

## 二元立木材积动态模型研究

骆期邦 宁辉 贺东北 蒋菊生

吴志德 曾伟生 张曙光

(林业部中南林业调查规划设计院)

**摘要** 为提高二元立木材积模型的精度,以雪岭云杉(*Picea schrenkiana* Fish. et Mey.)和杉木(*Cunninghamia lanceolata* (Lamb.) Hook.)两个树种为例,研究提出了能充分利用样本信息的二元立木材积动态模型(其通式为 $V = C_0 \cdot D^{f_1(D,H)} \cdot H^{f_2(D,H)}$ )及其建模原理与方法。这种模型不仅可显著提高精度,而且其参数可依据样本资料所反映的客观规律作出解释,模型稳定可靠。

**关键词** 二元立木材积动态模型 异方差 参数稳定性 残差分布图

二元立木材积模型,是森林蓄积量清查中最常用的计量依据。要保证清查结果准确,提高二元立木材积模型精度十分重要。考察已发表的二元立木材积模型,一个共同的缺陷是,不能充分利用样本资料所提供的信息而使精度的提高受到限制。为了提高精度,人们曾通过采用多项式逼近和分段建模的方法,但效果并不理想。本文的研究目的,旨在依据样本资料所反映的胸高形数变化规律,探求其参数能作出符合客观规律解释的动态模型,力求模型能充分利用样本所提供的信息,使参数随胸径和树高的变化而变化,追踪客观规律,从而把二元立木材积模型的精度提高到一个新的水平。

## 1 样本资料和研究方法

为便于深入探讨动态模型的结构和功能,本研究选用了样本含量较大(9 049株样木),胸径和树高变幅较宽,干形变化规律比较复杂的雪岭云杉为研究样本<sup>[1]</sup>,同时还收集了一套由990株样木组成的杉木资料为辅助研究样本。二元立木材积变化规律的研究,实质上是干形随胸径和树高变化规律的研究,为了便于更直观地分析样本所提供的干形变化信息,将样本资料按径阶(间距4 cm)和树高阶(间距2 m)归类,求出平均胸高形数( $f_{1.3}$ )和相应的株数( $n$ )列于表1。将材积转换为形数的另一重要目的是,它能消除立木材积建模中必然存在的异方差影响。在胸径( $D$ )和树高( $H$ )二元控制的条件下,材积方差随 $D^2H$ 的增大而增大的规律是客观存在的<sup>[1]</sup>。为了消除不等方差在建模中的影响,一般可采用两种方法<sup>[2]</sup>:一是采用以 $W = 1/(D^2H)^2$ 为权的加权最小二乘法求解模型参数;二是进行变量变换,即将材积除以 $\pi \cdot D^2 \cdot H / 40\,000$

本文于1991年8月16日收到。

1) 周林生. 试论雪岭云杉二元材积表的编制. 新疆林业科技文集, 1981, 第三辑第二分册, 26.

表1 雷岭云杉  $f_{1.5}$  随  $D$ 、 $H$  变化实测值

$D$	$H$																		
	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42
8	$\frac{0.594}{59}$	$\frac{0.576}{181}$	$\frac{0.545}{155}$	$\frac{0.528}{33}$															
12		$\frac{0.538}{86}$	$\frac{0.535}{221}$	$\frac{0.531}{308}$	$\frac{0.528}{196}$	$\frac{0.511}{49}$													
16		$\frac{0.553}{12}$	$\frac{0.508}{80}$	$\frac{0.513}{193}$	$\frac{0.514}{314}$	$\frac{0.508}{290}$	$\frac{0.509}{152}$	$\frac{0.508}{51}$	$\frac{0.516}{19}$										
20			$\frac{0.517}{23}$	$\frac{0.476}{86}$	$\frac{0.490}{204}$	$\frac{0.494}{294}$	$\frac{0.495}{298}$	$\frac{0.494}{228}$	$\frac{0.500}{149}$	$\frac{0.502}{30}$	$\frac{0.485}{12}$								
24				$\frac{0.460}{18}$	$\frac{0.475}{79}$	$\frac{0.464}{159}$	$\frac{0.475}{252}$	$\frac{0.480}{257}$	$\frac{0.480}{248}$	$\frac{0.482}{166}$	$\frac{0.484}{66}$	$\frac{0.508}{18}$							
28					$\frac{0.454}{12}$	$\frac{0.438}{58}$	$\frac{0.461}{95}$	$\frac{0.459}{157}$	$\frac{0.497}{198}$	$\frac{0.472}{202}$	$\frac{0.477}{180}$	$\frac{0.479}{82}$	$\frac{0.488}{44}$						
32						$\frac{0.439}{13}$	$\frac{0.441}{44}$	$\frac{0.432}{72}$	$\frac{0.452}{128}$	$\frac{0.458}{145}$	$\frac{0.460}{192}$	$\frac{0.465}{150}$	$\frac{0.462}{94}$	$\frac{0.479}{42}$					
36							$\frac{0.410}{26}$	$\frac{0.419}{39}$	$\frac{0.428}{54}$	$\frac{0.442}{84}$	$\frac{0.451}{110}$	$\frac{0.445}{138}$	$\frac{0.456}{120}$	$\frac{0.467}{101}$	$\frac{0.463}{46}$				
40								$\frac{0.408}{19}$	$\frac{0.417}{41}$	$\frac{0.426}{39}$	$\frac{0.425}{60}$	$\frac{0.425}{86}$	$\frac{0.442}{93}$	$\frac{0.444}{98}$	$\frac{0.451}{73}$	$\frac{0.457}{22}$			
44									$\frac{0.406}{13}$	$\frac{0.404}{24}$	$\frac{0.410}{29}$	$\frac{0.417}{44}$	$\frac{0.426}{54}$	$\frac{0.428}{86}$	$\frac{0.437}{71}$	$\frac{0.446}{41}$	$\frac{0.459}{14}$		
48										$\frac{0.398}{15}$	$\frac{0.397}{19}$	$\frac{0.397}{35}$	$\frac{0.405}{44}$	$\frac{0.414}{61}$	$\frac{0.421}{44}$	$\frac{0.427}{20}$	$\frac{0.445}{10}$	$\frac{0.478}{10}$	
52												$\frac{0.395}{13}$	$\frac{0.397}{16}$	$\frac{0.398}{25}$	$\frac{0.401}{42}$	$\frac{0.410}{32}$	$\frac{0.421}{16}$	$\frac{0.430}{10}$	
56													$\frac{0.370}{10}$	$\frac{0.400}{25}$	$\frac{0.416}{12}$	$\frac{0.426}{21}$			
60														$\frac{0.370}{13}$	$\frac{0.376}{13}$	$\frac{0.396}{13}$	$\frac{0.406}{9}$	$\frac{0.424}{12}$	

