

汪霖,蒲春玲,刘志有,等. 基于主成分分析和熵权法的区域土地资源集约利用研究——以新疆伊犁州直为例[J]. 江苏农业科学,2016,44(5):522-525.

doi:10.15889/j.issn.1002-1302.2016.05.149

# 基于主成分分析和熵权法的区域土地资源集约利用研究 ——以新疆伊犁州直为例

汪霖,蒲春玲,刘志有,黄晓东,潘丽萍

(新疆农业大学管理学院,新疆乌鲁木齐 830000)

**摘要:**丝绸之路经济带战略的实施,促进了新疆城市化、工业化的快速发展,区域土地资源集约利用已经成为新疆可持续发展的重要内容。本研究以新疆伊犁哈萨克自治州直为例,构建土地集约利用评价指标体系,基于 SPSS 21.0 统计软件,运用主成分分析对伊犁州直土地集约利用进行了时序方面的模型分析,同时采用熵权法确定指标权重。通过分析得出:(1)伊犁州直在 2003—2012 年,土地集约利用程度持续不断提高;(2)对伊犁州直土地集约利用贡献率高的因素依次是地均固定资产投资、地均农业产值、地均 GDP、地均工业产值等。

**关键词:**主成分分析;熵权法;土地集约利用;伊犁州直

**中图分类号:** F301.24 **文献标志码:** A **文章编号:** 1002-1302(2016)05-0522-04

土地是人类一直以来赖以生存和发展的基础,土地资源的有限性及稀缺性要求人们集约利用土地。近年来,随着中国人口不断增加,城市化及工业化的持续发展,人地矛盾越来越突出<sup>[1]</sup>。伴随着丝绸之路经济带建设的不断深入,新疆经济发展也迎来了更大的机遇。经济发展的同时,也不可避免地产生土地供求矛盾尖锐、土地利用粗放等问题。为了解决这些问题,缓解人地矛盾和土地供求矛盾,就需要对土地进行集约利用<sup>[2]</sup>。集约经营的概念最早来自农业,而在土地经济研究领域,土地利用集约度是指单位面积土地上劳力、资金、技术、物质等投入的密集程度<sup>[3]</sup>。

目前,国内研究学者们在评价方法上对土地集约利用做了深入的研究和分析,具有代表性的方法主要是回归分析法、德尔菲法、层次分析法、模糊评价法等数学模型和统计方法,分别计算土地利用程度、土地利用效益、土地投入强度和土地利用集约度,并解决土地集约利用的问题<sup>[4]</sup>。以前的土地集约利用评价中也采用过主成分分析法和熵值法,但是一般都是单纯地利用某一种。本研究则采用 2 种分析方法相结合,首先运用主成分分析法简化指标体系,进而用熵值法求取客观权重。利用主成分分析法合理实现了指标体系由多到少的降维和简化;利用熵值法确定了各指标的权重,一定程度上避免了赋权的片面性和主观性,使得到的权重更具客观性和科学性<sup>[5]</sup>。

本研究以新疆伊犁州直属县市为主要研究对象,选取了 11 项指标,通过查阅相关年鉴,结合 SPSS 统计软件进行主成

分分析,对伊犁州直的土地利用水平进行时序性评价,同时,运用 Excel 进行熵权的分析,对 11 项指标进行权重的计算,以便得出每项指标对于土地集约利用水平的影响度。对伊犁州直的土地集约利用进行评价,研究其土地集约利用水平变化特征,能够有助于改变该地区土地粗放利用的现状,提高土地利用效率,为土地合理利用提供科学依据。

## 1 研究区概况

伊犁哈萨克自治州直位于 80°09'~84°56'E,42°14'~44°50'N 之间,下辖 8 县 2 市,包括伊宁市、奎屯市、伊宁县、霍城县、巩留县、新源县、昭苏县、特克斯县、尼勒克县、察布查尔锡伯自治县等。2014 年 6 月,县级霍尔果斯市正式设立,由伊犁哈萨克自治州管辖,本研究中的数据截止到 2013 年,故未将霍尔果斯市也列入本研究中。该地区属于温带大陆性气候,夏季温暖,冬春季温和,空气湿润,具有塞外江南的特殊称号,地势西高东低,南高北低,地形由西南向东北倾斜,形成四谷一盆的伊犁河谷地区。伊犁河谷的水资源十分丰富,年径流量达到 167 亿 m<sup>3</sup>。至 2012 年年末,伊犁哈萨克自治州州直总人口为 291.72 万人,其中少数民族 183.30 万人,占总人口的 62.83%。伊犁州直国民经济发展迅速。2013 年,伊犁哈萨克自治州直属县市生产总值 681.93 亿元,比上年增长 14.8%;州直人均生产总值 24 895 元,增长了 14.3%<sup>[6]</sup>。

