

文章编号: 1007-4619(2006)02-0256-07

聚类分析在监测区域样地分类中的应用研究

李崇贵¹, 赵宪文²

(1. 深圳职业技术学院 3S 研究中心, 广东 深圳 518055; 2. 北京林业大学 资源与环境学院, 北京 100083)

摘 要: 为建立监测区域以遥感和 GIS 为基础的森林郁闭度及蓄积定量估测方程, 需在监测区域抽取一定数量有代表性的样地。如何有效抽取有代表性的样地属多目标优化问题。若根据监测区域已知样地的数量, 采用所有可能的组合方法, 按一定优化标准进行抽样, 计算工作量将非常庞大, 实际工作中很难进行。为此需先对监测区域的样地进行分类, 在分类的基础上选择一定数量有代表性的样地。因度量样地之间亲疏程度的统计量及类和类之间相似性测度指标有多种, 因此对同一监测区域, 不同的统计量及类和类之间相似性测度指标, 将得到不同的分类结果。设置不同的影响郁闭度及蓄积估测的遥感和 GIS 因子, 分类结果也会出现较大的差异。本文针对特定监测区域, 采用计算机仿真方法, 系统研究了不同影响因子对分类的影响规律, 实际工作中应如何选择度量样地间亲疏程度的统计量及类和类之间相似性测度指标, 所得结果可用于指导生产实践。

关键词: 聚类分析; 亲疏统计量; 相关系数**中图分类号:** P208/TP79**文献标识码:** A

An Application Research of Cluster Analysis on Sample Plot Classification in Monitoring Area

LI Chong-gui¹, ZHAO Xian-wen²

(1. 3S Research Center Shenzhen Polytechnic Guangdong, Shenzhen 518055, China;

2. Resource and Environment Institute Beijing Forestry University, Beijing 100083, China)

Abstract In order to establish the forest canopy density and stock volume estimation equation based on remote sensing and GIS in the monitoring area, it's needed to sample certain amounts of representative sample plots. How to rationally select certain amounts of representative sample plots belongs to the problem of multi-objective optimization. It's hard to do in practical work because of the heavy calculation workload from the selecting by certain optimizing standards with all the combination method according to the known amounts of sample plots in monitoring area. Therefore, to classify the sample plots firstly and then select certain representative ones is needed. Because of the manifold of statistic measuring distance between sample plots and concrete clustering method, different categorized results appeared in the same monitoring area result in different statistic and clustering methods. Designing different factors of remote sensing and GIS that influencing the estimation of canopy density and stock volume, there will be a large difference in classifying results. To a specific monitoring area, the influencing law of different factors to classifying is studied systematically by means of computer simulation. Meanwhile, how to choose the statistic measuring distance between sample plots and clustering methods in practical work is also studied in this paper. The results can be useful to real work.

Key words: cluster analysis; resemble extent statistical; correlation coefficient

收稿日期: 2004-09-30; 修订日期: 2005-01-17

基金项目: 国家自然科学基金项目 (30371159) 资助。

作者简介: 李崇贵 (1966—), 男, 副教授, 2001 年 7 月获中国林业科学研究院资源信息研究所博士学位。主要从事遥感、GIS 和 GPS 在

森林资源和土地资源等领域的应用研究和产业开发, 已发表论文 20 多篇, 专著 1 部。

1 引 言

为建立监测区域以像元为单位的森林郁闭度及蓄积估测方程,需在监测区域选择一定数量有代表性的样地,有代表性的样地是指能代表监测区域森林种类及分布状况的样地。为有效选择监测区域有代表性的样地,需根据一定的数学方法,对监测区域样地进行分类,然后再从这些类型的样地中选择参与建立估测方程的样地,使参与建立估测方程的样地尽可能代表监测区域森林资源的分布状况,以提高估测方程的稳健性及可靠性。

在确定参与建立郁闭度及蓄积估测方程的样地时,除考虑样地的种类外,还需考虑每种类型样地的数量。在考虑每种类型样地的数量时,应根据每种类型样地的数量占监测区域样地总数的权重来确定,不能按随机方式抽取。若按随机方式抽取样地,被抽取的样地不一定能代表监测区域森林资源的分布状况。为确定参与建立估测方程样地的种类及每种类型样地的数量,必须先对整个监测区域的样地进行分类。若根据监测区域已知样地的数量,在不事先分类的情况下,采用所有可能的组合方法,按某种标准抽取建立估测方程的样地,计算工作量将非常庞大,在实际工作中很难进行。

