

# 新疆非耕地系数相似性类型区划分

冯雪力<sup>1,2</sup>, 吴世新<sup>1</sup>, 陈 红<sup>1,2</sup>, 张良侠<sup>1,2</sup>

(1. 中国科学院新疆生态与地理研究所, 乌鲁木齐 830011; 2. 中国科学院研究生院, 北京 100049)

**摘 要:** 非耕地系数反映了田间耕地与非耕地的分布情况, 它的意义在于有助于获得较为准确的耕地面积。以影响非耕地系数的因素为指标对农区进行类型区划分是利用非耕地系数获取准确耕地面积的前提。首先以干旱区典型绿洲为研究区采用灰色关联方法探讨影响非耕地系数的主要因素, 使用层次分析法综合影响非耕地系数的主要因素, 并采用累计平方根法研究各个类型区分界点的确定。研究结果表明: 影响非耕地系数的主要因素有土地利用的斑块破碎度、水网密度、农业人口密度、人均农业生产值等。将它们作为类型区划分的指标, 并采用最优分层理论将类型区数目确定为 8。利用累计平方根法确定各类型区的分界点, 该方法使得分区结果较优, 各类型区内的方差较小。该类型区划分结果反映了人们对土地利用的强度和各类型区的非耕地系数逐渐增加的趋势。

**关 键 词:** 非耕地系数; 农业分区; 层次分析法; 累计平方根

土地是一个综合的自然地理概念。它是地表某一地段各种自然要素(地质、地貌、气候、水文、植被、土壤等)相互作用及人类活动影响在内的自然综合体, 是各种自然过程及人类活动最活跃的场所。耕地是土地的精华, 是与人类关系最为紧密的综合体<sup>[1-2]</sup>。随着经济的发展和人口的急剧增长, 耕地的数量不断下降。耕地的不足, 很可能危及地区或国家的粮食安全<sup>[3-5]</sup>。因此, 对耕地面积的统计监测显得尤为重要。新疆是我国最大的省份, 也是我国生态环境最为严酷和脆弱的地区<sup>[6-7]</sup>, 其绿洲农业的发展在干旱区具有十分重要的地位<sup>[8-11]</sup>, 因此对其耕地资源的统计也越来越重要。

获取耕地面积有多种方法。较为传统的方法是实地测量, 由此方法获得的耕地面积精度较高, 缺点是周期长、费用高。遥感监测方法是当前较为通用的方法, 它的时间周期较短, 缺点是由于遥感数据的空间分辨率、同物异谱以及异物同谱等问题, 致使耕地中的细小地物(耕地中的非耕地)被淹没在耕地之中。因此, 仅从遥感图像上无法将耕地中的非耕地识别、提取, 这样就影响了获得耕地面积的精度。

非耕地系数<sup>[12-13]</sup>是土地资源调查中耕地图斑内非耕地面积占图斑毛面积的比例, 通常以百分比表示, 它反映了田间耕地与非耕地分布的实际情况。

它存在的意义在于帮助人们获取较为准确的耕地面积, 通常结合遥感抽样方法<sup>[14-19]</sup>对遥感监测面积进行“纯化”。在利用非耕地系数获取精确的耕地面积时, 耕地图斑的尺度问题直接影响非耕地系数的大小, 进而影响“纯化”耕地面积的精度。因此, 选择耕地图斑的尺度时, 要使图斑尺度内的非耕地系数较为近似, 即图斑内的非耕地反映在某个指标上一般较为近似或非耕地的属性基本保持一致。

以影响非耕地系数的因素为指标对农区进行类型区划分是利用非耕地系数获取准确耕地面积的前提。影响非耕地系数的因素有地形、降雨量、水网密度、土地利用的斑块破碎度、土壤类型、土地利用类型、耕作制度、种植结构、复种指数、农业政策等。确定影响非耕地系数的主要因素后, 基于非耕地系数相似性的类型区划分方可进行。

由戴伦纽斯和霍奇斯(Hodges)于 1959 年提出的累积平方根法<sup>[20-21]</sup>是分层界限确定方法中较为著名的方法, 是最常用的并且效率较高的方法。本文首先以新疆典型绿洲农区为研究区, 采用灰色关联分析方法<sup>[22]</sup>探讨影响非耕地系数的主要因素, 然后采用层次分析法将各主要影响因素进行综合, 形成影响非耕地系数的综合因子并将其作为分区的指标。依据分区指标, 采用累计平方根法确定各个类型区的分层界限, 对全疆进行类型区划分。依据非

收稿日期: 2009-08; 修订日期: 2010-01.

