

文章编号: 1000-0585(2001)06-0731-08

基于因子分析的中国西部 土地利用程度分区

王秀红¹, 何书金¹, 张镜铨¹, 罗 明²

(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 北京大学城市与环境学系, 北京 100871)

摘要: 利用多元统计方法中的因子分析法, 对我国西部地区县市级行政区的土地利用程度进行了综合评价。基于评价单元的含有结构性映射的综合得分值, 以中国生态地理区域系统为宏观控制性框架, 利用 GIS 技术对评价单元进行了归并, 从而尝试了自上而下和自下而上区划方法的结合。研究结果将西部 12 个省、自治区和直辖市的所有县市级行政区归并成东部季风区、西北干旱区和青藏高寒区 3 个大区及 30 个小区。

关键词: 因子分析; 西部地区; 土地利用程度; 分区

中图分类号: F301.24; O213 **文献标识码:** A

1 引言

全国的土地整理工作方兴未艾, 西部大开发战略的实施又为此特殊地区的土地整理提出许多新的命题。西部土地整理所服务的国家目标是生态和资源安全^[1]; 因此, 土地整理开发过程中必须保证生态环境的改善, 同时不断提高土地的利用率和利用效益。显然, 就全国乃至全球而言, 生态环境脆弱的中国西部地区, 其土地整理必须符合当地特殊的生态地理条件。本分区的目的就是服务于西部土地整理模式的深入研究, 为在宏观上进一步发现土地利用中存在的问题和探索土地整理的方向提供科学依据。

在我国, 关于土地利用和土地生产力的区划与分区研究已经取得了丰硕的成果^[2~5], 特别是利用综合数值方法和 GIS 技术研究土地利用问题的最新进展^[6~9], 为本研究提供了许多重要启示。人为因素对土地利用具有很大影响, 同时由于人为因素的强烈影响, 相邻区域之间土地利用结构和利用程度的空间相互作用比较强烈。但所有这些人为因素都不可能大范围、长时间地离开生态地理环境的制约, 特别是在西部地区, 生态地理条件的制约作用更为重要。中国科学院地理研究所长期研究的中国综合自然区划^[10]和中国生态地理区域系统^[11, 12], 是许多野外考察、科学观测和分析、理论性宏观综合的结果; 其区域系统涵概了许多重要生态地理界线的有机组合。将这些重要生态地理界线作为本分区研究的宏观控制性框架, 是保证研究目标和结果不偏离总体方向的必要条件。

收稿日期: 2001-09-10; 修订日期: 2001-10-22

基金项目: 国土资源部重点项目 (2000301); 中国科学院地理科学与资源研究所知识创新工程领域前沿项目 (CXIOG-E 01-01); 国土资源部 2000 年科技专项计划 (200010303) 资助

作者简介: 王秀红 (1964-), 男, 山西昔阳人, 博士, 副研究员, 从事土壤、地生态学与土地资源研究, 发表论著 30 余篇 (部)。

有关土地利用程度的信息纷繁复杂,但分区定界最需要的是可以直接参考的、集中主要信息的综合性指数。在一定分区原则的指导下,利用主成分分析和聚类分析等多元统计方法对评价单元进行归并的研究已经出现^[14,15],这为土地利用的定量化研究做出了贡献。本文所利用的因子分析法,可对相关性或重叠性较高的数据资料进行有效的“提纯”和合理综合,在信息处理方面显示出许多的优势。这样,充分利用新近土地详查的成果,依托成熟的多元统计方法和 GIS 技术,可以实现自上而下和自下而上区划方法的结合。本研究结果将西部黔、滇、川、渝、桂、青、藏、陕、甘、宁、新、内蒙古 12 个省、自治区和直辖市的所有县市级行政区划单元归并成东部季风区、西北干旱区和青藏高原区 3 个大区及 30 个小区。

2 资料来源与研究方法

本分区采用的指导性框架是新近完成的“中国生态地理区域系统”^[11~13],特别注意参照其提供的水热格局在宏观上以及地形特征。在这种水热格局影响下的土地利用格局在宏观上是比较稳定的,而通过生物技术和生物环境技术而影响的土地利用状况则是局部的或不稳定的。只有在此总体宏观格局控制下,自下而上的数值归并方法才不至于迷失方向。

