

文章编号: 1000-5692(2001)03-0228-05

树高曲线聚类分析研究

刘安兴

(浙江省森林资源监测中心, 浙江 杭州 310020)

摘要: 针对在大区域森林资源连续清查中使用的一元立木材积表不能适用于该区域内任何县一级调查的问题, 提出了用树高曲线、相对欧氏距离和组平均法, 形式聚合图, 再截树的方法进行聚类分析研究, 使材积表在允许误差范围内, 对相近的县域进行归类, 为所编制的大区域的一元立木材积表适用于县级的调查提供理论依据。图2表1参10

关键词: 树高曲线; 一元立木材积表; 聚类分析

中图分类号: S758.62 **文献标识码:** A

立木胸径与树高是衡量林木数量生长和测定立木生物量及立木材积的2个重要指标。它们都随着年龄的累积而增长, 两者都与年龄呈正相关, 且两者自身也呈正相关。在测树学上, 若以某一林分中的某一树种胸径为横坐标, 与其相对应的树高为纵坐标, 即可得到相应的胸径-树高曲线图, 若用数学模型进行分析, 即可得树高曲线的数学表达式, 通常称树高曲线方程(或模型)。在测树制表上, 树高曲线是由二元立木材积表导算一元立木材积表的媒介, 因此树高曲线就成为二元立木材积表导算一元材积表的关键所在。同一种树种的林木在不同的时空中, 其树高曲线的数学模型是不同的。

在大区域的森林资源调查中使用的立木材积表, 过去通常按地形地貌、自然历史和社会经济, 人为地将一些地域上相连, 或者同一行政区内的几个县或地区合并为一个大区域的编表单元, 这就难以保证大区域的立木材积表适用于该区域内任一总体的调查。对此类问题, 有些学者在二元材积式等方面进行了一些研究^[1-4]。

一元立木材积表因其地方性较强而称其为地方性材积表^[5]。本文从树高曲线的聚类分析研究入手, 在材积表所允许的误差范围内, 对相近似的县域树高曲线进行归类, 为所编制的大区域一元立木材积表适用的地域范围与编表总体的地域范围相一致提供理论依据。

1 方法与分析

1.1 树高曲线的制作

首先在一类调查的总体范围内(通常以省为总体), 以县为单位, 根据导算一元立木材积表的要求抽取胸径树高样本。

设全省(G)有 j 个县(G_j), 第 i 个县的某一种木材有 in 株胸径树高样本(D_{in}, H_{in})(令 $i = 1, 2, 3, \dots, j$), 则全省的样本为:

收稿日期: 2001-02-27; 修回日期: 2001-04-17

作者简介: 刘安兴(1955-), 男, 浙江杭州人, 高级工程师, 从事森林资源及其调查监测技术研究。

$$D_{AB} = \frac{1}{n_A n_B} \sum_{\substack{i \in A \\ k \in B}} D_{ik} \quad (6)$$

即样本组 A 与 B 各含有 n_A 和 n_B 个样本, 从 A 组中任取一样本 j 与 B 组中任取一样本 k , 求出两者的距离 D_{jk} ; 样本对共有 $n_A n_B$ 个, 其平均距离就定义为组 A 和组 B 间的距离 D_{AB} 。

假定 A 与 B 合并成组 $A+B$, 则 $n_{A+B} = n_A + n_B$ 个样本, C 为另一个样本组, 有 n_C 个样本, 则

$$D_{C(A+B)} = \frac{1}{n_C n_{A+B}} \sum_{\substack{i \in A \\ k \in A+B}} D_{ik} = \frac{n_A}{n_{A+B}} D_{CA} + \frac{n_B}{n_{A+B}} D_{CB} \quad (7)$$

聚合树状图使全省各县的树高曲线根据其相互之间相对欧氏距离的大小, 最终汇合在一起。为了使所分的组(即编表的区域)能代表其中的各具体县, 即归类后各县的数据共同所得的树高曲线, 在允许误差范围内能适用于其中的各个县, 因此要求得到一个距离值作为截树状图的截值(R), 截后的各组, 即为我们所求的各区域。

从定积分的概念, 可得县 i 的树高曲线的平均高为:

$$H_i = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} f(D_i) dD \quad (8)$$