基金项目: 中国科学院知识创新工程重大项目; 中国科学院知识创新工程重大项目(KZCX1-YW-09-10)。

作者简介: 冯雪力(1985-), 男, 河南鹿邑人, 硕士研究生, 主要从事地理信息系统与遥感应用研究。E-mail: fxluyi@163.com

耕地系数相似性的类型区划分的研究对新疆耕地面积遥感监测,获得较为准确的耕地面积具有重要的服务和指导意义。

## 1 研究区概况

新疆位于我国的西北部,土地面积约占全国总面积的 1/6,与蒙古、俄罗斯、哈萨克斯坦、吉尔吉斯斯坦、塔吉克斯坦、阿富汗、巴基斯坦、印度等国相邻,国境线长达 5400 km,是我国面积最大的省区,地域辽阔,地处欧亚大陆中心,山地盆地相间排列,形成三山夹两盆的地貌。全区地势起伏巨大,地貌类型复杂多样。由于四周高山环绕,形成了典型的温带大陆干旱性气候,具有明显差异的南北疆两大气候区。

新疆土地资源丰富,类型多,人均占有量高,是我国重要的农垦地区之一。但新疆土地资源中宜用地比重低、质量差、结构不平衡。新疆是一个农业大省,是全国粮食基地和特大商品棉基地,保持一定数量的耕地,利用和保护好有限的耕地资源,是当地经济和环境可持续发展的重要前提。

## 2 数据来源和研究方法

### 2.1 数据来源

使用 2008 年 Quick Bird 遥感影像和中巴资源卫星影像。Quick Bird 遥感影像的空间分辨率是 0.6m,使用该高分辨率的遥感影像代替实地样方,以缩短调查时间,节省人力、物力等资源。中巴资源卫星影像的空间分辨率是 19.5 m,在此数据尺度上确定采样框的大小。

处理平台采用 ArcGIS 9.2 和 Erdas 9.1 软件。数据预处理包括对该区遥感影像的镶嵌、投影变换,然后在 ArcGIS 9.2 中进行图像解译,并进行矢量化。

粮食产量和农业人口数量的数据通过 2007 年新疆统计年鉴获得。斑块破碎度、水网密度数据、耕地面积是在新疆土地利用数据库源数据的基础上计算得到的。

### 2.2 研究方法

影响非耕地系数的因素有很多,包括自然因素和社会经济因素两方面,自然因素包括地貌类型、海拔高度、土壤类型、气候、降雨量等,社会经济因

素包括农业人口比重、耕作制度、复种指数及农业政策等因素。其中有些因素是可以量化的,而有些难以量化。

首先采用灰色关联分析方法通过研究伊犁河流域的典型绿洲,研究影响非耕地系数的主要因素。其次,采用层次分析方法将影响非耕地系数的多个因素综合成反映非耕地系数的一个综合因子,以此作为类型区划分的指标,这样就大大降低了类型区划分工作的复杂度。

类型区划分之前要确定划分类型区的数目。1972 年,William G.Cochran<sup>[23]</sup>推导出分层抽样的测

量方差和分层层数的关系: 
$$V(\bar{y}_{st}) \geq \frac{s_y^2}{n} \left[ \frac{\rho^2}{L} (1 - \rho^2) \right]$$

式中: $\rho$  表示分层标志变量  $x$  和被测量的变量  $y$  之间的简单相关系数。并且有证明分层数小于 6 是较好的,当层数大于 6 时,方差的减少速度大为减缓。分层层数越多,抽样的费用就越大。采用此方法确定类型区划分的数目。

