

基于主成分分析法的耕地整理潜力分析评价——以临川区为例

陈英,刘平辉* (东华理工大学地球科学学院,江西南昌 330013)

摘要 根据江西省抚州市临川区第二次调查更新数据和底图,结合土地整治规划和统计年鉴资料,建立了7个耕地潜力评价指标,运用SPSS软件进行主成分分析,计算特征值和累计贡献率,确定主成分个数;然后对主成分得分采用聚类分析方法计算,得到临川区耕地整理潜力等级,为临川区耕地整理提供科学依据。
关键词 耕地整理;潜力;评价;主成分分析
中图分类号 S28 **文献标识码** A **文章编号** 0517-6611(2014)32-11510-02

Analysis and Evaluation of Cultivated Land Consolidation Potential Based on Principal Component Analysis—A Case Study in Linchuan
CHEN Ying, LIU Ping-hui* (College of Earth Sciences, East China Institute of Technology, Nanchang, Jiangxi 30013)
Abstract According to the second survey update data and base maps in Linchuan Region, Fuzhou City of Jiangxi Province, combined with land remediation planning and statistical yearbook data, 7 potential evaluation indicators of cultivated land were established. Using SPSS software to conduct principal component analysis, calculate the eigenvalues and the cumulative contribution rate, determine the number of principal components. Then the principal component score was calculated using cluster analysis method to obtain Linchuan arable land readjustment potential level, which will provide a scientific basis for the Linchuan farmland consolidation.
Key words Arable land consolidation; Potential; Evaluation; Principle component analysis

随着经济全球化和城市化的快速发展,对不可再生的土地资源的需求量越来越大,同时对土地资源的过度开发和不当开垦造成了环境破坏,进一步加剧人与自然的矛盾。耕地整理是土地整理的重要内容之一,可以增加有效耕地面积,提升耕地整体质量,保障国家粮食安全,实现耕地总量动态平衡^[1]。

1 指标选取、数据来源与研究方法

1.1 指标选取与数据来源 根据2010年临川区第二次土地调查更新数据和统计年鉴、临川区土地整治规划(2011-2015)、临川区土地利用总体规划(2006-2020)及实际调研收集的数据和资料,构建了耕地整理潜力评价指标体系,包括新增耕地系数、耕地单产潜力指数、投入产出率、排水灌溉基础设施配套潜力、农民人均收入、农业人口、林地面积年增长量等7个评价指标^[2]。各指标原始数值见表1。

1.2 研究方法 运用主成分分析和聚类分析方法进行研究。耕地整理潜力评价指标体系是由多个指标组成,由于每个指标值具有纲量,数据差异大,为了让指标具有可比性,原始数据要经过标准化处理,即将指标值都转化为无量纲判别的标准值。可以运用标准化方法对各个指标进行标准化处理^[3],公式为:

