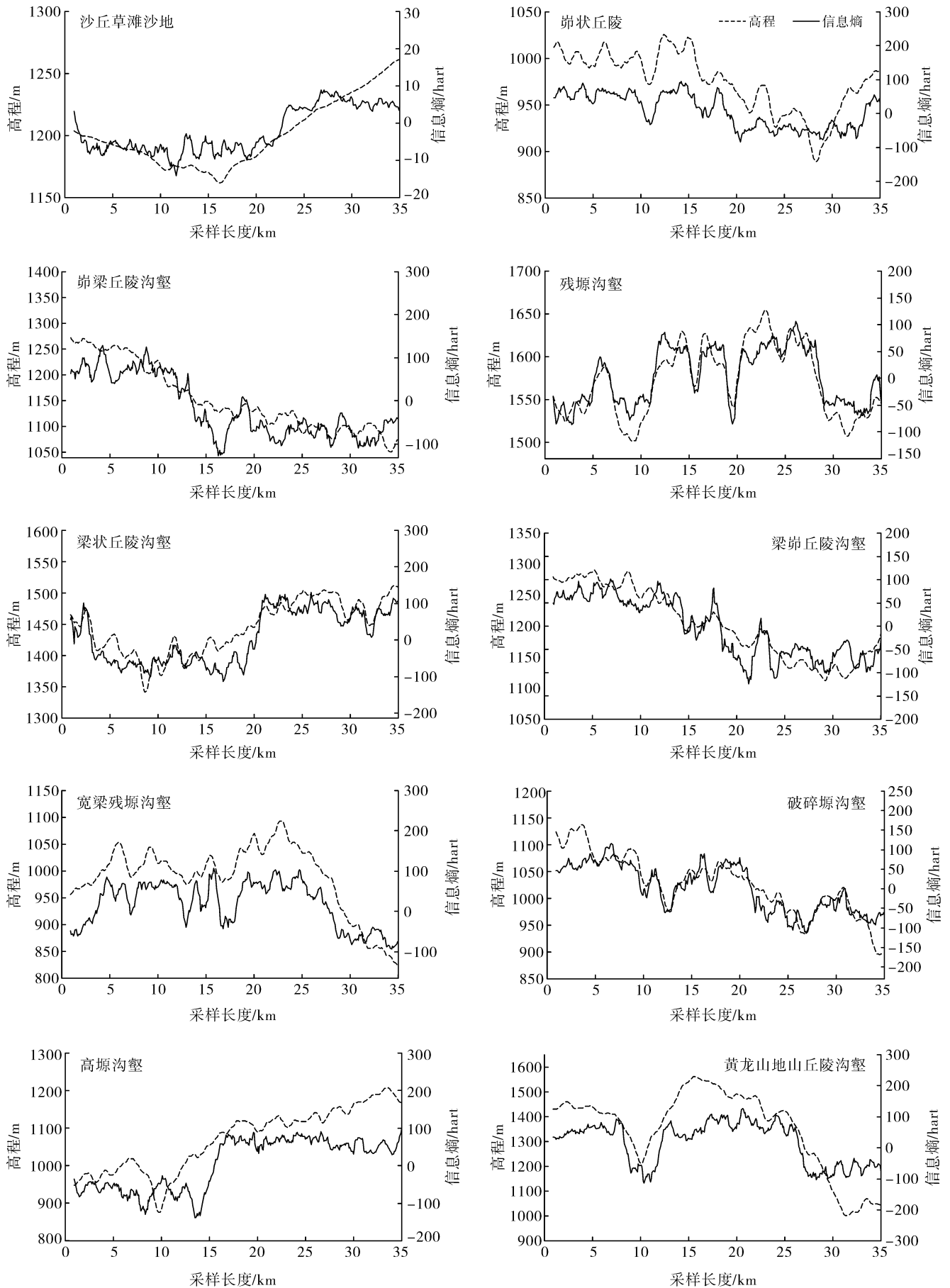


图 2 横向剖面线高程和信息熵对比

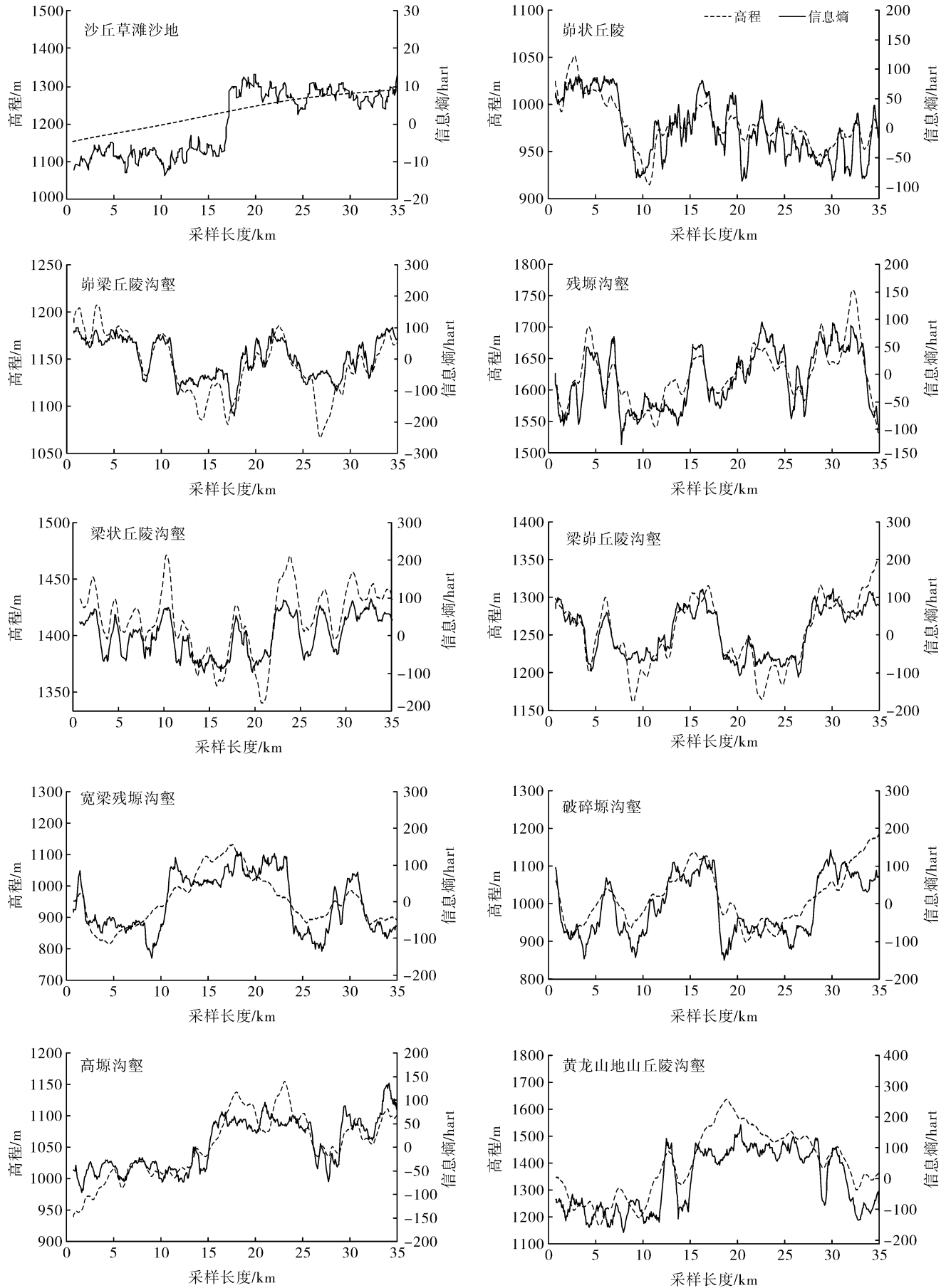


图 3 纵向剖面线高程和信息熵对比

致谢: 本文在写作过程中, 使用的数据包括 STRM 数据平台的 DEM 数据, 以及甘枝茂先生的黄土高原地区地貌分区图。在此一并表示衷心感谢!