用累积平方根法确定各类型区的分界点是最常用的并且是效率较高的方法。其做法是在奈曼分配方式下,将分区变量  $X$  的分布密度函数  $f(x)$  的累积平方根进行等分来获得最佳分区。该方法的具体操作步骤如下:①自然分组。通常按  $X$  自身的特征进行等距分组。②根据所给数据画出  $f(x)$  频率分布表,其中包括  $f(x)$  的累积频率。③确定类型区间隔。④找出  $f(x)$  的累积频率的界点。⑤找出最优界点  $X_h$  (类型区的界点)。

## 3 研究结果

### 3.1 影响非耕地系数的主要因素

为确定影响非耕地系数的主要因素,选择干旱区典型绿洲—伊犁河谷的农区为研究对象,该典型区在地貌类型、海拔高度、土壤类型等方面的差异较小。因此,主要从社会经济及景观的角度选用指标,并进行量化,确定它们与非耕地系数的关系。选取的因素包括人均耕地、土地垦殖指数、斑块破碎度、水网密度等。在典型区确定 30 个样方,用于统计计算。

一般地,土地利用的斑块破碎度<sup>[24-25]</sup>和水网密度反映土地利用的景观状况,描述土地利用被分割的程度,在一定程度上揭示人类活动对土地利用的影响,其数值越大表示土地利用越趋于破碎化,即

斑块破碎度越大,耕地中的非耕地部分(如防护林、田间道路、水渠、田埂等)就越多。

由于多指标的量纲不统一,不便于统计分析,因此对数据进行归一化处理,以消除量纲影响。

采用如下公式计算灰色关联度,在 Matlab 软件中编写程序,计算结果如表 1 所示。

$\xi_i(k)=$

$$\frac{\min_i \min_k |A_0(k)-A_i(k)| + \rho \max_i \max_k |A_0(k)-A_i(k)|}{|A_0(k)-A_i(k)| + \rho \max_i \max_k |A_0(k)-A_i(k)|}$$
$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k)$$

在地形、土壤类型等因素一定的状况下,影响非耕地系数的首要因素是土地利用的斑块破碎度。斑块破碎度越大,耕地中的非耕地部分就越多。水网密度是影响非耕地系数的又一主要因素,它反映单位面积内水渠的多少,干旱区耕地主要依赖水渠灌溉。影响非耕地系数的第 3 个主要因素是土地垦殖指数。

根据农业专家的知识,影响非耕地系数的因素还包括:农业人口密度和人均农业产值。农业人口密度反映人口对耕地利用的需求程度,人均农业产值反映土地的经济功能—土地利用的强度。农业人口密度越大,人类对土地的开发利用强度将加大,致使土地利用的斑块更破碎,非耕地数量将增加。人均农业产值越大,种植的多样性随之增加,非耕地就增多。

3.2 影响非耕地系数的综合因子

选取对类型区划分的指标要恰到好处,过多不利于工作的开展,过少不能反映区划的内容。选取 4 个主要指标:斑块破碎度、水网密度、农业人口密度、人均农业产值。

为便于对类型区划分,采用层次分析法将 4 个指标综合成一个反映分区要素的综合指标。利用层次分析法,首先构造层次结构模型。其次,需要农业专家为各主要指标的重要程度进行打分,构建判断矩阵。最后,判断判断矩阵的一致性,得到各个影响因素对综合因子的权重,结果如表 2 所示。

根据上述结果,可得到综合因子的表达式为:0.4321×斑块破碎度+0.3201×水网密度 +0.1301×农业人口密度+0.1177×人均农业产值。

3.3 非耕地系数相似性类型区划分

根据相关研究,类型区数目为 6 时较优。综合

考虑研究区的实际情况以及应用分层抽样获取非耕地系数的费用,将总类型区数目控制在 8 之内。

初步将类型区数定为 6,然后根据类型区划分的实际情况对初步区划结果进行优化。以县或自治县为单位进行统计数据,采用累积平方根法确定各类型区之间的分界点。综合因子的均值是 0.288235。首先,把综合因子的数据从小到大排列,选择 0.288235 (绝对平均值) 的 1/6 即 0.288235/6=0.048 作为组距进行分组。具体结果如表 3 所示。