土地整理是操作性极强的工作,并且要求行政区界的完整性。显然,省级区界过于笼统而乡镇级区界过于琐碎。尽管在研究区内,有的县区面积过大,跨越了部分生态地理区域系统中描述的水热格局控制线,但对西部总体而言,采用县市级行政区作为综合评价和归并的基本单元是比较适用的方法。有关土地利用的资料是新近土地详查的成果(1996 年)。反映土地利用程度指标的选择,主要依据新近完成的研究成果“中国土地资源”^[14],包括土地利用率(x_1)、农业用地指数(x_2)、垦殖率(x_3)、森林覆盖率(x_4)、林地利用率(x_5)和牧草地比重(x_6)。这些指标较为常用,也密切相关于当地综合生态地理条件;而复种指数和建设用地指数由于受技术和经济条件的直接影响,故暂未采用。

用数学方法进行综合评价,是建立一种从高维空间到低维空间的映射,这种映射能保持样本在高维空间的某种“结构”,其中最明显的是与“排序”有关的结构。根据图形资料和数据资料的匹配情况,本文利用因子分析方法,对 974 个样本(县市级行政区)有关土地利用程度的信息进行“提纯”(冗余信息的剔除),之后进行有效信息的综合。最后,每个样本均被赋予一个反映相应土地利用程度的、含有土地利用结构性特征的综合评价价值。这种结构性特征只有在各个样本相互比较时才存在,否则只是单纯的数值。

具体步骤:(1)利用主成分分析法提取公共因子,(2)根据方差累计贡献率提取特征根(一般认为应大于 70%,即可集中主要信息),(3)选择方差最大正交旋转进行变换,(4)依据回归法计算因子得分,(5)最后根据特征根的贡献率大小进行加权平均计算综合得分^[16,17]。综合得分值的公式为:

$$V = {}_1f_1 + {}_2f_2 + \dots + {}_nf_n$$

式中 V 为评价单元综合得分值, f_1 、 f_2 ... f_n 为前 n 个公因子, ${}_1$ 、 ${}_2$... ${}_n$ 为前 n 个特征根。

在 GIS 的技术支持下,首先将因子分析及综合评价的结果与各样本的空间属性进行

关联，然后对各个样本或评价单元进行归并（图 1）。因为土地利用程度在现实空间上是逐渐过渡的，所以根据样本在二维空间上的实际分布，参照综合得分值进行评价单元的归并，可以消除某些“结构”差异很大而综合得分值相近的单元的合并。主要有三种情形需要处理：（1）如果水热格局或自然区控制线穿过某评价单元，单元的归并注意相邻评价单元的综合得分值的比较；（2）如果评价单元面积过大，包含几个自然区，为了照顾县级区界的完整性，则对自然区进行适当的归并，如塔里木与吐鲁番盆地及昆仑山北翼；（3）如果自然区内各评价单元的综合得分值分异较大，而且自然区的面积较大，则对自然区进行必要的划分，如在塔里木与吐鲁番盆地区又划分出天山南坡。后两种情形极少。

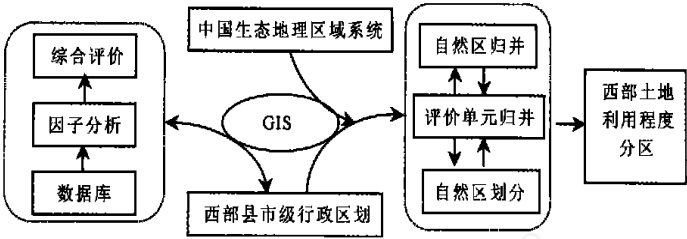


图 1 研究方法技术流程图

Fig. 1 Technological process of research method

3 结果与讨论

3.1 综合得分值的基本格局

因子分析将常用的土地利用率、农业用地指数、垦殖率、森林覆盖率、林地利用率和牧草地比重等指数中相关或重叠的信息，进行必要的剔除，也对上述 6 个变量进行了归并。由表 1 所示，前三个特征根的累计贡献率已经达到 90.962 %（一般认为应大于 70 %，即可提取主要信息），即 6 个变量反映的信息可由三个主成分（ $x_1 + x_2 + x_3$ 5.5 个变量）反映 90.962 %；因此，在 6 个特征根中提取前三个特征根。