对监测区域样地进行分类时,根据样地的多项调查指标和对应的遥感信息,定义能度量样地间相似程度(亲疏关系)的统计量,求出各样地间相似程度的度量值,按一定的类和类之间相似性测度指标对样地进行逐一归类。因有多种衡量样地间相似程度的统计量,且类和类之间相似性测度指标又有多种,对特定的监测区域,不同的亲疏关系统计量及类和类之间相似性测度指标的组合,将得到不同的分类结果。哪种分类结果能实际代表监测区域样地的种类,实际工作中应如何选择度量样地间亲疏关系的统计量及类和类之间相似性测度指标,值得分析研究。设置不同影响郁闭度及蓄积估测的遥感和 GIS 因子,其分类结果也会有较大差异,在分类时如何设置刻画样地的遥感和 GIS 因子,本文对此也进行了研究与分析。

2 基本数学原理

设监测区域有 n 个样地,每个样地可用 p 个影响森林郁闭度及蓄积估测的遥感和 GIS 因子描述。

其中 GIS 因子来自地面样地调查,主要包括海拔、地类(有林地、无林地)、坡向(阴坡、阳坡、无坡向)、优势树种等;遥感因子是地面样地在配准后遥感图像上对应像元各波段的灰度值及灰度比值,可将样地展绘在配准后的遥感图上,通过编程自动提取。系统聚类是首先将监测区域内的样地各自看成一类,然后根据样地的亲疏程度,将亲密程度最高的两个样地合并,考虑合并后的类与其他类之间的亲疏程度,再进行合并。重复这一过程直到分类结束。

2.1 样地间亲疏程度的测度

因监测区域每个样地可用其影响郁闭度及蓄积估测的 p 个因子来描述,每个样地可看成是 p 维欧氏空间 E^p 中的一点。可用距离测度样地之间的相似程度,距离越近,样地间亲密程度就越高。设监测区域 n 个样地 p 个因子组成的观测矩阵可表示为^[1-4]:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

对任意两个样地 x_i 和 x_j 之间的距离,需满足下列 3 个条件:

$$(1) d(x_i, x_j) \geq 0; \text{ 且仅当 } x_i = x_j \text{ 时, } d(x_i, x_j) = 0;$$

$$(2) d(x_i, x_j) = d(x_j, x_i);$$

$$(3) d(x_i, x_j) \leq d(x_i, x_k) + d(x_k, x_j).$$

在对监测区域的样地进行分类时,根据已有研究文献,本文主要考虑以下几种衡量样地间亲疏程度的距离指标^[1-3]。

2.1.1 欧氏距离

$$d_{ij}^2 = (x_i - x_j)(x_i - x_j)' \quad (2)$$

2.1.2 标准化欧氏距离

$$d_{ij}^2 = (x_i - x_j)D^{-1}(x_i - x_j)' \quad (3)$$

设:

$$\bar{x} = [\bar{x}_1 \quad \bar{x}_2 \quad \cdots \quad \bar{x}_p] \\ = \left[\frac{1}{n} \sum_{k=1}^n x_{k1} \quad \frac{1}{n} \sum_{k=1}^n x_{k2} \quad \cdots \quad \frac{1}{n} \sum_{k=1}^n x_{kp} \right]$$

为各样地 p 个因子的均值,令 $s_t =$

$$\sqrt{\frac{1}{n} \sum_{k=1}^n (x_{kt} - \bar{x}_t)^2}$$

为 n 个样地第 t 因子的标准差,

$$D^{-1} = \begin{bmatrix} s_1^{-1} & 0 & \cdots & 0 \\ 0 & s_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_p^{-1} \end{bmatrix}$$

为 n 个样地 p 个因子标

准差形成矩阵的逆阵。

2.1.3 马氏距离

$$d_{ij}^2 = (x_i - x_j) S^{-1} (x_i - x_j)' \quad (4)$$

式中, S 为样本协方差矩阵, 将式 (1) 中的样地观测矩阵中心化, 可得中心化后的矩阵

$$X = X - \bar{x} = \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \cdots & x_{np} - \bar{x}_p \end{bmatrix} \quad (5)$$