$\sqrt{f(y)}$  总的累加值是 143.35,按照等距将数据划分为 6 层,类型区的跨度是 23.89,分界累计值分别为:23.89,47.78,71.67,95.56,119.45。在上述表 3

表 1 各影响因素与非耕地系数之间的灰色关联度

Tab.1 The degrees of grey incidence between non-cultivated land coefficient and its affecting factors

非耕地系数	斑块破碎度	水网密度	土地垦殖指数	人均耕地
灰色关联度	0.98	0.97	0.94	0.48
关联次序	1	2	3	4

表 2 各影响因子对综合因子的权重

Tab.2 The coefficient of affecting factors to the composite factor

各影响因子	权重
斑块破碎度	0.4321
水网密度	0.3201
农业人口密度	0.1301
人均农业产值	0.1177

表 3 累计平方根法分层界限结果

Tab.3 The stratification thresholds based on accumulated  $\sqrt{f(y)}$

y(综合因子)	f(y)	f(y)累加	$\sqrt{f(y)}$ 累加	参考分层界限
0	0	0	0	
0.048	9	9	3	
0.096	11	20	7.47	
0.144	17	37	13.55	
0.192	9	46	20.33	23.89
0.240	14	60	28.076	
0.288	11	71	36.50	
0.336	2	73	45.046	47.78
0.384	1	74	53.65	
0.432	2	76	62.366	
.....	...	.....	.....	
0.576	1	77	71.14	71.67
0.624	1	78	79.97	
0.672	0	78	79.97	
0.72	1	79	88.86	95.56
.....	...	.....	.....	
0.96	1	80	97.8	
1.008	1	81	106.8	
.....	...	.....	.....	
1.344	1	82	115.86	119.45
.....	...	.....	.....	
1.536	1	83	124.97	
.....	.....	.....	.....	
2.112	1	84	134.135	
.....	.....	.....	.....	
3.504	1	85	143.35	
3.552				



中, 查找与界值最相近的值所对应的组, 则得到了以  $y$ (综合因子) 为标度的各类型区的分界值, 这 6 个类型区的分界值依次为 :0.192, 0.336, 0.576, 0.72, 1.344, 3.504。

依据上述计算数据, 可以得到各类型区的空间分布如图 1 所示, 各类型区的统计信息如表 4 所示。

从表 4 中可以看出, 第六类型区的方差(1.027)是最大的, 说明该区内部关于影响非耕地系数属性的差异较大。第六类型区包括喀什市、昌吉市、克拉玛依市。从统计的数据可以得出, 克拉玛依市的农业人口密度、人均农业总值均较大; 昌吉市的水网密度较大; 喀什市的斑块破碎度比其他的任何地方都大; 因此, 考虑将此三个市分别作为独立的类型区, 这样使得类型区内方差较小, 则非耕地系数较为近似。优化后的分区结果如图 2 所示。

类型区划分图显示出非耕地系数从小到大的变化趋势。由上述可知, 选用的指标斑块破碎度、水网密度、农业人口密度、农业人均生产总值对非耕地系数的影响是正向的。依据综合指标, 采用累计平方根法确定各个类型区的分界值, 然后按照综合指标的大小进行确定相似的类型区域。综合指标的差异反映了非耕地系数的不同。因此, 与第一类型区至第八类型区对应的非耕地系数是逐渐增加的。

4 结论与讨论

- (1) 本文通过灰色关联分析初步确定了影响非耕地系数的主要因素, 包括斑块破碎度、水网密度、农业人口密度、人均农业生产总值等, 并将它们作为研究基于非耕地系数相似性的新疆类型区划分的指标, 为进一步研究非耕地系数提供一定的参考。
- (2) 采用累计平方根法确定各个类型区的分界点简洁、快速, 该方法使得分区的结果较为合理, 使得非耕地在层内差异较小, 层间差异较大。同时, 各

表 4 各个类型区的统计信息

Tab.4 Descriptive statistics						
	样本数	最小值	最大值	层总值	层均值	层方差
	N	Minimum	Maximum	Sum	Mean	Variance
第一类型区	46	0.003	0.182	4.605	0.100	0.003
第二类型区	27	0.193	0.305	6.592	0.244	0.001
第三类型区	4	0.338	0.541	1.706	0.427	0.007
第四类型区	2	0.613	0.706	1.319	0.660	0.004
第五类型区	3	0.925	1.338	3.236	1.079	0.051
第六类型区	3	1.501	3.474	7.060	2.353	1.027