由表 2 可知，第一因子在 x_1 、 x_2 的载荷较高，这是从土地利用率和农业用地指数两方面反映土地利用的总体程度，可称为总体利用因子；第二因子在 x_4 、 x_5 、 x_6 的载荷较高，主要从森林覆盖率、林地利用率和牧草地比重方面反映林草地的覆被和利用状况，可称为林草地利用因子；第三因子在 x_3 的载荷较高，主要从垦殖率方面反映耕地利用程度，可称为耕地利用因子。这三个因子的物理意义可用上述方

表 1 因子分析总方差解释

Tab. 1 Total variance explained

因子	提取值			旋转提取值		
	特征根 ()	方差 %	累计方差 %	特征根 ()	方差 %	累计方差 %
1	2.450	40.827	40.827	2.166	36.106	36.106
2	1.893	31.556	72.383	1.938	32.303	68.409
3	1.115	18.579	90.962	1.353	22.553	90.962

提取方法：主成分分析；旋转方法：方差最大正交旋转

表 2 因子成分得分系数矩阵

Tab. 2 Component score coefficient matrix

变量	x_1	x_2	x_3	x_4	x_5	x_6
因子 1	0.452	0.458	0.039	0.055	0.061	0.222
因子 2	-0.046	-0.013	-0.124	0.54	0.303	-0.408
因子 3	0.12	-0.051	0.759	-0.377	0.162	-0.184

提取方法：主成分分析；旋转方法：方差最大正交旋转

法解释，它们是剔除相关信息后的提纯，是从 6 维空间到 3 维空间降维处理的结果。

根据表 2 可以列出前三个公因子 (f_1, f_2, f_3) 与相应变量 ($x_1 \sim x_6$) 的数学表达式。每个评价单元的综合得分值 V 可由原始数据与表 1 提供的特征根和表 2 提供的系数计算出。综合得分值的计算从某种意义上是做进一步的降维分析，当然这些综合得分值只具有相互比较的意义。 V 值是含有结构性特征的从多维空间到一维空间的映射，完全不同于一般的诸如土地利用率和农业用地指数等较为综合的指标。

综合得分值大致介于 $-11.5 \sim 5.0$ 之间。空间展示的综合得分值格局表明：得分值 $V > 0.1$ 的评价单元主要分布于东部季风影响区，是得分高值区； V 值介于 $-3.3 \sim 0.1$ 的评价单元主要分布于新疆北部、青藏高原东南端和内蒙古中东部； $V < -6.6$ 的评价单元集中分布于新疆的南部、青藏高原的北部和内蒙古的西北端，是得分低值区；其它地区的 V 值介于 $-3.3 \sim -6.6$ 。 V 值反映含有结构性映射的土地利用程度的综合评价状况，其基本格局反映了综合生态地理条件的总体分异特点，是具体分区定界的主要参考。 V 值越高，说明综合生态地理条件越好，土地利用率和土地利用集约性程度越高，主要表现为土地利用率高、农业用地指数、垦殖率、森林覆盖率和林地利用率高。 V 值越低，则与上述情况相反，主要表现为牧草地比重较高，而其它用地指数较低则反映未利用土地比重较大。