$$H_{ij} = \frac{X_i - \bar{X}_j}{S}$$

其中: $\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_i$

$$S = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}$$

式中, $i,j=1,2,3,\cdots,n$ (乡镇个数); H_{ij} 为第j个指标中第i个

数据的标准化值; X_{ij} 为第i个乡镇的第j个指标的实际值; $\bar{X}_j$ 为第j个指标的平均值。

表1 临川区耕地整理潜力指标原始值

行政区	新增耕地系数 %	耕地单产提高量 kg/hm ²	投入产出率 /%	农业机械总动力 kW	农民人均纯收入年 增长率/%	农业人口 人	林地面积 年增长量 hm ²
唱凯镇	3.55	240.63	5.12	42 161.63	23.96	45 292	0.079
崇岗镇	2.38	382.15	5.83	34 614.29	9.28	27 071	7.213
大岗镇	3.37	305.17	5.69	57 670.45	13.27	27 907	11.231
东馆镇	2.37	633.19	8.55	37 699.63	11.23	18 678	6.812
抚北镇	1.95	143.52	2.93	7 237.13	13.24	6 721	0.587
高坪镇	2.34	578.56	7.93	51 045.83	19.05	33 747	11.584
河埠乡	2.84	470.23	7.11	32 060.71	13.94	11 698	8.424
湖南乡	3.08	157.24	3.88	46 242.52	17.55	42 858	2.911
连城乡	2.54	211.18	4.09	23 202.39	20.19	20 072	3.333
龙溪镇	1.72	312.23	4.60	32 199.79	17.44	17 471	13.100
罗湖镇	2.78	247.27	4.64	61 337.40	14.73	51 952	2.417
罗针镇	3.14	649.31	9.27	38 230.24	13.09	48 511	0.195
茅排乡	2.34	392.31	5.91	3 670.04	19.50	6 461	14.568
鹏田乡	2.23	702.49	9.20	23 269.99	19.72	12 194	7.153
七里岗乡	3.97	383.34	6.96	22 628.12	17.52	22 246	3.272
青泥镇	3.70	388.89	6.83	27 467.36	12.39	29 193	3.110
秋溪镇	1.99	386.89	5.60	43 731.44	16.34	25 701	5.062
荣山镇	2.27	348.83	5.39	46 489.77	16.25	22 760	21.525
上顿渡镇	2.49	536.19	7.58	54 755.67	16.93	59 176	1.861
嵩湖乡	2.69	311.59	5.28	38 728.60	14.99	23 485	3.775
太阳镇	2.60	324.91	5.36	23 256.47	17.32	21 919	2.420
腾桥镇	2.78	363.42	5.90	59 597.61	11.24	26 214	13.717
桐源乡	2.75	328.83	5.51	37 207.85	12.93	25 550	11.612
温泉镇	3.02	716.95	9.91	43 360.57	14.62	31 500	4.613
孝桥镇	2.16	98.33	2.59	23 248.75	18.26	25 429	0.045
云山镇	3.15	329.99	5.81	44 608.00	16.19	35 349	9.297
展坪乡	3.33	593.87	8.80	22 277.73	15.18	19 738	5.551

2 结果与分析

2.1 耕地整理潜力主成分分析 将标准化处理后的数据导入SPSS16.0软件,可以得出前4个主成分特征值分为 $\lambda_1 = 2.232, \lambda_2 = 1.801, \lambda_3 = 1.277, \lambda_4 = 0.891$,累计贡献率 $E = e_1$

+ $e_2 + e_3 + e_4 = 88.604\% > 85\%$, 说明前 4 个主成分就可以反映原指标信息的 88.604%, 所以只要选取前 4 个指标即可。依据以下公式可以得到临川区 27 个乡镇的 4 个主成分得分值(表 2)。相应公式为:

$$F_1 = 0.207X_1 + 0.361X_2 + 0.234X_3 + 0.39X_4 - 0.134X_5 + 0.223X_6 - 0.008X_7$$
$$F_2 = 0.225X_1 - 0.29X_2 + 0.273X_3 - 0.233X_4 + 0.116X_5 + 0.439X_6 - 0.292X_7$$
$$F_3 = -0.29X_1 - 0.129X_2 + 0.504X_3 - 0.185X_4 - 0.339X_5 + 0.088X_6 + 0.520X_7$$
$$F_4 = -0.328X_1 + 0.182X_2 + 0.178X_3 + 0.108X_4 + 0.905X_5 + 0.189X_6 + 0.287X_7$$

表 2 临川区各乡镇耕地整理潜力主成分得分

乡镇名称	F_1	F_2	F_3	F_4
唱凯镇	-0.153 8	2.067 2	-1.365 8	1.549 8
崇岗镇	0.022 7	-0.395 1	0.908 2	-1.618 4
大岗镇	0.407 7	0.529 6	1.229 3	-0.678 0
东馆镇	0.939 1	-1.292 9	0.262 9	-0.754 0
抚北镇	-2.169 7	-0.489 0	-0.529 7	-1.663 7
高坪镇	0.822 8	-0.368 7	0.586 2	1.977 5
河埠乡	0.148 0	-0.978 9	-0.079 3	-0.613 8
湖南乡	-0.477 6	1.782 4	0.127 8	0.052 8
连城乡	-1.377 5	0.333 3	-0.827 3	0.584 6
龙溪镇	-1.167 2	-0.787 7	0.999 4	1.039 1
罗湖镇	0.269 1	1.922 1	0.962 5	-0.128 5
罗针镇	1.856 0	0.337 6	-0.838 7	-0.551 8
茅排乡	-1.221 5	-1.734 3	-0.630 9	0.982 6
鹏田乡	0.479 6	-1.747 9	-1.164 8	1.551 7
七里岗乡	0.241 3	0.229 9	-1.708 2	-0.621 1
青泥镇	0.532 7	0.275 6	-0.830 5	-1.740 1
秋溪镇	-0.322 8	-0.041 5	0.485 8	0.536 8
荣山镇	-0.369 2	-0.784 9	2.071 2	1.165 1
上顿渡镇	1.308 7	1.182 1	0.147 1	1.123 5
嵩湖乡	-0.348 6	0.265 4	0.040 0	-0.510 0
太阳镇	-0.702 9	0.015 1	-0.841 6	-0.071 0
腾桥镇	0.436 9	-0.064 8	1.980 3	-0.687 3
桐源乡	-0.158 9	-0.228 3	0.912 3	-0.654 7
温泉镇	1.809 4	-0.568 5	-0.557 6	0.107 6
孝桥镇	-1.888 8	0.846 9	-0.480 4	-0.031 8
云山镇	0.206 7	0.602 9	0.425 6	0.140 1
展坪乡	0.877 7	-0.907 8	-1.283 9	-0.487 2