全省的平均树高为:

$$H = \frac{1}{j} \sum_{i=1}^j H_i \quad (9)$$

原林业部部颁标准材积表采用山本材积式, 其数学式为:

$$V = \alpha D^b H^c \quad (10)$$

式中: V 为立木材积, a, b, c 为参数。

从(4)式已知县(G_A)与县(G_B)之间树高的标准差为 $d(G_A, G_B)$, 则材积的标准差为 $\alpha D^b [d(G_A, G_B)]^c$ 。若材积表的允许的相对误差限为 E , 则

$$E = \frac{t \alpha D^b [d(G_A, G_B)]^c / \sqrt{\beta - \alpha}}{\alpha D^b [(H_A + H_B)/2]^c} \quad (11)$$

t 为可靠性指标。

为了全省的截值统一, 我们以(9)式的 H 代替 $(H_A + H_B)/2$, 则

$$E = \frac{t \alpha D^b [d(G_A, G_B)]^c / \sqrt{\beta - \alpha}}{\alpha D^b H} \quad (12)$$

由(12)式得截值 R 为:

$$R = d(G_A, G_B) = H \left[\frac{E \sqrt{\beta - \alpha}}{t} \right] \quad (13)$$

以截值 R 截树, 得全省所分的区域。

2 结果与讨论

为了体现聚类分析研究的结果, 我们从导算浙江省一元立木材积表所收集的杉木(*Cunninghamia lanceolata*)树高样木中, 抽取径阶范围从6~30 cm, 分布均匀, 且株数在60株以上的7个县, 举例说明结果, 并展开讨论。

每个县的数据, 均以若干个一元线性回归方程用计算机比选。以相关系数最为紧密的数学模型作为该县树高曲线的数学表达式。选择出的最佳树高曲线式类型如下:

$$1/H = a + b/D \quad (14)$$

$$\ln H = a + b \ln D \quad (15)$$

具体参数见表1。

按公式(4), 其 α 取6(即胸径6 cm), β 取30(即胸径30 cm), 计算县与县之间的相对欧氏距离, 得7×7阶相对欧氏距离矩阵。即

$$\begin{bmatrix} G_1 & G_2 & G_3 & G_4 & G_5 & G_6 & G_7 \\ 0 & 1.689\ 25 & 0.687\ 19 & 1.069\ 53 & 1.447\ 67 & 0.836\ 67 & 0.429\ 35 \\ & 0 & 1.184\ 02 & 0.892\ 43 & 0.307\ 49 & 2.163\ 80 & 1.800\ 91 \\ & & 0 & 0.418\ 28 & 1.027\ 89 & 0.983\ 14 & 0.633\ 25 \\ & & & 0 & 0.805\ 61 & 1.315\ 91 & 1.015\ 24 \\ & & & & 0 & 1.015\ 24 & 1.608\ 04 \\ & & & & & 0 & 0.437\ 80 \\ & & & & & & 0 \end{bmatrix} \quad (16)$$

表 1 各县树高曲线参数

Table 1 Parameter of height curve

县名 (代号)	公 式	参数 a	参数 b	相关系数 r
衢县 (G_1)	(14)	-0.002 32	1 837 17	0.948 61
诸暨 (G_2)	(14)	0.031 79	1 344 26	0.903 33
临海 (G_3)	(14)	0.016 64	1 454 82	0.899 19
仙居 (G_4)	(15)	0.183 19	0 735 97	0.899 35
云和 (G_5)	(15)	-0.071 79	0 794 08	0.924 02
宁海 (G_6)	(15)	-0.198 87	0 897 09	0.926 09
武义 (G_7)	(15)	0.336 15	0 929 70	0.960 81

由公式 (8) 分别计算 7 个县的平均树高：衢县为 10.060 92 m，诸暨为 9.093 68 m，临海为 10.043 37 m，仙居为 9.922 60 m，云和为 9.118 94 m，宁海为 10.878 37 m，武义为 10.442 63 m，由公式 (9) 计算各县的总平均树高为 9.937 22 m。

由公式 (13) 计算截值 R 。取 $H=9.937\ 22$ ； $E=0.05^{[9]}$ ； α 与 β 同前，即 $\alpha=6$ ， $\beta=30$ ， $t=1.96$ ， $c=0.896\ 462^{[10]}$ 。得：

$$R = 9.937\ 22 \left[\frac{0.05 \sqrt{30-6}}{1.96} \right]^{1/0.896\ 462} = 0.976\ 72.$$