#### [ 参 考 文 献 ]

- [1] 李炳元, 潘保田, 程维明, 等. 中国地貌区划新论[J]. 地理学报, 2013, 68(3): 291-306.
- [2] Shannon C E. A mathematical theory of communication [J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(1): 3-55.
- [3] 张维. 基于 DEM 的陕北黄土高原流域剖面谱研究[D]. 南京: 南京师范大学, 2011.
- [4] 张会平, 刘少峰. 利用 DEM 进行地形高程剖面分析的新方法[J]. 地学前缘, 2004, 11(3): 226-226.
- [5] 郭兰勤, 丑述仁. 基于 DEM 不同路径算法的沟壑密度提取[J]. 地下水, 2011(6): 145-147.
- [6] 张磊, 汤国安, 李发源, 等. 黄土地貌沟沿线研究综述[J]. 地理与地理信息科学, 2012, 28(6): 44-48.
- [7] 张法. 地图信息量的量测和评价[J]. 测绘, 2006, 29(1): 21-24.
- [8] Zhilin Li, Peizhi Huang. Quantitative measures for spatial information of maps[J]. International Journal of Geographical Information Science, 2002, 16(7): 699-709.
- [9] 汤国安. ArcGIS 地理信息系统空间分析实验教程[M]. 北京: 科学出版社, 2012.
- [10] Peters S. Interactive scale-dependent multidimensional point data selection using enhanced polarization transformation[J]. Journal of Dental Research, 2011, 81(4): 253-258.
- [11] 邓敏, 樊子德, 刘慧敏. 层次信息量的线要素化简算法评价研究[J]. 测绘学报, 2013, 42(5): 767-773.
- [12] 罗来兴. 划分晋西、陕北、陇东黄土区域沟间地与沟谷的地貌类型[J]. 地理学报, 1956, 22(03): 201-222.
- [13] 王万忠, 焦菊英, 马丽梅, 等. 黄土高原不同侵蚀类型区侵蚀产沙强度变化及其治理目标[J]. 水土保持通报, 2012, 32(5): 1-7.
- [14] 邹斌文, 马维峰, 龙昱, 等. 基于 ArcGIS 的条带剖面提取方法在地貌分析中的应用[J]. 地理与地理信息科学, 2011, 27(3): 42-44.
- [15] 张翼. 黄土高原丘陵沟壑区土壤侵蚀研究[J]. 水土保持研究, 2000, 7(2): 121-124.
- [16] 管华. 秦岭—黄淮平原交界带地貌信息熵分析[J]. 河南大学学报: 自然科学版, 1996, 26(4): 67-71.
- [17] 徐彬彬. 我国土壤光谱线之研究[J]. 环境遥感, 1991, 6(1): 61-71.
- [18] 毕海芸, 王思远, 曾江源, 等. 基于 TM 影像的几种常用水体提取方法的比较和分析[J]. 遥感信息, 2012, 27(5): 77-82.
- [19] 程磊, 徐宗学, 左德鹏, 等. 基于 Landsat TM 数据的黄土高原区水体识别方法研究[J]. 北京师范大学学报: 自然科学版, 2010, 46(3): 424-430.
- [20] 王伟武, 朱霞, 孙跃池, 等. 基于 ETM 图像的山地水体提取方法研究[J]. 系统仿真学报, 2013, 25(9): 2196-2200.
- [21] Mcfeeters S K. The use of the normalized difference water index (NDWI) in the delineation of open water features[J]. International Journal of Remote Sensing, 1996, 17(7): 1425-1432.
- [22] 田庆久, 闵祥军. 植被指数研究进展[J]. 地球科学进展, 1998, 13(4): 327-333.
- [23] Qi J, Chehbouni A, Huete A R, et al. A modified soil adjusted vegetation index[J]. Remote Sensing of Environment, 1994, 8(2): 119-126.
- [24] 池宏康. 黄土高原地区提取植被信息方法的研究[J]. 植物学报, 1996, 38(1): 40-44.

(上接第 165 页)

- [9] 詹志明, 秦其明, 阿布都瓦斯提·吾拉木, 等. 基于 NIR-Red 光谱特征空间的土壤水分监测新方法[J]. 中国科学: D 辑, 2006, 36(11): 1020-1026.
- [10] 刘焕军, 宇万太, 张新乐, 等. 中国东北黑土带土壤线空间变异规律[J]. 农业工程学报, 2009, 25(10): 166-170.
- [11] Fox G A, Sabbagh G J, Searcy S W. An automated soil line identification routine for remotely sensed images[J]. Soil Science Society of America Journal, 2004, 68(4): 1326-1331.
- [12] 秦其明, 游林, 赵越, 等. 基于二维光谱特征空间的土壤线自动提取算法[J]. 农业工程学报, 2012, 28(3): 167-171.
- [13] XU Dandan, GUO Xulin. A Study of Soil Line Simulation from Landsat Images in mixed grassland[J]. Remote Sensing, 2013, 5(9): 4533-4550.
- [14] 杨国范, 翟光耀, 张婷婷, 等. 基于环境小卫星的土壤线提取研究[J]. 遥感技术与应用, 2014, 29(3): 428-432.
- [15] 朱乐天, 焦峰, 刘源鑫, 等. 黄土丘陵区不同土地利用类型土壤水分时空变异特征[J]. 水土保持研究, 2011, 18(6): 115-118.
- [16] 赵婷. 陕北延河流域结构性植被盖度遥感模型研究[D]. 西安: 西北大学, 2013.