

蒙古栎林全林整体生长模型及其应用

洪玲霞¹, 雷相东¹, 李永慈²

(1. 中国林业科学研究院资源信息研究所, 北京 100091; 2. 北京林业大学基础学院, 北京 100083)

摘要:利用汪清林业局 1997 年、2007 年森林经理调查的 61 块蒙古栎林的固定样地数据建立了蒙古栎林全林整体模型。全林整体模型是一组非线性联立方程组, 它由 8 个非线性模型组成。利用 Forstat2.0 软件中的“非线性误差变量联立方程组”方法求解模型参数, 保证了模型的无偏性。“刀切法”检验结果表明各林分因子的平均相对误差和相对均方误差均在 15% 以下。应用全林整体模型可以进行不同初始条件林分的生长分析及制定不同立地条件的林分密度控制图, 为蒙古栎林的经营提供参考。

关键词:蒙古栎林; 全林整体模型; 林分密度控制图

中图分类号: S711

文献标识码: A

Integrated Stand Growth Model of Mongolian Oak and It's Application

HONG Ling-xia, LEI Xiang-dong, LI Yong-ci

(1. Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China;

2. Beijing Forestry University, Beijing 100083, China)

Abstract: This paper established an integrated stand growth model of Mongolian oak (ISGM_oak) using the data from 61 permanent sample plots measured in 1997 and 2007. ISGM_oak is a group of nonlinear simultaneous equations. The method of nonlinear error-in-variable simultaneous equations is used to estimate the parameters of ISGM_oak with the statistical software Forstat 2.0, so the parameter estimation of the group of correlated equations in ISGM_oak is unbiased and the equations are compatible. Model validation using bootstrap method showed that both the average relative error and square error are less than 15 percent. The ISGM_oak model can be used to simulate the stand growth with different values of site index, stand density and to draw stand density management diagram for decision making.

Key word: *Quercus mongolica*; integrated stand growth model; stand density management diagram

蒙古栎 (*Quercus mongolica* Fisch) 是我国东北地区分布最广泛的树种之一。在我国东北有大面积的蒙古栎天然次生林。蒙古栎不但具有很高的经济价值, 而且还具有保持水土、涵养水源等作用。一些学者对蒙古栎的生长规律进行了研究, 例如王春霞等^[1]和王世忠等^[2]对蒙古栎的生长过程进行了研究; 亢新刚等^[3]编制了冀北蒙古栎天然次生林的经验生长过程表; 杜纪山^[4]以优势树种(组)的方法建立了蒙古栎的全林整体模型; 陈新美^[5]建立了蒙古

栎林直径分布模型; 但在解决林分生长模型的相容性方面, 均没有考虑自变量也具有误差的情形。当自变量和因变量的观测值都具有度量误差时, 通常的回归模型参数估计方法不再适用, 而度量误差模型可以解决这一问题^[6]。李永慈^[7]的研究也表明, 非线性度量误差联立方程组方法明显优于最小二乘法。由于以前没有求解非线性联立方程组的方法, 各种文献上都是采用逐个非线性方程求解的方法来建立模型。这种建模型方法得到的参数是有偏

估计^[8-9]。在世纪之交,已经出现了比较好的非线性联立方程组的解法^[6,8,10-11],唐守正、李永慈等又研究了度量误差对模型参数估计的影响及带度量误差的全林整体模型参数估计方法^[7,9,12-15],并且这种方法包含在 Forstat2.0 软件^[8]中。本研究用吉林省汪清林业局 1997 年、2007 年森林经理调查的 61 块蒙古栎林的固定样地数据和唐守正提出的全林整体模型^[16],采用 Forstat2.0 软件中的“非线性误差变量联立方程组”方法建立了蒙古栎林全林整体模型,保证了参数估计的无偏性。利用刀切法验证了模型的精度及模型参数的稳定性,应用全林整体模型建立了林分密度控制图、模拟了某一个现实蒙古栎林分(已知立地指数、年龄、株数、平均直径)今后的林分生长过程。为该地区蒙古栎的森林经营决策提供方