根据协方差传播律, 可得 n 个样地的样本协方差矩阵为:

$$S = \frac{1}{n} X'X \quad (6)$$

2.1.4 绝对距离

$$d_{ij} = \sum_{k=1}^p |x_{ki} - x_{kj}| \quad (7)$$

2.1.5 明可斯基 (Minkowski) 距离^[3]

$$d_{ij} = \left\{ \sum_{k=1}^p |x_{ki} - x_{kj}|^q \right\}^{\frac{1}{q}} \quad (8)$$

在上面所列的距离中, 绝对距离和欧氏距离分别是 Minkowski 距离在 $q=1$ 和 2 时的特殊情况。在实际中欧氏距离、布洛克距离及明可斯基距离应用较多, 当应用于森林资源监测时, 其主要缺点是它与描述样地的遥感和 GIS 因子的量纲有关, 有一定的人为性; 没有考虑各遥感和 GIS 因子间可能存在的相关性。马氏距离满足对一切线性变换的不变性, 不受各遥感和 GIS 因子量纲的影响, 但它夸大了变化微小因子的作用。在对不同监测区域样地进行分类时, 因描述样地的遥感和 GIS 因子的种类及取值可能存在差异, 如何合理选择衡量样地间亲疏程度的指标, 现有研究文献没有说明, 本文将采用仿真方式确定。

2.2 类和类之间的相似性测度

根据现有研究文献, 在本研究中, 所采用的类和类之间相似性测度包括以下几种:

2.2.1 最短距离法

$$D(G_1, G_2) = \min_{\substack{x_i \in G_1 \\ x_j \in G_2}} \{d(x_i, x_j)\} \quad (9)$$

式中, G_1, G_2 表示 2 个不同的类, 以下相同。其直观意义为两个类中最近两点之间的距离。

2.2.2 最长距离法

$$D(G_1, G_2) = \max_{\substack{x_i \in G_1 \\ x_j \in G_2}} \{d(x_i, x_j)\} \quad (10)$$

其直观意义为两个类中最远两点之间的距离。

2.2.3 中间距离法

该法在定义类与类之间的距离时, 采用的是最短距离与最长距离之间的中间距离, 当将两类 G_p 和 G_q 合并成新类 G_r 时, 任一类 G_i 和 G_r 之间的距离用下式计算:

$$D_{ir} = \sqrt{\frac{1}{2} D_{ip}^2 + \frac{1}{2} D_{iq}^2 - \frac{1}{4} D_{pq}^2} \quad (11)$$

式中, D_{ip}, D_{iq} 和 D_{pq} 分别为 G_i, G_p 和 G_q 之间的距离。

2.2.4 重心法

该法将两类之间的距离定义为两类重心之间的距离, 它考虑了每一类所包含的样地数, 每一类的重心为该类的均值。

设类 G_p 和 G_q 合并成为 G_r 后, 它们的样地数目分别为 n_p, n_q 和 $n_r = n_p + n_q$ 。则任一类 G_i 与 G_r 之间的距离为:

$$D_{ir}^2 = \frac{n_p}{n_r} D_{ip}^2 + \frac{n_q}{n_r} D_{iq}^2 - \frac{n_p n_q}{n_r n_r} D_{pq}^2 \quad (12)$$

2.2.5 离差平方和法

该法是 Ward 根据方差分析的原理得到的, 如果分类比较合理, 则同类样地之间的离差平方和较小, 类与类之间的离差平方和较大。设将类 G_p 与 G_q 合并成新类 G_r , 则任一类 G_i 与 G_r 之间的距离递推公式为:

$$D_{ir}^2 = \frac{n_i + n_p}{n_r + n_i} D_{ip}^2 + \frac{n_i + n_q}{n_r + n_i} D_{iq}^2 - \frac{n_i}{n_r + n_i} D_{pq}^2 \quad (13)$$

对上述各种类间相似性测度指标, 根据已有研究文献, 没有说明各种方法的优缺点。因为各类间相似性测度指标所采用的类间距离定义不同, 可能使聚类过程不尽相同, 聚类结果也可能存在差异。将上述衡量样地间亲疏程度的多种指标及不同类和类之间的相似性测度进行组合, 可得 25 样地分类方法, 在实际工作中如何合理选择分类组合方法, 下面将通过实例进行分析与研究。

2.3 分类效果评价

不同的亲疏关系测度指标与类和类之间相似性测度指标的组合, 将得到不同的分类结果, 到底哪种分类结果效果好, 可与实际情况比较来确定。为使分类方法在实际工作中得到应用, 必须用一个适当的数学指标描述分类效果, 以便通过计算机仿真自动确定最优的分类组合方法。在本研究中, 采用 Cophenetic 相关系数作为分类效果的评价指标^[5]。