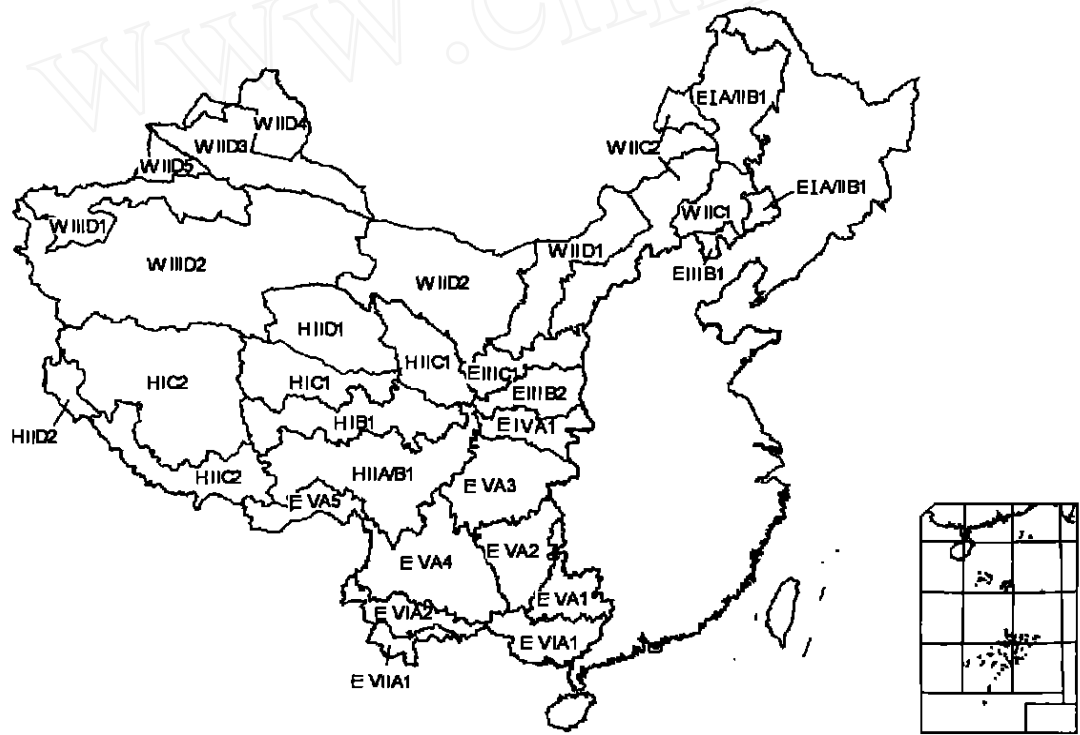


图 2 西部土地利用程度分区示意图

Fig. 2 Division of land use degree in western China

表 3 土地利用程度分区及各区土地利用程度 (%)

Tab. 3 Division of land use degree and land use degrees in each region (%)