法和参考依据。

1 研究地点和实验数据概况

研究地区为吉林省汪清林业局,该局始建于 1947 年,全局经营面积 30.4 万 hm^2 ,地处吉林省延边朝鲜族自治州东北部汪清县城内,位于 $129^\circ56' \sim 131^\circ04' \text{ E}$, $43^\circ05' \sim 43^\circ40' \text{ N}$,属长白山系老爷岭山脉雪岭支脉,地貌属低山丘陵,植被属长白山植物区系。本研究所采用的数据为汪清林业局 1997 年、2007 年森林经理调查的 61 块蒙古栎纯林(蒙古栎占 10 成)的固定样地数据,样地面积为 0.06 hm^2 ,平均年龄、平均树高是每块样地抽取 3 棵平均木进行实际测量后的平均值,平均直径是断面积平均直径,具体情况见表 1。其中有 6 块样地进行过采伐。

表 1 样地基本因子统计量

调查时间	株数/(株· hm^{-2})	平均年龄/a	平均直径/cm	平均树高/m	蓄积/($\text{m}^3 \cdot \text{hm}^{-2}$)
1997 年	650~4 100	25~76	7.4~20.3	7.2~12.5	48.7~235.4
2007 年	633~3 733	35~86	8.7~22.2	8.1~13.2	40.1~267.7

2 研究方法

2.1 全林整体生长模型及参数估计

全林整体模型是 1991 年由唐守正提出^[16],以后又逐渐完善的一个人工林生长、经营模型^[17-21]。这个模型是一组由 8 个非线性模型组成的非线性联立方程组,其中包括:1)断面积模型;2)密度指数定义;3)断面积公式;4)自稀疏模型;5)优势高生长模型;6)平均高模型;7)蓄积公式;8)形高模型。

由于森林经理调查数据中各样地只有平均高数据,没有调查优势高,所以将上述全林整体模型中的立地指数 L 用地位级代替,优势高生长模型(5)改为平均高生长模型,本研究所建立的全林整体模型由下面 7 个非线性的联立方程组模型组成:

$$(1) \text{断面积模型: } G = b_1 \times L^{b_2} \times [1 - \exp(-b_4 \left(\frac{S}{10\,000}\right)^{b_5} \times (age - t_0))]^{b_3} \quad (1)$$

$$(2) \text{密度指数定义: } S = N \times \left(\frac{D}{20}\right)^\beta \quad (2)$$

$$(3) \text{断面积公式: } G = N \times D^2 \times \left(\frac{\pi}{40\,000}\right) \quad (3)$$

$$(4) \text{自稀疏模型: } \left(\frac{sf}{N}\right)^\beta - \left(\frac{D}{20}\right)^{\beta\gamma} = C \quad (4)$$

$$(5) \text{平均高生长模型: } PH = L \times \exp\left(-\frac{b}{A} + \frac{b}{A_0}\right) \quad (5)$$

$$(6) \text{蓄积公式: } M = fH \times G \quad (6)$$

$$(7) \text{形高模型: } F = \left(c_1 + \frac{c_2}{H+2}\right) \times H \quad (7)$$

式中变量符号: A (年龄), N (株数), D (胸径), H (平均高), F (形高), L (地位级), t_0 (平均高达到胸高的年龄), A_0 (立地基准年龄), 其他: $b_1, b_2, b_3, b_4, b_5, sf, \beta, \gamma, b, c, c_1, c_2$ 是待估参数。C 为常数,对同一林分,由初始密度决定。