2.4 不同影响因子对分类的影响

因监测区域每个样地可用影响郁闭度及蓄积估测

的 p 个因子来描述,分类是在描述样地 p 个因子确定的前提下进行。不同的影响因子对分类结果有何影响,本文根据特定的监测区域,设置多种类型描述样地的影响因子方案,通过仿真方式对此进行研究。

3 试验分析

取云南思茅地区 7500km^2 范围内的 129 个一类调查样地进行分析^[6]。每个样地考虑的 GIS 因子包括纵坐标、横坐标、海拔、有林地、阴坡、阳坡、郁闭度、蓄积;遥感因子为样地在校正后图像上对应像元的 TM1、TM2、TM3、TM4、TM5 和 TM7 波段的灰度值。对每个样地考虑的因子总数为 14 个,监测区域每个样地对应 14 维欧氏空间中的一点。

采用 Matlab6.5 软件系统进行编程分析^[5],利用内嵌函数 $Y = \text{pdist}(X, 'metric')$ 计算监测区域各样地之间亲疏程度的指标值,参数 'metric' 确定不同的指标种类。所得结果 Y 是长度为 $n(n-1)/2$ 的距离向量。用 $Z = \text{linkage}(Y, 'method')$ 创建系统聚类树,参数 'method' 确定类和类之间相似性测度指标。 Z 为包含聚类树信息的 $(n-1) \times 3$ 矩阵,该矩阵的第 3 列每一行代表类中配对对象之间对应的联结距离。

对每种度量样地间亲疏程度的指标和不同类和类相似性测度指标的结合,将得到不同的分类结果,每种分类的效果可用 Cophenet 相关系数来评价^[5]:

$$C = \frac{\sum_{i < j} (Y_{ij} - \bar{y})(Z_{ij} - \bar{z})}{\sqrt{\sum_{i < j} (Y_{ij} - \bar{y})^2 \sum_{i < j} (Z_{ij} - \bar{z})^2}} \quad (14)$$

式中, Y_{ij} 是 Y 中对象 i 和对象 j 之间的距离; Z_{ij} 是 Z 矩阵第 3 列中对象 i 和对象 j 之间的距离; \bar{y} 和 \bar{z} 分别是 Y 和 Z 矩阵第 3 列的平均值。

C 值越接近于 1,分类效果就越好,特定监测区域样地的具体聚类步骤如图 1 所示。对特定监测区域样地进行分类,分类效果受很多因素影响,本文主要研究不同的分类组合方法及描述样地的遥感和 GIS 因子的差异对分类的影响。

3.1 不同组合的分类方法

不同样地间亲疏程度的测度指标和各种类与类之间相似性测度指标的组合,可得到不同的分类方法。在每个样地所考虑影响因子确定的前提下,不同的分类方法将得到不同的分类效果。设(欧氏距离,马氏距离,绝对距离,明可斯基距离,标准化欧氏

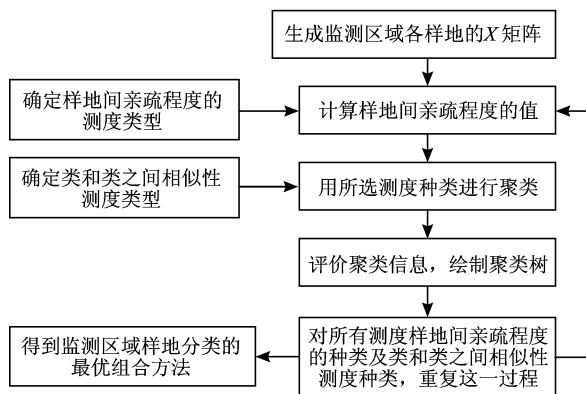


图 1 监测区域样地分类的流程

Fig 1 The flow chart of sample plot classifying in monitoring area

距离}为不同样地间亲疏程度测度指标的集合。{最短距离法,最长距离法,平均距离法,重心距离法,离差平方和法}为类和类相似性测度指标的集合,依次从度量样地亲疏程度的指标集合中取一种指标,与所有类和类测度指标组合,可得 25 种组合分类方法。