以截值 $R=0.976\ 72$ 截 7 个县 的 聚 合 树 状 图，7 个县 的 杉 木 树 高 曲 线 的 数 学 表 达 式 在 95% 的 可 靠 性，材积相对误差不超过 5% 的前提下，可以归为 2 类，即衢县、武义、临海、仙居和宁海为一类，诸暨和云和为另一类（图 2）。2 类即可分别将其中各县的树高样木数据合在一起导算树高曲线的回归方程。

结果显示树高曲线之间的比较分析是可行的。这使得我们将过去所作的成对值比较的基础上，可进一步通过拟合曲线，进行连续性函数的比较分析，为材积表式相互之间的直接比较与分析提供了一种技术方法和技术思路。

树高曲线之间的相对欧氏距离，不受公式类型的影响，相同的公式类型受其参数的影响。结果显示，诸暨和云和的树高曲线并不相同，但它们之间的相对欧氏距离却很小，表明 2 条树高曲线非常接近。在树高曲线的公式类型相同的县之间，可以由直观的参数之间的差异判其相对欧氏距离的大小，如武义分别与云和、仙居的参数情况，与相对欧氏距离的大小相对应。

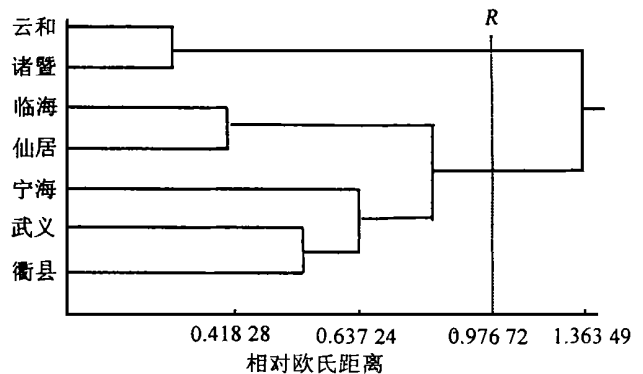


图 2 7 个县组平均法聚合树状图

Figure 2 Cluster tree diagram of class mean method about 7 counties

比较完备可行的。在对于截树的截值导算上, 仅作了初步的探讨, 根据已建立的二元材积式导算一元材积式, 以及与截值相关的误差计算方法的导算方面上, 还有待进一步的研究。

参考文献:

- [1] 贺东北, 曾伟生, 骆期邦. 通用性二元立木材积模型的建模样本研究[J]. 中南林业调查规划, 2001, 20(1): 1-9.
- [2] 骆期邦, 吴志德. 相对树高曲线模型的研究和应用[J]. 中南林业调查规划, 1995, 14(2): 1-3.
- [3] 谢利玉. 林业数表的概念及其特点[J]. 林业资源管理, 1998, (4): 34-37.
- [4] 何宗明, 洪伟. MSM法优化材积方程的研究[J]. 林业科学, 1993, 35(6): 558-562.
- [5] Avcy T E. *Natural Resources Measurements* [M]. New York: McGraw-Hill Book Company, 1975. 93-95.
- [6] 贺仲雄. 模糊数学及其应用[M]. 天津: 天津科学技术出版社, 1983.
- [7] M. 费史. 概率论及数理统计[M]. 王福保, 译. 上海: 上海科学技术出版社, 1962. 63-70.
- [8] 阳含熙, 卢泽愚. 植物生态学的数量分类方法[M]. 北京: 科学出版社, 1981. 111-114.
- [9] 林业部调查规划院. 森林调查手册[M]. 北京: 中国林业出版社, 1980. 58-63.
- [10] 刘安兴, 龚赛凤, 孙源和. 浙江省林木根径材积表及其数学模型[J]. 华东森林经理, 1992, 6(3): 12-18.

On cluster analysis of tree height curves

LIU An-xing

(Monitoring Center for Forest Resources in Zhejiang Province, Hangzhou 310020, Zhejiang China)

Abstract: One-way volume table used in large scale continuous forest inventory is unsuitable for forest inventory of county-class in this region. To counter this problem, the cluster analysis method is offered. The procedure of cluster is as following: the cluster diagram is firstly made with tree height curve, relative Euclidean distance and class mean method, and then is divided into units. This method provided theoretical basis for classifying the similar counties within range of error, and for making regional one-way volume table being suitable for county-class forest inventory.

Key words: tree height curve; one-way volume table; cluster analysis