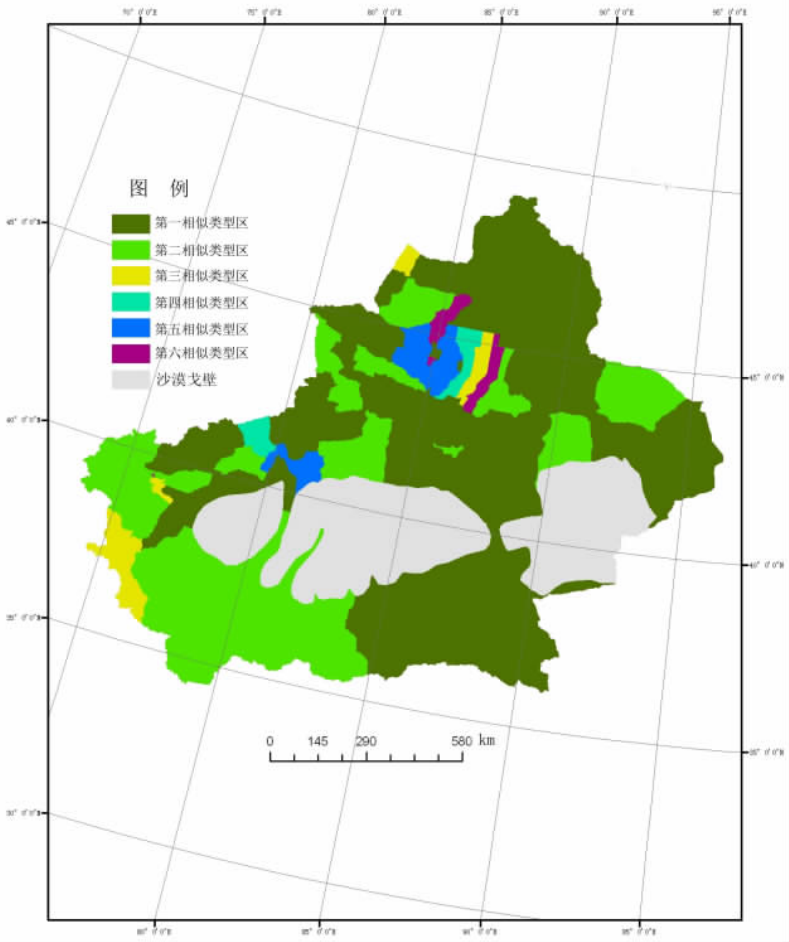


图 1 初步分区图

Fig.1 The result of the first division

- 个类型区的差异也反映了各区的农业发展水平、土地利用与开发的程度。依据该方法的机理可知, 综合指标对非耕地系数的影响反映在各个类型区上是逐渐增强的。因而, 可以预见各类型区在非耕地系数大小方面, 应该存在着逐渐上升的趋势。
- (3) 对基于非耕地系数相似性的类型区划分的研究探讨, 为“纯化”由遥感监测获得的耕地面积作了准备工作, 即解决了应用非耕地系数时的耕地图斑的尺度问题。

