

空间分层抽样在森林覆盖面积监测中的应用

张小伟^{1,2}, 余光辉^{2*}, 温小荣², 林国忠², 曹林², 罗扬³

(1. 浙江省森林资源监测中心, 浙江 杭州 310020; 2. 南京林业大学森林资源与环境学院, 江苏 南京 210037;
3. 贵州省林业科学院, 贵州 贵阳 550000)

摘要:为探讨空间分层抽样在森林资源调查中的应用,以对贵阳市花溪区的森林覆盖面积调查为例,利用融合后遥感数据 SPOT5 作为底图,布设 100 m × 100 m 的抽样框,并从 SPOT5 影像上提取辅助因子进行先验抽样。采用 Moran's I 指数进行空间自相关分析,结果表明,调查目标具有显著空间自相关性,可以进行空间抽样。利用 460 个样本单元的观测值对整个研究区进行普通克里格空间插值,并以插值结果作为分层辅助因子进行空间分层,结果初步表明空间分层抽样与经典抽样和随机空间抽样相比具有明显优越性,可以降低成本,提高效率。

关键词:空间分层抽样;森林覆盖面积;Moran's I 指数;普通克里格

中图分类号:S757 **文献标志码:**A **文章编号:**1000-2006(2012)03-0081-04

The application of spatial stratified sampling in remote sensing monitoring of forest cover

ZHANG Xiaowei^{1,2}, SHE Guanghui^{2*}, WEN Xiaorong², LIN Guozhong², CAO Lin², LUO Yang³

(1. Monitoring Center of Forest Resources of Zhejiang Province, Hangzhou 310020, China; 2. College of Forest Resources and Environment, Nanjing Forestry University, Nanjing 210037, China; 3. Forestry Science Institute of Guizhou Province, Guiyang 550000, China)

Abstract: In this paper, we selected the forest cover per hectare in Huaxi district, Guiyang city as an example to explore the application of spatial stratified sampling for forest resources. By using the information from SPOT5 image as foundation, a 100 m × 100 m sampling frame was constructed, the information extracted from SPOT5 image were used to make priori sampling. The results showed that there is significant spatial autocorrelation of the forest cover per hectare from Moran's I. The result of ordinary Kriging interpolation is used as the auxiliary factor to stratify. Our research initially show that spatial stratified sampling is more efficient compared with classic sampling or spatial random sampling.

Key words: spatial stratified sampling; forest cover; Moran's I; ordinary Kriging

从 20 世纪 30 年代开始,随机抽样的概念和原则得到了确立,系统的抽样理论和方法开始形成和发展。经过 1 个多世纪的发展,抽样技术已经相当成熟和完善^[1]。森林资源作为一种可再生资源,其分布面广、自然地理环境复杂并且多数林区交通不便。正是由于森林资源的这些特征,抽样技术在森林资源调查中被广泛应用^[2-6]。经典抽样技术主要调查研究“属性”的特征,一般不涉及“属性”

的空间特征,经典抽样技术假设调查目标彼此独立。但是,在森林生长发育过程中,由于森林会受到环境条件、自然干扰、人为干扰、更新树种的生理生态学特性、现存树种与更新树种的关系、竞争植物种和其他生物种的特性等因素及其相互作用的影响,往往使森林具有显著的空间自相关性。因此,经典抽样技术中样本间彼此独立的假设很难成立,而空间抽样技术的原理更加切合实际。在同一

收稿日期:2011-04-19 修回日期:2012-03-12

基金项目:国家自然科学基金项目(30571491);国家林业公益性行业科研专项项目(200804006);江苏省林业三项工程项目(lyxs(2009)46)

第一作者:张小伟,硕士。* 通信作者:余光辉,教授。E-mail: ghsh@njfu.edu.cn。

引文格式:张小伟,余光辉,温小荣,等. 空间分层抽样在森林覆盖面积监测中的应用[J]. 南京林业大学学报:自然科学版,2012,36(3):81-84.