## 2 数据来源与研究方法

### 2.1 数据来源

本研究采用的数据来源于《新疆统计年鉴(2004—2013 年)》和《伊犁哈萨克自治州统计年鉴(2004—2013 年)》。

### 2.2 研究方法

**2.2.1 主成分分析** 主成分分析法是利用降维的思想,把多个指标和数据简化为少量的综合指标,同时又使这少量指标尽可能地包含原指标群中的信息资料。这些综合指标能更合

收稿日期:2015-04-17

基金项目:国家社会科学基金(编号:14XGL005);新疆社会科学基金(编号:12BJL130)。

作者简介:汪霖(1989—),女,湖北武汉人,硕士研究生,研究方向为土地资源可持续发展。E-mail:380509133@qq.com。

通信作者:蒲春玲,博士,教授,博士生导师,研究方向为土地资源可持续发展。E-mail:puchunling@163.com。

理地反映样本之间的差别,而在统计意义上是相互独立的<sup>[7]</sup>。

主成分分析步骤如下:(1)建立指标体系和原始数据矩阵  $X$ 。(2)对原始数据进行标准化处理。本研究采用 SPSS 21.0 统计分析软件进行数据处理,数据标准化方法是 SPSS 默认的标准方法即  $Z$ -Score 标准化。该方法是基于原始数据的均值和标准差进行数据的标准化。标准化计算公式如下:

$$Z_{ij} = \frac{X_{ij} - X_i}{S_i} (i = 1, 2, \dots, m; j = 1, 2, \dots, n)。(1)$$

式中: $Z_{ij}$ 为标准化后的变量值; $X_{ij}$ 为第  $i$  个样本第  $j$  个指标的原始数据; $X_i$  为第  $j$  个指标数据的算术平均值; $S_i$  为第  $j$  个指标数据的标准差。(3)由标准化数据矩阵计算相关系数矩阵  $R$ 。(4)计算相关矩阵  $R$  的特征值  $\lambda$ 、主成分的贡献率及累积贡献率。(5)根据主成分方差累积贡献率(一般取值 85% 以上)确定主成分个数  $N$ 。

2.2.2 熵权法 熵权法是一种客观赋权法,它的基本思路是根据指标变异性的确定客观权重,即它主要是根据各指标对于目标层所包含的信息量的多少来确定其权重<sup>[8]</sup>。

熵权法计算步骤如下:(1)数据标准化,为了使各年份各项指标之间具有可比性,就需要消除指标之间的差异性,即对初始数据进行无量纲化处理。本研究中所选指标,均是正向指标,即指标数据越大,对目标层的影响就越大。正向指标数据标准化的计算公式如下:

$$Y_{ij} = \frac{X'_{ij} - \min(X'_i)}{\max(X'_i) - \min(X'_i)}。(2)$$

熵权法的计算采用的是原始数据矩阵  $X$  矩阵的转置矩阵  $X'$  矩阵。其中  $i = 1, 2, \dots, n; j = 1, 2, \dots, m$  (本研究采用

2003—2012 年的数据,共 10 年,即  $i = 1, 2, \dots, 10$ ;采用了 11 项指标,即  $j = 1, 2, \dots, 11$ )。(2)计算第  $i$  年第  $j$  项指标占有年份该指标和的比重:

$$P_{ij} = \frac{Y_{ij}}{\sum_{i=1}^n Y_{ij}}。(3)$$

(3)计算第  $j$  项指标的熵值:

$$e_j, e_j = -\ln(n)^{-1} \times \sum_{i=1}^n P_{ij} \ln P_{ij}。(4)$$

(4)求各因子的熵权:

$$w_j = \frac{1 - e_j}{m - \sum_{j=1}^m e_j}。(5)$$

一般而言,指标的信息熵越小,表明指标值的变异程度越大,提供的信息量越多,在综合评价中能起到的作用也越大,其权重也就越大。相反,指标的信息熵越大,表明指标值的变异程度越小,提供的信息量也越少,在综合评价中起到的作用也越小,其权重也就越小<sup>[9]</sup>。