2.2 主成分得分的聚类分析 运用 SPSS16.0 软件进行聚类分析,将临川区 27 个乡镇按照 4 个主成分得分采用系统聚类分析方法进行分类。其中分析过程采用类平均法,选择欧氏距离,根据聚类谱系图得出临川区土地整治耕地潜力等级分布(图 1)和临川区 27 个乡镇的耕地整理潜力级别划分

结果(表 3)^[4]。

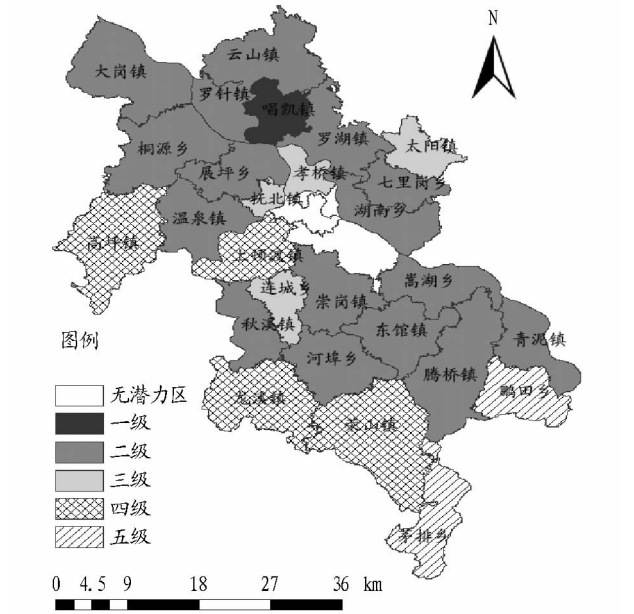


图 1 临川区耕地整理潜力等级分布

表 3 临川区耕地整理潜力级别划分结果

序号	潜力等级	乡镇名称
1	一级	唱凯镇
2	二级	崇岗镇 大岗镇 东馆镇 河埠乡 湖南乡 罗湖镇 罗针镇 七里岗乡 青泥镇 秋溪镇 嵩湖乡 腾桥镇 桐源乡 温泉镇 云山镇 展坪乡
3	三级	抚北镇 连城乡 太阳镇 孝桥镇
4	四级	高坪镇 龙溪镇 荣山镇 上顿渡镇
5	五级	茅排乡 鹏田乡

3 结语

主成分分析方法和聚类分析方法相结合,不仅完善了以往只按照“新增耕地系数”这一指标来计算耕地整理潜力大小的传统方法,而且能够更加科学客观地评价耕地整理潜力的等级。主成分分析法是通过线性变化,将原来多个指标组合成少数几个能反映总体信息的指标,然后对几个主成分的得分做聚类分析,与以往按照主成分总分值来确定潜力大小有很大区别,有利于减少人为干扰因素^[5]。

参考文献

[1] 刘黎明. 土地资源学[M]. 北京: 中国农业大学出版社, 2004.
[2] 唐秀美,潘瑜春,刘玉,等. 基于耕地系数和预评价法的耕地整治潜力测算方法[J]. 农业工程学报,2014,30(1):211-218.
[3] 徐建华. 现代地理学中的数学方法[M]. 北京:高等教育出版社,1994.
[4] 张仕超,魏朝富. 区域土地开发整理新增耕地的潜力及其贡献分析[J]. 农业工程学报,2010,26(S1):312-319.
[5] 朱恒峰. 大连旅顺口区耕地整理潜力分级方法探讨[J]. 国土资源导刊,2004(3):16-19.