3.2 不同的影响因子

对特定监测区域样地进行分类,除分类方法的影响外,分类效果还与每个样地所考虑的影响因子个数及种类有关,本文取云南思茅地区 7500km^2 范围内 129 个地面调查样地对应的遥感和 GIS 因子来进行分析^[6],根据地面调查资料,129 个样地可分为 6 类。现设置以下 5 种方案探讨不同遥感和 GIS 因子对分类的影响^[7,8]:

- (1)仅考虑 GIS 因子的影响, GIS 因子中不包含纵、横坐标,每个样地所考虑的影响因子为海拔、有林地、阴坡、阳坡、郁闭度和蓄积;
 - (2)考虑遥感和 GIS 因子的影响,但 GIS 因子中不包含纵、横坐标,每个样地所考虑的影响因子为海拔、有林地、阴坡、阳坡、郁闭度、蓄积、TM1、TM2、TM3、TM4、TM5 和 TM7;
 - (3)仅考虑 GIS 因子的影响, GIS 因子中包含样地的纵、横坐标;
 - (4)仅考虑遥感因子的影响,每个样地包含的因子为 TM1、TM2、TM3、TM4、TM5 和 TM7;
 - (5)仅考虑通过特定数学方法筛选的影响郁闭度估测的主要因子及郁闭度本身^[8],本文中为有林地、阳坡、TM4、TM5 和郁闭度。
- 对上述 5 种情况,每种又对应 25 种分类方法,作

者设计的软件系统能自动计算各种分类方法所得结果的 Cophenet 相关系数, 绘制 Cophenet 相关系数的变化曲线, 确定最优分类组合方法并用填充的方块表

示, 如图 2 中的 (a), (b), (c), (d), (e) 所示。根据 Cophenet 相关系数的变化及分类效果评价指标, 每种方案都对应一个如表 1 所示的最优分类组合。