3 结果与分析

3.1 指标体系构建

土地集约利用是一个内涵丰富的概念,涉及经济、社会等诸多方面<sup>[10]</sup>。依据土地集约利用的概念和内涵,并结合新疆伊犁州直的实际情况,在遵循科学性、前瞻性、指导性和可量化性等原则的基础上,从土地利用投入强度、土地利用程度、土地利用效益、土地可持续利用度 4 个方面,选取了共计 11 项指标来构建伊犁州直土地集约利用的评价体系(表 1)。

表 1 土地集约利用评价指标体系

目标层	准则层	指标层
土地利用集约度	土地利用投入强度	地均固定资产投资(万元/km <sup>2</sup> ) 人均道路面积(m <sup>2</sup> )
	土地利用程度	在岗职工平均工资(元) 人口密度(人/km <sup>2</sup> ) 城镇化率(%)
	土地利用效益	地均 GDP(万元/km <sup>2</sup> ) 地均农业产值(万元/km <sup>2</sup> ) 地均工业产值(万元/km <sup>2</sup> ) 地均社会消费品零售额(万元/km <sup>2</sup> )
	土地可持续利用度	人均公共绿地面积(m <sup>2</sup> ) 森林覆盖率(%)

3.2 主成分分析

运用 SPSS 21.0 统计分析软件,对 2003—2012 年 11 项指标进行分析,得到其特征值、方差贡献率及累计贡献率(表 2)。根据累计贡献率大于 85% 的原则,选取前 3 项作为主成分。第 1 主成分的方差贡献率为 83.746%,第 2 主成分的方差贡献率为 12.148%,第 3 主成分的方差贡献率为 2.663%。

通过分析得知,第 1 主成分在地均 GDP、地均社会消费品零售总额、地均工业产值、地均农业产值、在岗职工平均工资、人口密度的载荷最大,主要体现了土地的利用效益;第 2 主成分在城镇化率的载荷最大,反映了土地的利用程度;第 3 主成分在人均公共绿地面积、森林覆盖率的载荷最大,体现了土地可持续利用程度。

计算选取的 3 个主成分对应的特征向量,写出主成分表达式。主成分载荷表示各项指标与主成分之间的相关系数<sup>[11]</sup>。特征向量表示不同主成分中各项指标的权重。主成分载荷系数  $u_{ij}$ 、特征值  $\lambda_i$ 、特征向量  $\vec{a}$  之间的关系为:  
 $\vec{a} = \frac{u_{ij}}{\sqrt{\lambda_i}} (i = 1, 2, 3 \dots)$ 。因此,根据表 3 中的数据除以主成分对应的特征值的平方根就可以得到各项指标对应的特征向量,特征向量即是主成分表达式系数。

成分系数的求法是:各自主成分载荷向量除以各自主成分特征值的算术平方根<sup>[12]</sup>。

第 1 主成分的函数表达式:

表 2 总方差分解

成分	初始特征值			提取平方和载入		
	合计	方差 (%)	累积 (%)	合计	方差 (%)	累积 (%)
1	9.212	83.746	83.746	9.212	83.746	83.746
2	1.336	12.148	95.893	1.336	12.148	95.893
3	0.293	2.663	98.556	0.293	2.663	98.556
4	0.112	1.018	99.574			
5	0.025	0.223	99.797			
6	0.015	0.137	99.934			
7	0.003	0.029	99.964			
8	0.003	0.028	99.992			
9	0.001	0.008	100.000			
10	$1.002 \times 10^{13}$	$1.017 \times 10^{13}$	100.000			
11	$-1.002 \times 10^{13}$	$-1.019 \times 10^{13}$	100.000			