编号	代号	小区	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
1	EIA/ IIB ₁	内蒙古东北端	92.4	90.0	9.2	53.2	91.3	27.6	1.3
2	EIIB ₁	赤峰及其以南地区	95.5	90.2	29.9	25.8	70.3	34.1	4.2
3	EIIIC ₁	陕北甘东高原丘陵	88.3	84.6	33.8	14.5	50.7	35.7	3.1
4	EIIB ₂	关中盆地	89.5	82.5	36.3	31.8	71.2	13.3	5.7
5	EIVA ₁	汉中盆地	93.6	90.3	17.4	66.7	76.7	5.7	1.9
6	EVA ₁	江南与南岭山地	77.6	73.2	15.4	52.0	71.1	5.1	2.4
7	EVA ₂	贵州高原	81.7	77.8	26.8	43.4	55.3	7.2	2.8
8	EVA ₃	四川盆地	83.9	73.4	33.0	36.4	69.5	1.9	7.0
9	EVA ₄	云南高原	82.1	78.5	17.1	53.6	63.8	7.1	2.2
10	EVA ₅	东喜马拉雅南翼	78.9	78.5	0.2	68.1	76.5	10.2	-
11	EVIA ₁	广西丘陵平原	81.1	78.0	18.4	57.6	72.2	0.5	2.0
12	EVIA ₂	滇中山地丘陵	78.6	71.4	23.1	44.2	64.3	2.6	3.8
13	EVIA ₁	滇南谷地丘陵	78.3	76.1	15.1	56.4	82.0	0.5	1.2
14	WIIC ₁	松辽平原西南部及大兴安岭南端	96.7	93.6	11.7	12.8	47.5	68.9	2.0
15	WIIC ₂	内蒙古高平原东部	96.3	93.3	11.1	6.2	34.0	76.0	1.3
16	WIID ₁	内蒙古高平原西部及河套	91.4	88.7	7.0	1.2	39.6	80.4	1.3
17	WIID ₂	阿拉善与河西走廊	54.0	52.2	2.4	2.9	22.3	46.9	1.2
18	WIID ₃	准噶尔盆地	67.6	65.4	4.8	5.4	29.8	55.3	1.0
19	WIID ₄	阿尔泰山与塔城盆地	91.1	89.5	3.4	6.6	58.3	79.5	0.5
20	WIID ₅	伊犁盆地	86.4	83.5	10.1	8.9	78.4	64.4	1.6
21	WIID ₁	天山南坡	58.5	55.4	3.9	6.7	38.7	44.6	1.2
22	WIID ₂	塔里木与吐鲁番盆地	15.0	14.1	0.7	2.3	15.8	11.0	0.2
23	HB ₁	果洛那曲丘状高原	85.3	84.1	0.2	5.5	20.8	78.4	0.1
24	HIC ₁	青南高原宽谷	48.2	44.3	-	-	-	44.3	-
25	HIC ₂	羌塘高原湖盆及昆仑高山高原	56.4	53.1	-	-	-	53.1	-
26	HIIA/ B ₁	川西藏东高山深谷	82.2	81.5	0.9	36.1	55.7	44.4	0.1
27	HIIC ₁	青东祁连山地	76.8	72.7	4.3	7.7	30.7	60.7	1.2
28	HIIC ₂	藏南山地	77.0	74.3	1.1	4.1	43.3	69.1	0.1
29	HIID ₁	柴达木盆地	19.2	17.5	0.1	0.9	1.6	16.4	0.2
30	HIID ₂	阿里山地	64.1	62.5	-	-	-	62.5	-

注：E、W、H 分别表示：东部季风区、西北干旱区和青藏高寒区；I~VII 分别表示：寒温带（I）、中温带（II）、暖温带（III）、北亚热带（IV）、中亚热带（V）、南亚热带（VI）、边缘热带（VII）；A、B、C、D 分别表示：湿润、半湿润、半干旱、干旱；x₁~x₇ 表示：土地利用率（x₁）、农业用地指数（x₂）、垦殖率（x₃）、森林覆盖率（x₄）、林地利用率（x₅）和牧草地比重（x₆）、非农建设用地比重（x₇）；表中的 -，指其值不足 0.06 %。

3.2 土地利用程度分区

应用自上而下和自下而上的分区方法，基于因子分析综合得分值的分布格局，叠加中国生态地理区域系统图，结果先将研究范围分为三大区，即东部季风区（E）、西北干旱区（W）和青藏高寒区（H）。然后对各大区内的评价单元做进一步归并，东部季风区含 13 个小区，西北干旱区含 9 个小区，青藏高寒区含 8 个小区（图 2、表 3）。此分区结果是生态地理区域系统在土地利用程度方面的具体化，其具体界线是参照生态地理界线的县市级行政区界的归并。为增强分区结果应用的可操作性，保持了县市行政界线的完整，这样在土地面积较大的县市内，其生态地理条件仍然存在较大差异，但这是顾此失彼的必然结果。分区格局既反映自然因素方面的水热状况，也表征人为因素方面的土地利用特点。

3.3 区间土地利用程度的静态比较

根据表 3 和图 3，总体而看，土地利用率、农业用地指数和林地利用率的空间变化趋

势大致相似，它们在 E 区内的变异相对平缓，而在 W 和 H 区内的变异起伏较大。说明与 E 区相比，W 和 H 区的土地利用率有待于提高。森林覆盖率和牧草地比重，基本上是负相关的，而且森林覆盖率的变异于 E 区内的较大，在 W 区和 H 内较小，但靠近 E 区的川西藏东高山深谷区例外。说明东部区有些地方的森林破坏严重，容易出现水土流失。垦殖率在 E 区较高而且变异较大，在 W 区和 H 区较低而且变异很小。垦殖率与非农建设用地指数基本上呈空间正相关，说明耕种条件好的地区，也是非农建设用地（主要是居民用地）较多的地方。但这并不表明垦殖率越高，非农建设用地指数也高这种动态关系；相反，随着非农建设用地的不断增多，耕地会逐渐减少。