精度和置信区间内,空间抽样技术比传统抽样技术所需的样本容量更小,能够达到降低抽样成本,提高抽样效率的效果。

1 空间分层抽样

1.1 空间分层抽样概念及原理

空间统计学又称地统计学,任何事物之间都有相关性,相距近的事物比相距远的事物之间更加相关,对这种相关性的研究和量化构成了空间统计理论的基础,许多学者对空间相关性和空间变异等问题做了大量的研究^[7-12],使得基于样本不独立假设的空间抽样调查技术得以迅速发展。

空间抽样简而言之是在考虑调查目标空间相关性的基础上所进行抽样的调查。空间抽样技术与经典抽样技术最大的区别在于认为样本单元不是相互独立的,而是具有相关性的。空间分层抽样就是将总体单位按其属性特征和空间连续性分成若干类型或层,然后在各类型或层中独立抽取样本单元的一种抽样方法。

传统分层抽样中,样本点没有空间坐标信息,根据分层标准,将层内方差小、层间方差大、分层属性值近似的分到同一层。根据这个分层标准对空间对象分层时,分层结果在空间上是离散的。各层分布散布在整个研究区域中。根据 Tobler 第一定律^[13],距离越近的对象,对象的相似程度越高。如果按照传统的分层抽样方法,同一层的对象在空间上可能相距很远,甚至各层混杂分布。因此空间分层抽样除了要求层内方差小、层间方差大以外,还要求兼顾同一个层的对象能够在空间上积聚。

在空间分层抽样中,空间研究区域 A 分为 L 层,首先按照分层随机模型中计算总样本量的方法计算研究区域总样本量,其次将样本量按权重 W_h 分配到各个层内,可以平均分配、按各层单元数分配、按各层离散方差和总体规模乘积分配,等等,其抽样有效性按此顺序提高。权重可以依据抽样有效性和参数可获取性选择;最后在各层内部进行独立随机抽样。经过野外调查后,计算各层样本均值和方差。

第 h 层的均值为: $\overline{Z}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} Z(x_{hi})$ 。 (1)

第 h 层的均值方差为: $V(\overline{Z}_h) = \text{var}\{\overline{Z}_h - Z(h)\} = \frac{1}{n} [\sigma^2 - E\{C_h(X,Y)\}]$ 。 (2)

在得到各个层的均值和方差后,研究区的均值和方差为:

$\overline{Z} = \sum_{h=1}^L W_h \overline{Z}_h$; (3)

$V(\overline{Z}) = \sum_{h=1}^L W_h^2 V(\overline{Z}_h)$ 。 (4)

1.2 空间自相关分析

空间自相关实际上是一种地理特性,因此空间数据中它几乎总是存在。Tobler 将其称为地理学第一定律^[13],即所有的事物都是相互联系的,但是离得越近的事物,彼此之间的联系越强。空间自相关是进行空间抽样的前提条件。笔者采用 Moran's I 来进行空间自相关分析^[14]。

Moran's I 可以简单定义如下:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sigma^2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}} = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} S_{ij}}{\sigma^2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}}$$
 (5)

I 指数的取值在 -1 到 1 之间,其中 -1 表示极强的负空间自相关,1 表示极强的正空间自相关。如果不存在空间自相关,则随机情况下的 Moran's I 为:

$E_I = \frac{-1}{n-1}$ 。 (6)

I 指数作为空间自相关的量化指标,要对其进行显著性检验。其标准化 Z 值为:

$Z = \frac{I - E_I}{\sqrt{\text{VAR}(I)}}$ 。 (7)

1.3 空间分层因子

在空间分层抽样中如何确定分层因子,是空间分层抽样的难点和重点,与随机分层抽样不同,空间分层抽样的每一层在空间上必须是连续的。在此选用普通克里格空间插值方法对整个研究区的每公顷森林覆盖面积进行插值^[15],以插值结果作为分层辅助因子进行空间分层。

使用克里格插值的基本条件是调查目标具备空间自相关性。普通克里格作为克里格大家族中应用最广泛的一种方法,它既可以对点进行估值,也可以对块进行估值。 x_0 处的估值可表示为:

$Z(x_0)B = Y'W^{-1}B$; (8)

$W =$

$$\begin{bmatrix} \text{cov}(x_1, x_1) & \text{cov}(x_1, x_2) & \cdots & \text{cov}(x_1, x_k) & 1 \\ \text{cov}(x_2, x_1) & \text{cov}(x_2, x_2) & & \text{cov}(x_2, x_k) & 1 \\ \vdots & & \ddots & \vdots & \\ \text{cov}(x_k, x_1) & \text{cov}(x_k, x_2) & \cdots & \text{cov}(x_k, x_k) & 1 \\ 1 & 1 & & 1 & 0 \end{bmatrix}$$