公式(1)到(7)是非线性联立方程组。前 5 个模型采用对数形式,以便使剩余方差接近等方差。然后按下述步骤利用 Forstat 2.0 软件估计参数。

由立地类型、年龄和平均树高,采用哑变量方法估计各类型的地位级指数,以及平均高生长模型的参数 b (舒马克型)。把密度指数定义(2)和断面积公式(3)代入断面积模型(1),把形高公式(7)代入蓄积公式(6),再和自稀疏模型(4)联立,得到下面的联立方程组:

$$\begin{aligned} 2y_1 + y_2 &= b_1 + b_2x_2 + b_3 \log(1 - \exp \\ &\quad (-b_4(\exp(y_2 + \beta(y_2 - g_{20})))^{b_5} \\ &\quad (x_1 - t_0))) - \log(\pi_0) \end{aligned} \quad (8)$$

$$y_2 = \alpha - \frac{1}{\gamma} \log[\exp(y_1 - \log(A_0))^{\beta\gamma} + \exp(\alpha - x_4)^\gamma - \exp(x_5)^{\beta\gamma}] \quad (9)$$

$$y_3 = \log(x_3) + \log(c_1 + \frac{c_2}{x_3 + 2}) + y_2 + 2y_1 + \log(\pi_0) \quad (10)$$

其中: $\pi_0 = \pi/40\ 000$, $g_{20} = \log(20)$, t_0 = 平均高达到 1.3 m 的年龄(选项给出), A_0 = 立地指数基准年龄(选项给出), $\alpha = \log(sf)$ 。

不含误差的变量: $x_1 = A$, $x_2 =$ 立地指数(已经由步 1 算出), $x_3 =$ 平均高, $x_4 = \log(\text{每个样地第一次观测的株数})$, $x_5 = \log(\text{每个样地第一次观测的直径}/20)$ 。

含误差的变量: $y_1 = \log(D)$, $y_2 = \log(N)$, $y_3 = \log(M)$ 。

应用 Forstat2.0 软件计算全林整体模型参数需要同一树种的多个时间复测的样地的观测数据。观测数据包括 2 个分类因子和 5 个林分测树因子。分类因子为样地编号和立地类型编号。5 个林分测树因子是:林分年龄、公顷株数、断面积平均直径、平均高和公顷蓄积。相同立地类型的样地具有同一立地类型编号。如果不输入立地类型编号,则认为每个样地具有不同的立地类型。每一个样地内最少有两次不同年龄的观测值。详见参考文献[8]。

2.2 模型检验

由于受样本数量和观测次数所限,本文采用刀切法来检验模型的精度。刀切法(jackknife)是由 Quenouille 提出的一种再抽样分析统计量的工具^[6]。具体是将 61 块样地数据依次剔除一块样地数据,例如,剔除第 i 个样地(不包括 5 个间伐样地),用其余的 60 块样地计算模型参数(包括间伐样地),然后再用该套模型参数和第 i 个样地 1997 年的数据作为初始值(株数、年龄、断面积),估计第 i 个样地 2007 年的林分因子。模型检验统计量包括:平均偏差(E_{MD})、平均绝对偏差(E_{MAD})、平均相对误差(E_{MRD})、均方误(E_{MSE})、相对均方误(E_{RMSE})和决定系数 R^2 。数学表达式如下:

$$\text{平均误差 } E_{MD} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (11)$$

$$\text{平均绝对误差 } E_{MAD} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (12)$$

$$\text{平均相对误差 } E_{MRD} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n y_i} \quad (13)$$

$$\text{均方误 } E_{MSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}} \quad (14)$$

$$\text{相对均方误 } E_{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-1)}{\sum_{i=1}^n y_i / n}} \quad (15)$$

$$\text{决定系数 } R^2 = 1 - \frac{\sum_{i=1}^n |y_i - \hat{y}_i|^2}{\sum_{i=1}^n |y_i - \bar{y}_i|^2} \quad (16)$$