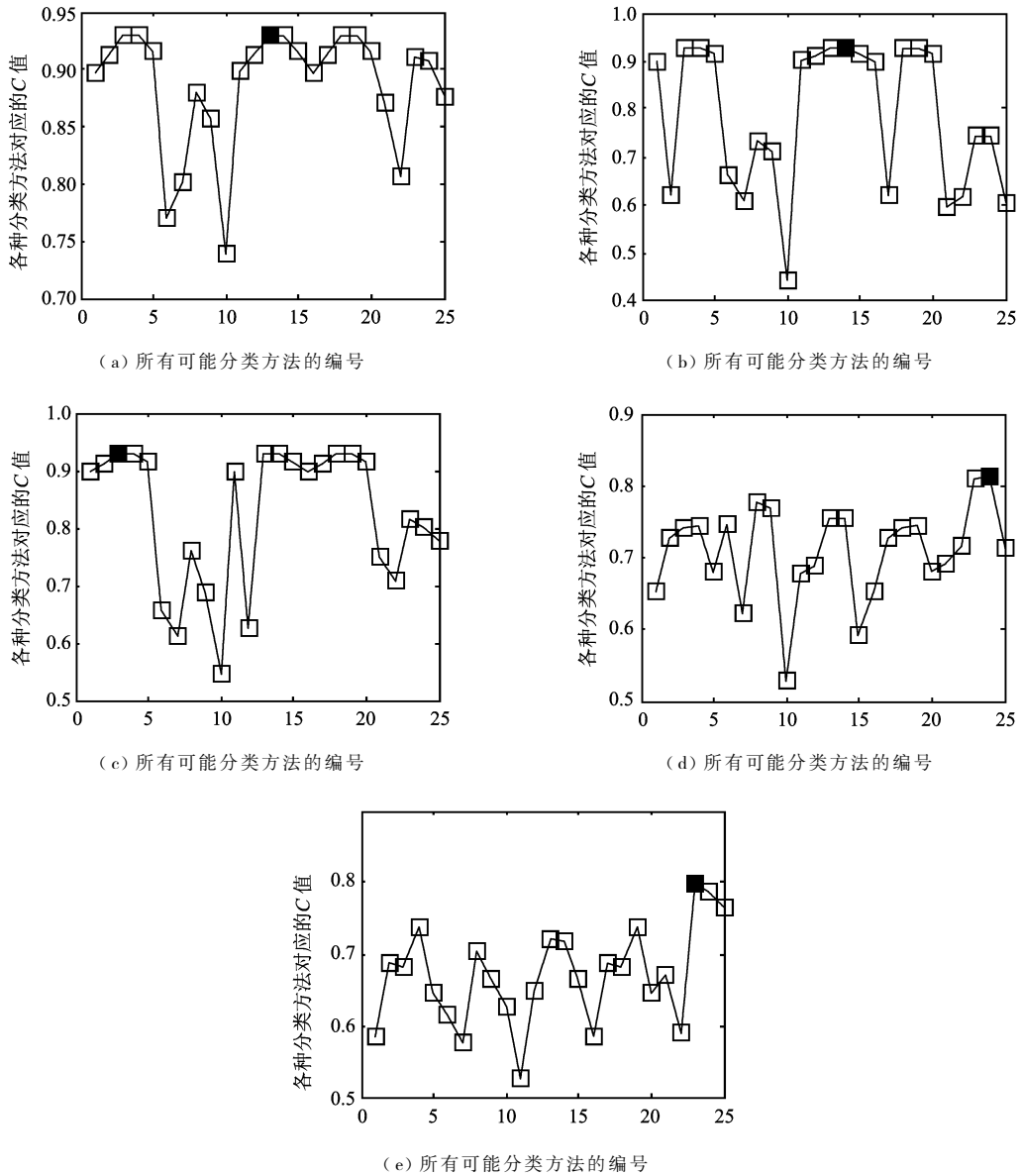


图 2 各种聚类方法对应 C 值的变化

Fig 2 The change of C value for all kinds of classifying methods

表 1 5 种方案所对应的最佳分类组合

Table 1 The optimum classifying combination of five schemes

方案编号	亲疏关系指标种类	聚类方法	Conpenet 相关系数
1	布洛克距离	平均距离法	0.9298
2	布洛克距离	重心距离法	0.9291
3	欧氏距离	平均距离法	0.9204
4	标准化欧氏距离	重心距离法	0.9128
5	标准化欧氏距离	平均距离法	0.7990

分析图表及根据监测区域样地的实际类型可得:

(1) 在监测区域用于描述影响样地分类的遥感和 GIS 因子确定的前提下, 不同的分类组合方法, 分类效果彼此间有一定差异。总的情况是采用马氏距离作为样本间亲疏程度的测度指标与各种类和类之间相似性测度指标的组合, 分类效果较差, 其中马氏距离与离差平方和法的分类效果最差;

(2) 欧氏距离、绝对距离和明可斯基距离分别

与类和类之间相似性测度指标的平均距离法和重心距离法的组合,其分类效果要明显优于其他组合方法;

(3)影响森林郁闭度及蓄积估测的遥感和 GIS 因子的种类及数量对监测区域样地的分类效果影响较大。其影响大小已无法通过 Cophenet 相关系数进行衡量,根据监测区域样地实际类型及 5 种方案的分类结果比较,仅考虑遥感信息的分类效果最差;

(4)在 GIS 信息中,代表样地地理位置的纵、横坐标对分类效果将产生不利影响;

(5)根据特定数学方法选择影响郁闭度估测的

主要遥感和 GIS 因子及郁闭度本身作自变量进行分类,分类效果与监测区域样地实际类型吻合最好,其聚类树如图 3 所示;

(6)在进行分类时,每个样地随着所考虑影响因子的增加, Cophenet 相关系数有增大的趋势,但并不代表其分类效果与实际吻合就好。Cophenet 相关系数只能在样地所考虑影响因子确定的前提下,对 25 种分类组合方法的分类效果进行评价;

(7)进行森林郁闭度估测时,样地的郁闭度应作为影响分类的因子,进行蓄积估测时,样地蓄积也应作为影响分类的因子。

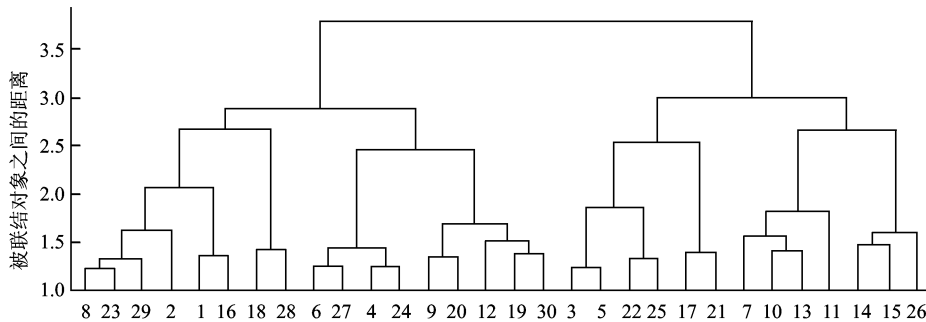


图 3 仅考虑主要影响因子最优分类组合的聚类树

Fig 3 The cluster tree of the optimum clustering combination only thinking main influencing factors