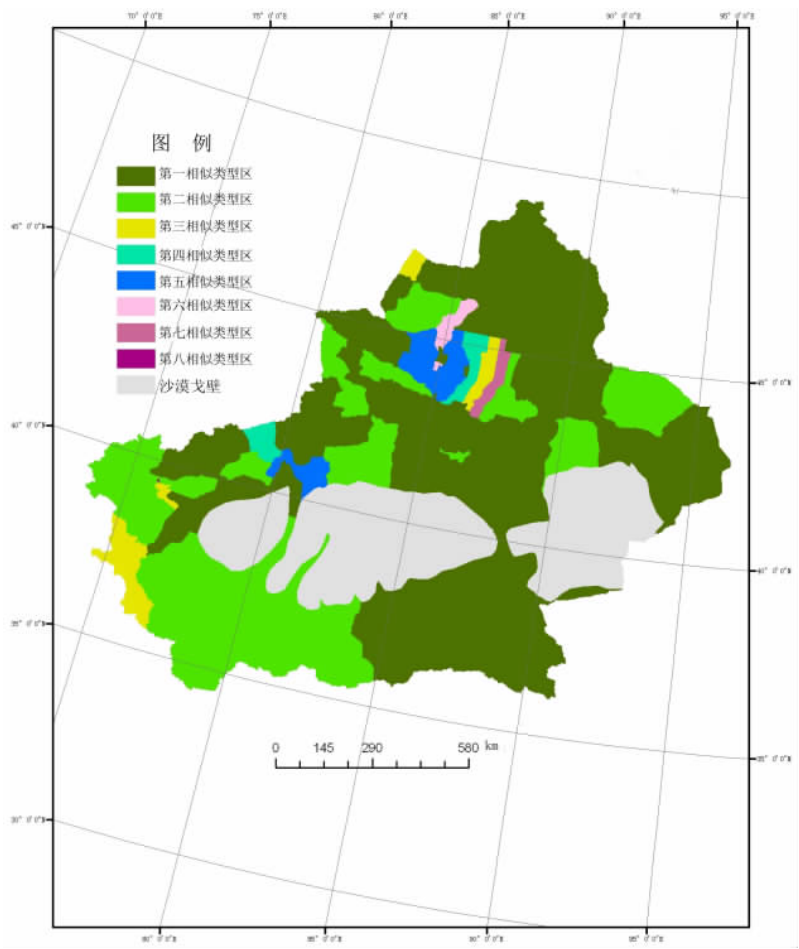


图 2 优化分区图  
Fig.2 The result of the last division

参考文献

[1] 王俊玲, 吴克宁, 吕巧玲. 县级耕地利用中存在的问题及对策研究. 国土资源科技管理, 2008, 25(1): 147-150.

[2] 赵文智, 程国栋. 人类土地利用的主要生态后果及其缓解对策. 中国沙漠, 2000, 20(4): 369-374.

[3] 刘旭华, 王劲峰, 刘纪远, 等. 国家尺度耕地变化驱动力的定量分析方法. 农业工程学报, 2005, 21(4): 56-60.

[4] 余振国, 胡小平. 我国的粮食安全与耕地的数量和质量研究. 地理与地理信息科学, 2003, 19(3): 45-49.

[5] 李茂. 河南省耕地和粮食灰色关联分析. 地理科学进展, 2002, 21(2): 163-172.

[6] 曹文栋, 郭青. 我国农村耕地利用及法律保护对策研究. 安徽农业科学, 2007, 35(4): 1134-1135.

[7] 李宪文, 林培. 国内外耕地利用与保护的理论基础及其进展. 地理科学进展, 2001, 20(4): 305-312.

[8] 陈东景, 徐中民. 干旱区农业生态经济系统的能值分析: 以黑河流域中游地区张掖地区为例. 冰川冻土, 2002, 24(4): 374-379.

[9] 黄青, 孙洪波, 王让会, 等. 干旱区典型山地—绿洲—荒漠系统中绿洲土地利用/覆盖变化对生态系统服务价值的影响. 中国沙漠, 2007, 27(1): 76-81.

[10] 王涛. 干旱区绿洲化、荒漠化研究的进展和趋势. 中国沙漠, 2009, 29(1): 1-9.

[11] 吕新, 杨磊, 张凤华, 等. 荒漠绿洲区农业特征及其可持续发展战略. 中国沙漠, 2005, 25(4): 599-603.

[12] 石承苍, 雍国玮, 周华茂, 等. 耕地系数测算方法研究与应用推广. 西南农业学报, 1993, 6(2): 44-48.

[13] 刘纪远. 彩色红外航空遥感技术在西藏农田非耕地系数测算中的应用研究与实践. 遥感学报, 1990, 5(1): 27-37.

[14] 陈仲新, 刘海启, 周清波, 等. 全国冬小麦面积变化遥感监测抽样外推方法的研究. 农业工程学报, 2000, 16(5): 126-129.