表 3 成分矩阵

指标	成份		
	1	2	3
地均固定资产投资(万元/km <sup>2</sup> )	0.926	-0.185	-0.281
人均道路面积(m <sup>2</sup> )	0.795	0.601	0.020
在岗职工平均工资(元)	0.981	-0.136	-0.117
人口密度(人/km <sup>2</sup> )	0.980	0.080	0.022
城镇化率(%)	0.532	0.835	0.069
地均 GDP(万元/km <sup>2</sup> )	0.995	-0.075	-0.051
地均农业产值(万元/km <sup>2</sup> )	0.988	-0.134	-0.065
地均工业产值(万元/km <sup>2</sup> )	0.991	-0.089	-0.056
地均社会消费品零售总额(万元/km <sup>2</sup> )	0.995	-0.072	-0.028
人均公共绿地面积(m <sup>2</sup> )	0.804	-0.426	0.379
森林覆盖率(%)	0.966	0.015	0.201

$$Y_1 = 0.926ZX_1 + 0.795ZX_2 + 0.981ZX_3 + 0.980ZX_4 + 0.532ZX_5 + 0.995ZX_6 + 0.988ZX_7 + 0.991ZX_8 + 0.995ZX_9 + 0.804ZX_{10} + 0.966ZX_{11} / \sqrt{9.212};$$

第 2 主成分的函数表达式:

$$Y_2 = -0.185ZX_1 + 0.601ZX_2 - 0.136ZX_3 + 0.080ZX_4 + 0.835ZX_5 - 0.075ZX_6 - 0.134ZX_7 - 0.089ZX_8 - 0.072ZX_9 - 0.426ZX_{10} + 0.015ZX_{11} / \sqrt{1.336};$$

第 3 主成分的函数表达式:

$$Y_3 = -0.281ZX_1 + 0.020ZX_2 - 0.117ZX_3 + 0.022ZX_4 + 0.069ZX_5 - 0.051ZX_6 - 0.065ZX_7 - 0.056ZX_8 - 0.028ZX_9 + 0.379ZX_{10} + 0.201ZX_{11} / \sqrt{0.293};$$

$$\text{综合得分} = 0.8375Y_1 + 0.1215Y_2 + 0.0266Y_3。$$

通过 SPSS 21.0 测算出的 3 个主成分及其所占权重,并且计算出 2003—2012 年伊犁州直土地利用得分,计算结果见表 4。

本研究运用主成分分析法评价 2003—2012 年伊犁州直土地集约利用的时序变化规律,综合以上分析,在伊犁州直土地集约利用分析中,充分考虑了人口密度、城镇化率、地均工业产值、地均农业产值等诸多条件及因素,构建了伊犁州直土地集约利用评价指标体系。根据 2003—2012 年数据分析结果,伊犁州直的土地利用集约度的综合得分由 -3.31 不断上升,至 2012 年达到 4.25,说明伊犁州直的土地利用集约度逐年增高。总体来看,伊犁州直的土地利用集约情况良好。这

表 4 不同年份主成分因子和综合得分

年份	$Y_1$	$Y_2$	$Y_3$	$Y$
2003	-3.68	-2.02	0.51	-3.31
2004	-3.67	-1.30	-0.67	-3.25
2005	-2.31	1.25	-0.44	-1.79
2006	-1.56	1.23	0.03	-1.16
2007	-1.03	1.26	-0.37	-0.72
2008	0.01	0.67	0.40	0.10
2009	1.16	0.19	0.88	1.02
2010	2.14	0.05	0.38	1.81
2011	3.71	-0.35	0	3.06
2012	5.24	-0.98	-0.72	4.25

正是由于伊犁州直属县市正处于城市化加速发展的阶段,经济社会发展对土地资源需求十分强烈。伊犁州直在“十一五”期间,基础设施建设不断加强,土地利用程度不断提高,建成了精伊霍铁路、奎北铁路、那拉提机场、清伊高速公路、果霍高速公路,以及完成了大中型灌区续建与改造等重大工程;城镇化进程加快,城镇化率达到 42%。在“十二五”期间,伊犁州直的发展面临难得的机遇,拥有着诸多有利条件,例如西部大开发战略以及丝绸之路经济带的建设,在 2011—2015 年间,伊犁州直加速推进新型工业化,促进了土地集约节约利用。但伊犁州直土地集约利用程度距我国东部地区还有很大距离,说明其土地集约利用发展空间仍具有较大潜力。