其中: $\text{cov}(x_i, x_j)$ 表示点 x_i, x_j 之间的变异量,因为变异量只与两点之间的相对距离有关,所以 W 是

对称矩阵。 B 和 Y 分别为

$$B = \begin{bmatrix} \text{cov}(x_0, x_1) \\ \text{cov}(x_0, x_2) \\ \vdots \\ \text{cov}(x_0, x_k) \\ 1 \end{bmatrix}, Y = \begin{bmatrix} Z(x_1) \\ Z(x_2) \\ \vdots \\ Z(x_k) \\ 0 \end{bmatrix}。$$

2 空间分层抽样在花溪区森林覆盖面积调查中的应用

2.1 研究区概况及数据来源

研究区设在贵阳市花溪区 (106° 07' ~ 107°17 'E, 26°11' ~ 27°22'N)。北临南明区、乌当区,东与龙里县相邻,南接长顺县、惠水县,西与清镇市、平坝县接壤,距贵阳市中心 17 km。全区地貌以丘陵和山地为主,土地总面积为 971.5 km²。全区森林面积 3.6 万 hm²,森林覆盖率 36.03 %,生物多样性丰富。常年气温温差不大,年均气温 14.9 ℃,年均降水量 1 100 ~ 1 200 mm。

所使用的遥感数据为 SPOT5 HRG 数据。全色

分辨率提高到 2.5 m,多光谱达到 10 m,幅宽为 60 km × 60 km。研究中还用到 1:10 000 的地形图,用来配准遥感图像和切割研究区。SPOT5 影像经校正后,多光谱影像和全色影像融合。融合后图像能较清晰将森林和其他地类区分开来。

2.2 花溪区森林覆盖面积抽样调查

以贵阳市花溪区为研究区,采用预处理过的遥感数据 SPOT5 的提取森林

覆盖面积信息作为辅助因子进行先验抽样。采用 100 m × 100 m 的正方形编制抽样框,花溪区共分为 97 156 个样本单元。预抽样随机抽取 150 个样本单元,得到其每公顷森林覆盖面积为 4 046.6 m²,总体标准差估计为 3 620,总体均值方差估计为 87 360。由 460 样本单元观测数据计算空间自相关 I 指数为 0.19, Z 值为 10.37,森林覆盖面积具有明显自相关。460 样本单元普通克里格插值结果见图 1A,依据插值结果的分层效果见图 1B,所抽取的 460 个样本单元分布见图 1C。

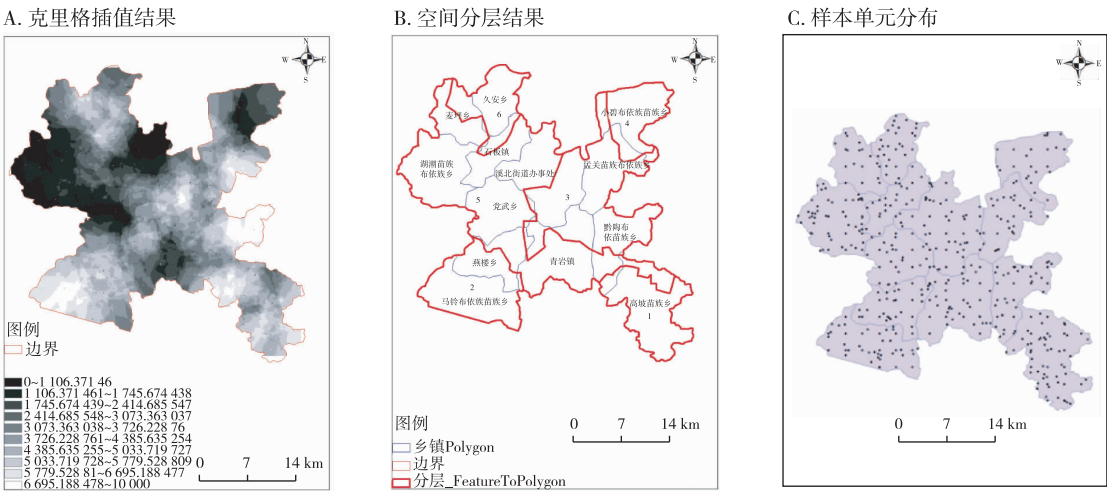


图1 样本单元分布及处理结果

Fig.1 The distribution of sample units and the results of treatment

在可靠性 95 % 和精度 85 % 的水平下,分别进行简单随机抽样、空间随机抽样和空间分层抽样。

简单随机抽样的样本容量为: $n = \frac{Z^2_{1-\alpha} \sigma^2}{d^2} (1 + 10 \%) = 151。$

简单随机抽取 151 个样本单元,调查样本单元并对总体进行估计,均值为 3 667.7 m²,相对误差为 15.3 %。

空间随机抽样的样本容量为:

$$n = \frac{1}{V} (\sigma^2 - R) = n_{\text{classic}} [1 - r] = 123。$$

在研究区随机抽取 123 个样本单元,调查样本单元并对总体进行估计,均值为 3 797.1 m²,相对误差为 18 %。

空间分层抽样的样本容量为: $n = \frac{\sum_{h=1}^1 (W_h S_h)^2}{\sigma^2_{y_{st}} + \frac{1}{N} \sum_{h=1}^1 W_h S_h^2} (1 + 10 \%) = 104。$