东部季风区是评价单元综合得分值 $V > 0.1$ 的集中空间分布区。反映了东部季风区在比较好的生态地理条件下，大部分地区的土地利用率很高，而且非农建设用地指数比其它区高，但牧草地比重很低。E 区的土地利用率和农业用地指数大致表现为亚热带以北较高，而其以南较低。垦殖率和森林覆盖率基本上呈空间负相关性，如四川盆地与东喜马拉雅南翼相比，前者的垦殖率较高，而森林覆盖率较低；后者则正好相反。垦殖率与牧草地比重在汉中盆地以北基本上呈空间正相关，在汉中盆地以南则基本上呈空间负相关。这些都反映了土地利用存在一定的纬向分布特征。

西北干旱区的内，综合得分值的变化较大。反映了区域空间的生态地理条件变化较大。具体表现为土地利用率、农业用地指数和林地利用率变异很大。垦殖率、森林覆盖率以及非农建设用地指数都较东部季风区低，且变异很小。牧草地比重明显高于东部区，而且变异性大，其中阿拉善与河西走廊区以及塔里木与吐鲁番盆地区的牧草地比重较低。草地在西北区具有重要的生态作用。各小区内各项指标值均呈正相关，只是程度不同。

青藏高寒区内，各种指数的变化特征与西北干旱区非常接近，只是垦殖指数、非农建设用地指数更低；除川西藏东高山深谷区外，总体的森林覆盖率也很低；除柴达木盆地外，牧草地比重较高而且空间变异较大。本大区从东南向西北，垦殖率和森林覆盖率不断减少，牧草地比重则呈现先增加后减少的趋势。

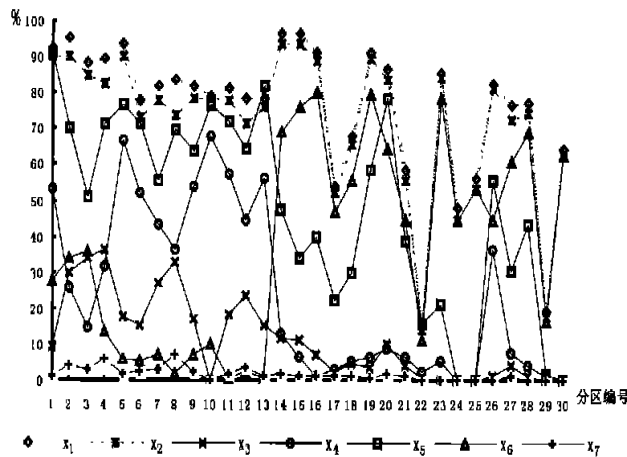


图 3 不同区土地利用率比较

Fig. 3 Comparison of land use degrees among different regions

4 结语

常用的有关土地利用程度的各种指标,只是从几个侧面反映土地利用情况。因此,单纯利用这些指标进行数学处理和分区研究,往往会偏离主导方向。以中国生态地理区域系统研究成果作为宏观控制性框架,自下而上的归并才能有层次性和方向性的指导。

利用多元统计中的因子分析方法,使多指标提供的许多信息得以提纯和综合,降维的结果还能保持样本在高维空间的某种“结构”,其中最明显的是与“排序”有关的结构。这样,各个样本间的综合比较和归并有了直接参考的、含有主要信息的单一指标。GIS技术的应用,使样本之间的综合数值排序与样本的空间属性紧密相关,也使综合得分值相近,而土地利用结构相差甚远的样本的不合理归并得以消除。GIS技术更使自上而下和自下而上的区划方法得以结合。

相邻各小区之间许多属性是逐渐过渡的,因此分区的结果也是相对的。为增强分区结果的应用必须照顾县级行政区界;而有的县区面积很大,其内的生态地理条件存在较大分异。随着土地利用调查信息的逐步细化,研究手段的不断完善,县域面积较大的土地利用程度分区研究有待进一步深入。