3.3 熵权法分析

根据熵权法的计算步骤,对伊犁州直 2003—2012 年的 11 项指标的信息熵进行了计算。根据 11 项评价指标的统计数据,建立原始数据的矩阵,根据公式(2)对其进行标准化,将标准化结果代入公式(3)计算各指标所占比重  $P_{ij}$ ,然后根据公式(4)进行各指标信息熵  $e_j$  的计算,最后将各指标的  $e_j$  代入公式(5)得到各指标的熵权  $w_i$ 。各评价指标权重及准则层权重见表 5。从表 5 可以看出,土地利用效益对于土地集约利用的影响最大,所占权重为 0.382 5。其中地均农业产值的影响最大,所占权重为 0.114 2,其次为地均 GDP,权重为 0.092 7,再次是地均工业产值。在所有指标当中,地均固定资产投资的权重最大,达到 0.144 7。目前,伊犁州直地均固定资产投资、地均农业产值、在岗职工平均工资、森林覆盖率、地均 GDP、地均工业产值等因素对伊犁州直土地集约利用的贡献位于所有指标前列,因此,在今后经济发展中更应注重经济发展的内涵,大力发展集约经济,杜绝土地资源闲置和低效利用的情况发生。从土地可持续利用度看,其权重只有 0.199 0,表明该指标对生态环境变化反应敏感,在未来的经济发展中,应重点采取相应措施,避免因环境污染而导致土地资源的浪费,以提高土地资源的集约利用程度。

4 结论

利用主成分分析和熵权法对 2003—2012 年伊犁州直的土地集约利用情况进行了综合分析,利用主成分分析法,该方法旨在利用降维的思路将指标数量简化,从 11 项数据指标中提取了前 3 项指标,能够综合反映伊犁州直的土地集约利用情况。并将标准化之后的数据导入 SPSS 进行计算,得出了 2003—2012 年伊犁州直的土地集约利用情况是随着年份的

表 5 各项指标权重

指标	准则层	指标代号	各指标权重	准则层权重
地均固定资产投资(万元/km <sup>2</sup> )	土地利用投入强度	$X_1$	0.144 7	0.190 2
人均道路面积(m <sup>2</sup> )		$X_2$	0.045 5	
在岗职工平均工资(元)	土地利用程度	$X_3$	0.110 7	0.227 7
人口密度(人/km <sup>2</sup> )		$X_4$	0.061 0	
城镇化率(%)		$X_5$	0.056 0	
地均 GDP(万元/km <sup>2</sup> )	土地利用效益	$X_6$	0.092 7	0.382 5
地均农业产值(万元/km <sup>2</sup> )		$X_7$	0.114 2	
地均工业产值(万元/km <sup>2</sup> )		$X_8$	0.090 2	
地均社会消费品零售总额(万元/km <sup>2</sup> )	土地可持续利用度	$X_9$	0.085 4	
人均公共绿地面积(m <sup>2</sup> )		$X_{10}$	0.090 2	0.199 0
森林覆盖率(%)		$X_{11}$	0.108 8	

推进,土地利用更加集约。伊犁州直土地集约利用度从 2003 年的 -3.31 上升至 2012 年的 4.25,土地集约利用程度有了大幅度提高,能够合理有效地使用土地。而这与土地投入的不断增加有着密切的关系。采用熵权法计算得出各指标所占权重,在伊犁州直城市化迅速发展的阶段,地均固定资产投资不断增加,它是促进伊犁州直土地集约利用的主要因素,此外,地均农业产值、在岗职工平均工资、森林覆盖率、地均 GDP、地均工业产值等因素对伊犁州直土地集约利用也有巨大贡献。

在城市发展开发过程中还需要注意生态环境等问题,避免因污染而影响土地集约利用;在今后发展过程中,还应充分利用闲置土地,注重土地整理,进行土地挖潜,提高城市土地的空间利用率和容积率;合理统筹不同产业对土地的需求,实行区域用地年度管理制度和用地总量控制等,以提高伊犁州直的土地利用效益,实现土地资源的可持续发展。

由于耕地集约利用涉及面广、影响因素众多,加之可借鉴的相关研究成果不多和部分指标数据的不可获取,本研究只选取了部分主要指标对伊犁州直土地集约利用进行综合评价,其评价指标体系有待进一步完善。

参考文献:

[1]朱天明,杨桂山,万荣荣. 城市土地集约利用国内外研究进展[J]. 经济地理,2009,29(6):977-983.