其中 n 为样本数, y_i 为观测值, \hat{y}_i 为预测值, \bar{y}_i 为观测值的平均值。

3 结果与分析

3.1 地位级指数

本研究所使用的 61 块样地中有 5 块样地进行了间伐,所以将这 5 块间伐样地 2 个年度的样地号分别编写。用 Forstat2.0 软件中的“地位级曲线”模块,将 61 块样地分为 4 个立地类型,地位级曲线方程为 $H = a(1 - \exp(-cA))^b$,其中 A 为年龄, a, b, c 为待估参数,计算结果见表 2。其中 2 和 3 地位级有较多的样地。

表 2 地位级曲线参数及立地类型划分结果

地位级编号	a	b	c	样地数
1	9.578 055	1.076 871	0.052 524	9
2	10.666 979	1.233 781	0.052 524	18
3	11.879 702	1.390 692	0.052 524	26
4	13.230 299	1.547 603	0.052 524	8

3.2 模型参数

采用全部数据及采用刀切法估算的参数见表 3。其中达到胸高的年龄 $t_0 = 4$,立地基准年龄 $A_0 = 40$ 。建议模型的使用范围为:地位级 7~11;年龄 12~129;密度指数 336~1 714。可以看出,参数 b_1 和 b_4 变动较大,其它参数比较稳定,二者相差不大。

3.3 模型精度验证

使用全部数据所建立的蒙古栎林全林整体生长模型经回归适应性检查结果见表 4,回归方程: $y = a + bx$ (y 为实测值, x 为估计值,零假设 $a = 0, b = 1$)。

表3 蒙古栎全林整体生长模型及参数

项目	全部数据	刀切法计算参数($n=56$)			
		平均值	最大值	最小值	标准差
断面积模型参数					
b_1	18.220 1	12.719 4	20.154 3	6.834 5	2.812 4
b_2	0.713 4	0.674 3	0.886 5	0.511 7	0.089 4
b_3	0.401 5	0.411 9	0.428 0	0.395 3	0.008 6
b_4	0.992 2	4.080 4	5.232 5	0.776 9	1.169 8
b_5	2.490 5	2.428 7	2.529 7	2.336 3	0.051 1
自稀疏模型参数					
sf	1 714	1 714	1 714	1 714	
β	1.38	1.38	1.38	1.38	
γ	3.963 4	5.069 0	6.716 3	3.980 0	0.861 9
形高模型参数					
C_1	0.219 9	0.220 1	0.230 6	0.192 3	0.005 7
C_2	4.481 3	4.481 6	4.835 0	4.357 0	0.072 8
平均高模型参数					
b	14.03	14.127 2	15.623 7	12.916 5	0.603 8

表4 株数、直径、蓄积的回归方程适应性检查结果

检查因子	a	b	相关系数	F -统计量	检验结果
直径	0.69	0.96	0.918	0.763 7	差异不显著
株数	55.82	0.98	0.982	2.069	差异不显著
蓄积	16.71	0.91	0.912	1.645	差异不显著

临界值 $F=3.071\ 779, p=0.050\ 000$, 一自由度=2, 二自由度=120

以1997年林分初始状态为初值,利用刀切法计算的模型参数估算2007年各样地的林分因子(林分平均胸径、平均树高、林分断面积、株数和蓄积),模型拟合及误差统计量见表5。可以看出,各林分因

子的平均相对误差和相对均方误差均在15%以下,平均胸径和平均树高的估计较为准确,模型均过高估计了其它林分因子。

表5 2007年的实测数据与全林整体模型估计结果的比较

统计量	蓄积/ $(\text{m}^3 \cdot \text{hm}^{-2})$	断面积/ $(\text{m}^2 \cdot \text{hm}^{-2})$	树高/m	直径/cm	株数/ $(\text{株} \cdot \text{hm}^{-2})$
平均偏差(E_{MD})	-16.12	-2.38	-0.18	0.05	-155.26
平均绝对偏差(E_{MAD})	20.46	2.52	0.68	0.45	155.76
平均相对误差($E_{\text{MRD}})/\%$	-11.25	-9.58	-1.24	0.14	-10.04
均方误(E_{MSE})	23.97	2.95	0.81	0.57	208.31
相对均方误($E_{\text{RMSE}})/\%$	14.85	11.23	7.61	3.62	14.10
决定系数 R^2	0.98	0.99	0.99	0.99	0.98