参考文献:

- [1] 罗明,王军.中国土地整理的区域差异和对策[J].地理科学进展,2001,20(2):97~103.
- [2] 吴传钧,郭焕成.中国土地利用[M].北京:科学出版社,1994.
- [3] 徐继填,陈百明,张雪芹.中国生态系统生产力区划[J].地理学报,2001,56(4):1~8.
- [4] 全国农业区划委员会.中国农业自然资源和农业区划[M].北京:农业出版社,1989.
- [5] 刘明光.中国自然地理图集[M].北京:中国地图出版社,1998.
- [6] 刘彦随,冯得显.三峡库区土地持续利用潜力与途径模式[J].地理研究,2001,20(2):139~145.
- [7] 陈彦光,刘继生.城市土地利用结构和形态的定量描述:从信息熵到分数维[J].地理研究,2001,20(2):146~152.
- [8] 王学军.空间分析技术与地理信息系统的结合[J].地理研究,1997,16(3):70~74.
- [9] 庄大方,刘纪远.中国土地利用程度的区域分异模型研究[J].自然资源学报,1997,12(2):105~111.
- [10] 黄秉维.中国综合自然区划纲要(1986年)[A].地理集刊(21)[C].北京:科学出版社,1999.10~20.
- [11] Zheng Du. A study on the eco-geographic regional system of China[A]. FAO FRA 2000 Global Ecological Zoning Workshop, Combrige, U K, July 28~30, 1999.
- [12] 吴绍洪,郑度,杨勤业.我国西北地区生态地理区域系统与生态建设战略初步研究[J].地理科学进展,2001,20(1):10~20.
- [13] 杨勤业,郑度,吴绍洪.中国的生态地域系统研究[J].自然科学进展,2002(待刊).
- [14] 李元.中国土地资源[M].北京:中国大地出版社,2000.
- [15] 唐华俊,陈佑启,Eric Van Ranst.中国土地资源可持续利用的理论与实践[M].北京:中国农业科技出版社,2000.
- [16] 于秀林,任雪松.多元统计分析[M].北京:中国统计出版社,1999.
- [17] 胡永宏,贺思辉.综合评价方法[M].北京:科学出版社,2000.

Division of land use degree in western China using factor analysis

WANG Xiu-hong¹, HE Shu-jin¹, ZHANG Yi-li¹, LUO Ming²

(1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. Department of Urban and Environmental Sciences, Peking University, Beijing 100871, China)

Abstract : Integrated evaluation on land use degree for each county in western China was made by using factor analysis. Based on the integrated score of each evaluated unit (county or county-level city), which reflected the structure of land use, and taken the eco-geographical system of China as a controlling framework, evaluated units were united with the GIS technique; thus, the combined division method of From top-to-bottom (TB) and From bottom-to-top (BT) was tested. Western China, which included 12 provinces, autonomous regions and municipality, were divided into three large regions of East monsoon region, Northwest arid region and Tibetan alpine region and further divided into 30 subregions.

The various types of indices commonly used to illuminate land use degree just reflect the situation of land use from a few on the sides. Therefore, the simply using of these indices to proceed mathematical analyses and to do division research would usually deviate from the right direction. Taken the eco-geographical system of China as a controlling framework, the combined division method of TB and BT can be used with hierarchical structure and suggestive direction. Using the factor analysis, information provided by many indices can be purified and synthesized. The result of data deduction also keeps the structure of samples in high dimension, especially the structure related to ordering. Thus, the comparison and union of the evaluated units have one directly referenced index rather than lots of indices. The application of GIS technique makes the ordering of integrated score closely related to the space attributes among samples, and avoid the irrational union of evaluated units which have almost the same integrated score but have very different land use structure. Also the GIS technique makes the division methods combined between that of TB and that of BT.

Many attributes of the divided regions are gradually changing; therefore, the result of division is opposite. County boundary must be considered for maneuverability; however, some counties have large areas, resulting in rather different ecogeographic conditions. With more detailed inquisition information about land use and perfecting study method, the division result would be more reasonable.

Key words factor analysis; western China; land use degree; division