[16]周仕千,陈张雷,郭 丽. 西南地区 3 种护坡灌木对干旱胁迫的响应研究[J]. 安徽农业科学,2011,39(35):21712-21713,21782.

[17]李慧卿,马文元. 沙生植物抗旱性比较的主要指标及分析方法[J]. 干旱区研究,1998,15(4):12-15.

[18]蒋宣斌,耿养会,彭 秀,等. 常绿岩垂草抗旱生理特性分析[J]. 西南林业大学学报,2012,32(4):17-21.

[19]李 源,师尚礼,王 赞,等. 干旱胁迫下鸭茅苗期抗旱性生理研究[J]. 中国草地学报,2007,29(2):35-40.

[20]芦 静,曹俊梅,周安定,等. PEG 处理下新疆冬小麦品种幼苗期生理指标的抗旱性研究[J]. 新疆农业科学,2014,51(3):393-402.

[2]毕宝德. 土地经济学[M]. 6 版. 北京:中国人民大学出版社,2010.

[3]陈柯斯. 报酬递减规律在土地集约利用中的应用[J]. 商业文化:学术版,2008(6):152.

[4]杨红梅,邱道持,张传华,等. 基于因子分析的城市土地集约利用比较研究[J]. 西南师范大学学报:自然科学版,2006,31(1):165-169.

[5]温秀琴,汪应宏,陈 晨,等. 主成分分析法与熵值法结合用于土地集约利用评价——以徐州市为例[J]. 安徽农业科学,2008,36(28):12372-12373,12538.

[6]伊犁哈萨克自治州 2013 年国民经济和社会发展统计公报[R]. (2014-02-25)[2015-01-15]. <http://www.xjyl.gov.cn/info/1072/46598.htm>.

[7]渠丽萍,张丽琴,胡伟艳. 城市土地集约利用变化影响因素研究——以武汉市为例[J]. 资源科学,2010,32(5):970-975.

[8]赵小凤,黄贤金,陈 逸,等. 城市土地集约利用研究进展[J]. 自然资源学报,2010(11):1979-1996.

[9]张富刚,郝晋珉,姜广辉,等. 中国城市土地利用集约度时空变异分析[J]. 中国土地科学,2005,19(1):23-29.

[10]张田鑫. 西安市城市土地集约利用评价及提高途径研究[D]. 西安:西安建筑科技大学,2013.

[11]李洪波. 城市土地集约利用途径探索[D]. 北京:北京林业大学,2010.

[12]曹银贵,周 伟,王 静,等. 基于主成分分析与层次分析的三峡库区耕地集约利用对比[J]. 农业工程学报,2010,26(4):291-296.

[21]周锋利,宋西德,齐高强,等. 臭柏抗旱生理特性研究[J]. 西南林学院学报,2005,25(3):1-4.

[22]孟 林,杨宏新,毛培春,等. 偃麦草属植物种间苗期抗旱性评价[J]. 草业学报,2011,20(5):34-41.

[23]陈 娟,陈其兵,潘远志,等. 6 种野生灌木的抗旱性研究[J]. 四川林业科技,2007,28(5):50-54,49.

[24]张立军,胡文玉,戴俊英. 渗透胁迫下玉米幼苗离体叶片膜透性变化机理的研究[J]. 沈阳农业大学学报,1996,27(3):207-210.

[25]陈志成,王志伟,王荣荣,等. 3 种阔叶树种对持续干旱的生理响应及抗旱性评价[J]. 中国水土保持科学,2013,11(2):65-71.

[26]孟 鹏,李玉灵,张柏习,等. 沙地彩武松与樟子松苗木抗旱生理特性比较[J]. 林业科学,2010,46(12):56-63.