3.4 应用

在计算出蒙古栎全林整体模型的参数后,可以反复用此程序模拟蒙古栎在不同条件下(在使用范围内)林分生长收获过程,或绘制各种立地条件下的林分密度控制图。其中,模拟生长过程包括:情况1:模拟蒙古栎在某一个立地(指定的地位指数)的条件下,各种初始密度的林分生长过程。情况2:模拟蒙古栎的某一个现实林分(已知地位指数、年龄、株数、平均直径)今后的生长过程。详细计算方法和

原理见参考文献[8]。在此,以情况2为例,模拟一个现实林分(306号样地,地位指数为10)的生长过程见表6及图1。图1中 N (公顷株数), M (公顷蓄积), G (公顷断面积), D (平均直径), H (平均树高)为306号样地从30年生长到80年的自然生长曲线,横坐标是年龄,纵坐标是以10为底的对数坐标,这样可以将公顷株数、公顷蓄积、公顷断面积、平均直径及平均树高这几个数值范围不同的林分因子的生长曲线画在同一个图上以便进行比较。

表6 现实林分生长收获简表

年龄/ a	蓄积/ ($\text{m}^3 \cdot \text{hm}^{-2}$)	蓄积平均生长量/ ($\text{m}^3 \cdot \text{hm}^{-2} \cdot \text{a}^{-1}$)	蓄积连年生长量/ ($\text{m}^3 \cdot \text{hm}^{-2} \cdot \text{a}^{-1}$)	断面面积/ ($\text{m}^2 \cdot \text{hm}^{-2}$)	树高/ m	直径/ cm	株数/ (株 $\cdot \text{hm}^{-2}$)	密度 指数
30	93.15			16.59	8.90	10.79	1 816	774.44
40	145.90	3.70	5.64	24.59	10.00	13.28	1 777	1 009.19
50	196.38	3.98	5.15	32.01	10.73	15.43	1 713	1 196.84
60	240.99	4.06	4.44	38.40	11.24	17.31	1 632	1 336.88
70	278.87	4.02	3.74	43.72	11.62	18.99	1 544	1 437.09
80	310.68	3.91	3.15	48.11	11.92	20.51	1 456	1 507.63