4 结 论

(1)在本文探讨的 25 种分类组合方法中,用马氏距离作为样地间亲疏关系的评价指标和各种类和类之间相似性测度指标的组合,分类效果最差。用欧氏距离、绝对距离和明可斯基距离作为样地间亲疏关系的测度指标,分别与类和类之间相似性测度指标的平均距离法和重心距离法进行组合,其分类效果要明显优于其他分类组合方法。

(2)用于描述样地的遥感和 GIS 因子的种类、个数及具体取值对分类效果影响较大。研究结果表明,代表样地地理位置的纵、横坐标不能作为描述样地的 GIS 因子参与分类,仅考虑样地遥感信息的分类效果最差,应根据特定数学方法优选影响郁闭度及蓄积估测的主要遥感和 GIS 因子对样地进行分类。

(3)在进行分类时,每个样地随着所考虑影响因子的增加, Cophenet 相关系数有增大的趋势,但并不代表其分类效果与实际吻合就好。Cophenet 相关系数只能在样地所考虑影响因子确定的前提下,对

各种可能的分类组合方法的分类效果进行评价。

(4)已有研究文献对各种可能的分类方法进行描述,缺少对不同分类方法分类效果的比较和针对具体分类问题如何自动寻找最优分类方法。在进行森林蓄积估测时,对不同的监测区域,通过数学方法优选用于描述样地的主要遥感和 GIS 因子的种类、数量及具体的取值均存在差异,最优分类组合方法可能不尽相同。采用本文所述的方法,设计软件系统(森林蓄积量遥感估测理论与实现,已于 2004 年 12 月在北京通过鉴定)能通过仿真方式自动确定最优分类组合方法,有效解决了生产应用问题。

参 考 文 献 (References)

- [1] Tang S Z. Multivariate Statistical Analysis[M]. Beijing: China Forestry Press, 1989. [唐守正. 多元统计分析[M]. 北京:中国林业出版社, 1989.]
- [2] Zhang J H. Parameter Estimation of Linear Model and Improvement[M]. Changsha: National University of Defense Technology Press, 1999. [张今槐. 线性模型参数估计及其改进[M]. 长沙:国防科技大学出版社, 1999.]
- [3] Ren R E, Wang H W. Multivariate Statistical Data Analysis[M]. Beijing: National Defense Industry Press, 1997. [任若恩,王惠文. 多元统计数据分析——理论、方法、实例[M].

- 北京:国防工业出版社,1997.]
- [4] Mei C L, Zhou J L. Practical Statistical Method[M]. Beijing: Science Press, 2002 [梅长林,周家良. 实用统计方法[M]. 北京:科学出版社,2002.]
- [5] Su J M, Ruan S Y. MATLAB 6.1 Practical Manual[M]. Beijing: Electronic Industry Press, 2002 [苏金明,阮沈勇. MATLAB 6.1 实用指南[M]. 北京:电子工业出版社,2002.]
- [6] Zhao X W, Li C G. Quantitative Estimation of Forest Resource Based on “3S” Theory Method Application and Software[M]. Beijing: China Science and Technology Press, 2001. [赵宪文,李崇贵. 基于“3S”的森林资源定量估测——原理、方法、应用及软件实现[M]. 北京:中国科学技术出版社,2001.]
- [7] Li C G, Zhao X W, Tian Y L, et al. On Forestry Canopy Density Estimation Based on RS and GIS with Ridge Estimation [J]. Forestry Science, 2001, 37(5): 24-30 [李崇贵,赵宪文,田永林等. 用岭估计研究以RS和GIS为基础的森林郁闭度估测[J]. 林业科学,2001,37(5):24-30.]
- [8] Li C G, Zhao X W. A Research on Variable Selection of Forest Volume Least Square Estimation Based on RS and GIS [J]. Journal of Remote Sensing, 2001, 5(4): 277-281. [李崇贵,赵宪文. 以遥感和地理信息系统为基础的森林蓄积LS估计自变量选择研究[J]. 遥感学报,2001,5(4):277-281.]