如图 1B 所示,研究区共分为 6 层,按照样本最优分配原则,各层样本数量为: $n_1 = 13, n_2 = 15, n_3 = 30, n_4 = 7, n_5 = 30, n_6 = 9。$

在各层中独立抽取样本,调查样本单元并对总

体进行估计,每公顷森林覆盖面积均值为 3 871.5m²,相对误差为 14.96 %。

由此可见,在同一可靠性和精度条件下,空间分层抽样样本容量最小,空间随机抽样次之,简单随机抽样样本容量最大。

3 结 论

(1)花溪区森林覆盖面积具有明显空间自相关性。通过 Moran's I 指数量化花溪区每公顷森林覆盖面积的空间自相关性,其空间自相关系数为 0.19,显著性检验指标 Z 值为 10.37 (远大于 2.58),所以花溪区森林覆盖面积的空间自相关性是显著的。

(2)依据抽样原理,空间分层抽样优于经典抽样、空间随机抽样,而空间随机抽样优于经典抽样。空间抽样方法量化样本单元的空间自相关性并用以缩小样本容量。空间分层抽样进一步考虑到相似样本单元丛聚性,使空间上连续的各层中方差尽量小,并在层间独立抽样。

(3)依据抽样成本,空间分层抽样成本最低、效率最高。抽样调查的样本容量越大,成本越高。经典抽样不考虑空间自相关性,相当于一部分样本被重复抽取,造成样本冗余,扩大成本。简单随机抽样样本容量 151,而空间随机抽样的样本容量为 123,空间分层抽样样本容量 104。

(4)空间分层抽样在森林资源动态监测中具有良好的应用前景。动态监测过程可以为空间分层抽样提供先验知识,空间分层抽样可以降低动态监测成本、提高动态监测效率。

参考文献:

[1] 李金昌. 应用抽样技术[M]. 北京:科学出版社,2006.

[2] 廖桂宗,彭世揆. 试验设计与抽样技术[M]. 北京:中国林业出版社,1993.

[3] 贾云奇. 系统抽样法在更新造林成活率调查中的应用[J]. 东北林业大学学报,1989,17(2):84-87.

[4] 于峰,张彬,代启光. 简述系统抽样在三类调查中的应用[J]. 林业勘查设计,2003(2):41.

[5] 向小芹. 分层抽样法在退耕还林工程水土流失监测中的应用[J]. 陕西林业科技,2006(2):9-11.

[6] 肖银松. “3S”及抽样技术在森林资源动态监测中的应用[J]. 西南林学院学报,2004,24(2):60-64.

[7] 王劲峰,姜成晟,李连发,等. 空间抽样及统计推断[M]. 北京:科学出版社,2009.

[8] Fisher M, Scholten H J, Unwin D. Spatial Analytical Perspectives on GIS [M]. London: Taylor & Francis, 1996.

[9] Anselin L. Spatial Econometrics: Methods and Models [M]. Dordrecht: Kluwer Academic, 1988.

[10] Haining R P. Spatial Data Analysis: Theory and Practice [M]. Cambridge: Cambridge University, 2003.

[11] Ripley B D. Spatial Statistics [M]. New York: John Wiley & Sons Inc,1981.

[12] Griffith D A. Spatial Autocorrelation and Spatial Filtering[M]. NewYork:Springer, 2003.

[13] Tobler W. A computer movie simulating urban growth in the Detroit region[J]. Economic Geography,1970,46:234-240.

[14] David W S, Wong Jay Lee. Statistics Analysis of Geographic Information with Arc View GIS and ArcGIS [M]. Beijing: China Financial & Economic Publishing House,2008.

[15] 苏姝,林爱文,刘庆华. 普通 Kriging 法在空间内插中的运用[J]. 江南大学学报,2004,3(1):18-21.

(责任编辑 郑瑛燧)