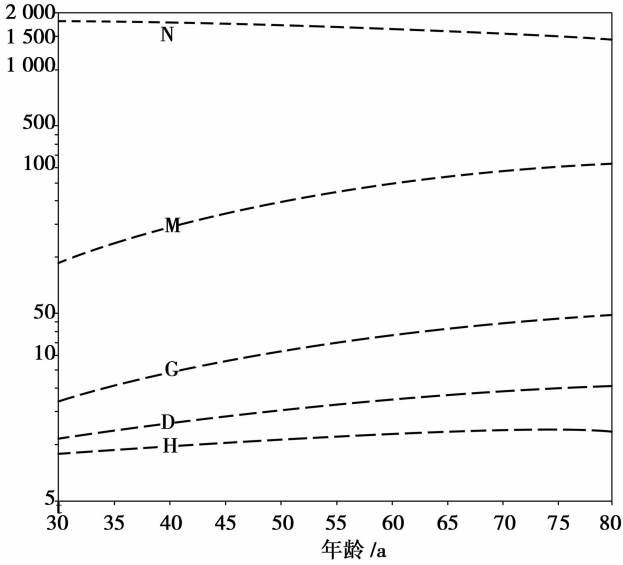


图1 林分因子生长过程

传统的林分密度控制图是以 $\ln(N)$ 为横坐标, $\ln(M)$ 为纵坐标的 4 族等值线, 不分立地指数。利用 Forstat2.0 软件可以计算输出不同地位级(立地指数)林分密度控制图的 4 族等值线的值, 然后以林分的 $\ln(N)$ 为横坐标, $\ln(M)$ 为坐标轴绘制林分密度图(见图 2), 图中等值线包括: 等优势高曲线族(本文为平均高 H)、等直径曲线组(D)、自然稀疏线族(N)和等经营度线族(J)。经营度 $J=1$ 的曲线就是完满立木度曲线。但是对于相同直径、相同株数的林分, 由于立地指数的不同其树高并不相同, 因而蓄积不同。这是传统林分密度控制图误差较大的一个重要原因。由全林整体模型推导出的林分密度控制图, 是分别地位级或立地指数编制的。它的精度与模型精度完全一致。

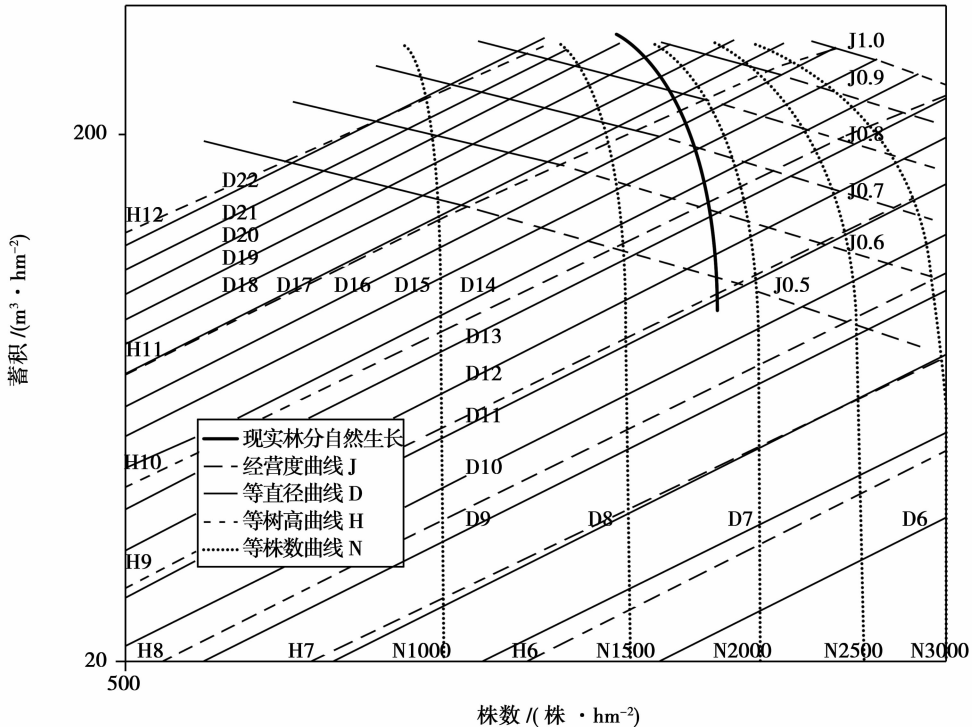


图2 地位级为 10 的蒙古栎林林分密度控制图

4 结论与讨论

(1) 利用 61 块蒙古栎林复测样地数据,采用非线性度量误差联立方程组方法,建立了蒙古栎林全林整体生长模型,保证了模型的相容性和参数估计的无偏性。

(2) 采用刀切法对模型的精度及模型参数的稳定性进行了检验,表明各林分因子的平均相对误差和相对均方误差均在 15% 以下,平均胸径和平均树高的估计较为准确,模型均过高估计了其它林分因子。因此可以用于该地区蒙古栎林的生长预测和经营决策。与杜纪山的研究相比,在参数估计方法上进行了改进。