[15] 焦险峰, 杨邦杰, 裴志远. 基于分层抽样的中国水稻种植面积遥感调查方法研究. 农业工程学报, 2006, 22(5): 105-110.

[16] 李强子, 吴炳方. 作物种植成数的遥感监测精度评价. 遥感学报, 2004, 8(6): 581-587.

[17] 焦险峰, 杨邦杰, 裴志远. 全国棉花种植面积遥感监测抽样方法设计. 农业工程学报, 2002, 18(4): 159-162.

[18] 曹志东, 王劲峰, 李连发, 等. 地理空间中不同分层抽样方式的分层效率与优化策略. 地理科学进展, 2008, 27(3): 152-160.

[19] 姜成晟, 王劲峰, 曹志东. 地理空间抽样理论研究综述. 地理学报, 2009, 64(3): 368-380.

[20] 吴建华, 张颖. “累积平方根法”最优解严格的数学证明及应用. 济南大学学报: 自然科学版, 2006, 20(3): 267-271.

[21] 冯士雍, 倪加勋, 邹国华. 抽样调查理论与方法, 北京: 中国统计出版社, 1998.

[22] 梁国业, 廖健平. 数学建模. 北京: 冶金工业出版社, 2004: 143-146.

[23] 刘卫江, 白磊, 景泉, 等. 改进型分层抽样技术及性能研究. 计算机工程与应用, 2007, 43(8): 114-117.

[24] 鲁春阳, 齐磊钢, 桑超杰. 土地利用变化的数学模型解析. 资源开发与市场, 2007, 23(1): 25-28.

[25] 王思远, 张增祥, 周全斌, 等. 近 10 年中国土地利用格局及其演变. 地理学报, 2002, 57(5): 523-530.

## Study on Type Division Based on Non-cultivated Land Coefficient Similitude in Xinjiang

FENG Xueli<sup>1,2</sup>, WU Shixin<sup>1</sup>, CHEN Hong<sup>1,2</sup>, ZHANG Liangxia<sup>1,2</sup>

(1. Xinjiang Institute of Ecology and Geography, CAS, Urumqi 830011, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Non-cultivated land in cultivated areas is the key factor affecting monitoring cultivated land area using remote sensing technique on a large scale. In this paper, the conception of non-cultivated land coefficient was proposed to express the percentage of non-cultivated land in cultivated areas. Non-cultivated land coefficient shows the effects of human activities on the cultivated land. In addition, Precision of surveying results could be improved with non-cultivated land coefficient and the cultivated land area by remote sensing. Taking the typical agriculture area in the oasis in Yining County in Xinjiang as a study area, the paper designed 30 samples based on the Quick Bird data in 2008. Based on the platform of ArcGIS, spatial database was built to store the sampling data and calculate the non-cultivated land coefficient. It is time-saving to use the Quick Bird data to substitute for filed measurements. Firstly, the paper used grey correlation method to study the relationship between non-cultivated land coefficient and its influencing factors. Simplification of these factors was done by analytic hierarchy process. Based on the simplified factors, differences between type zones can be found to complete the agriculture division using accumulated  $\sqrt{f(y)}$  method. The results are as follows: the main factors influencing non-cultivated land coefficient were land use fragmentation, density of water network, agricultural population density and per-capita cultivated land. These factors were used as agricultural division index. In the paper, 8 zones are divided using the theory of optimal allocation. Finally, the results of agricultural division in the paper can meet the requirements of zone division that there is little difference between the samples in one division zone and quite obvious between two different division zones. Using accumulated method to find the thresholds of different type zones was very effective and convenient. The agricultural division, which reflects interference degree of human activities to land, is useful in guiding land use and land consolidation. Study on agricultural division based on non-cultivated land coefficient similitude in Xinjiang is quite useful for improving the cultivated areas by remote sensing.

**Key words:** non-cultivated land coefficient; agricultural division; analytic hierarchy process;

accumulated  $\sqrt{f(y)}$

本文引用格式:

冯雪力, 吴世新, 陈红, 等. 新疆非耕地系数相似性类型区划分. 地理科学进展, 2010, 29(3): 301-306.