(3) 利用建立的蒙古栎林全林分生长模型可以模拟蒙古栎在某一个立地(指定的立地指数)的条件下,各种初始密度的林分生长过程,本文模拟了一个现实林分的生长过程,给出了生长过程表及图示。绘制了林分密度控制图。

(4) 本文初步尝试将全林整体生长模型用于模拟异龄近似纯林的生长情况,并给出了精度验证,今后,在样本量增加的情况下,可以进一步修正模型的参数及验证模型的精度。

(5) 由于异龄林的主伐收获方式比较复杂,如何模拟抚育间伐的生长情况留待以后讨论。

参考文献:

- [1] 王春霞,刘万成,刘瑰琦,等. 大兴安岭林区蒙古栎生长过程研究[J]. 中国林副特产,2005(5):12-14
- [2] 王世忠,步兆东,丁玉武,等. 辽西地区蒙古栎生长规律的研究[J]. 辽宁林业科技,1993(3):26-27
- [3] 亢新刚,崔相慧,王虹. 冀北次生林 3 个树种林分生长过程表的编制[J]. 北京林业大学学报,2001,23(3):39-42
- [4] 杜纪山,唐守正,王洪良. 天然林区小班森林资源数据的更新模型[J]. 林业科学,2000,36(2):26-32
- [5] 陈新美,张会儒,武纪成,等. 蒙古栎林直径分布模拟研究[J].

- 林业资源管理,2008(1):39-43
- [6] 唐守正,李勇. 生物数学模型的统计学基础[M]. 北京:科学出版社,2002
- [7] 李永慈,唐守正. 度量误差对模型参数估计值的影响及参数估计方法的比较研究[J]. 生物数学学报,2006,21(2):285-290
- [8] 唐守正,郎奎建,李海奎. 统计和生物数学模型计算(ForStat 教程)[M]. 北京:科学出版社,2009
- [9] 李永慈,唐守正. 带度量误差的全林整体模型参数估计研究[J]. 北京林业大学学报,2006,28(1):23-27
- [10] Tang S Z, Li Y, Wang Y H. Simultaneous equations, error-in-variable models, and model integration in systems ecology Source[J]. Ecological Modelling, 2001, 142(3): 285-294
- [11] Tang S Z, Wang Y H. A parameter estimation program for the error-in-variable model[J], Ecological Modelling, 2002, 156(23): 225-236
- [12] 李永慈,唐守正. 度量误差对全林整体模型的影响研究[J]. 林业科学,2005,41(6):166-169
- [13] 李永慈,唐守正,李海奎,等. 用度量误差模型方法编制相容的生长过程表和材积表[J]. 生物数学学报,2004,19(2):199-204
- [14] 李永慈,唐守正,徐松. 线性度量误差模型的参数估计法与最小二乘法的关系[J]. 生物数学学报,2008,23(1):139-142
- [15] Hong L X, Tang S Z, Li Y C. Integrated Stand Growth Model (ISGM) and Its Application[R]. IEEE Proceeding of Second International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications. 2007:223-230
- [16] 唐守正. 广西大青山马尾松全林整体模型及其应用[J]. 林业科学研究,1991,4(增刊):8-13
- [17] 唐守正. 同龄纯林自然稀疏规律的研究[J]. 林业科学,1993,29(3):234-241
- [18] 李希菲. 大青山实验局主要树种(组)全林整体模型及精度验证[J]. 林业科学研究,1991,4(增刊):14-21
- [19] 洪玲霞. 由全林整体生长模型推导林分密度控制图的方法[J]. 林业科学研究,1993,5(6):510-516
- [20] 唐守正,杜纪山. 利用树冠竞争因子确定同龄间伐林分的断面面积生长过程[J]. 林业科学,1999,35(6):35-41
- [21] 唐守正,李希菲. 用全林整体生长模型计算林分纯生长量的方法及精度分析[J]. 林业科学研究,1995,8(5):471-476