注: 表内分数为  $f_{1.5}/n$  的比值。

转换为胸高形数,再用普通最小二乘法求解模型参数,两者是等价的。将样本材积转化为形数以后,依据  $f_{1.3}$  随胸径和树高的变化信息,便可研究确定动态模型的合理结构,然后建模并求出复相关系数  $r$ 、剩余标准差  $s$ 、参数变动系数,并绘制残差分布图作为比较分析的基础。用以比较研究的模型选用山本式、分段模型和以往认为较好的 4 种多项式(含混合多项式)模型。

## 2 动态模型的建立及其与山本式的效果比较

$$V = C_0 \cdot D^{C_1} \cdot H^{C_2} \quad (1)$$

式中,  $V$ ——立木材积;  $C_0$ 、 $C_1$ 、 $C_2$ ——模型参数。

长期实践证明,就二元立木材积方程而言,山本式(1)具有较强的灵活性,它较好地体现了构成材积三要素之间的关系,用幂函数体现干形的变化规律,当  $C_1 > 2$ 、 $C_2 > 1$  为上升趋势,反之为下降趋势,参数可对干形的客观规律作出直观的解释。因而选用(1)式为建立形数动态模型的基本框架是恰当的。依据表 1 数据按(2)式用 Marquardt 迭代法求解参数建立一个总模型作为对比研究的基础。其参数值及有关统计指标列于表 2。模型的拟合残差为随机分布(图 1)。

$$f_{1.3} = a \cdot D^b \cdot H^c \quad (2)$$

( $a$ 、 $b$ 、 $c$ ——模型参数)

表 2 模型(2)统计指标

参 数 值	参数变动系数 (%)	标 准 差 和 复相关系数
$\hat{a} = 0.798\ 03$	2.5	$s = 0.017\ 3$
$\hat{b} = -0.211\ 02$	6.3	$r = 0.926\ 8$
$\hat{c} = 0.051\ 098$	31.1	

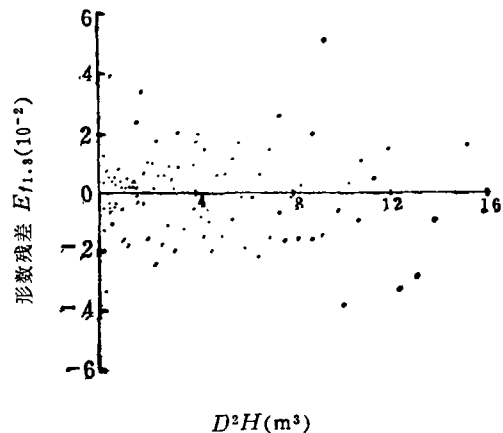


图 1 模型(2)残差分布

如表 1 所示,  $f_{1.3}$  随不同径阶和树高阶的变化率是不相同的,其中相同径阶  $f_{1.3}$  随树高的变化率,由低径阶到高径阶的变化规律是,先随树高的增加而下降逐步过渡到随树高的增加而升高。对于这种变化规律,模型(2)显然不能作出全面描述。要充分利用样本信息来提高模型精度,只有使参数  $b$ 、 $c$  随  $D$ 、 $H$  的变化而变。为确定其变化的相关表达式,现固定  $a = 0.798\ 03$ ,分别径阶和树高阶,用 Marquardt 迭代法求解出参数  $b$ 、 $c$  值列于表 3 和表 4,根据表 3 作散点图。

$b$ 、 $c$  与  $D$  呈线性相关,经拟合模型为

$$\begin{aligned} b &= 0.123\ 3 - 0.013\ 5D \quad (s = 0.101\ 5, r = 0.911\ 4) \\ c &= -0.306\ 6 + 0.014\ 8D \quad (s = 0.116\ 3, r = 0.903\ 5) \end{aligned} \quad (3)$$

式中,  $s$ ——标准差;  $r$ ——相关系数。

表 3 各不同径阶的参数  $b$ 、 $c$  值

项 目	8	12	16	20	24	28	32
$\hat{b}$	0.011	-0.103	-0.106	-0.149	-0.240	-0.266	-0.277
$\hat{c}$	-0.174	-0.063	-0.053	-0.011	0.084	0.116	0.126

项 目	36	40	44	48	52	56	60
$\hat{b}$	-0.346	-0.332	-0.374	-0.525	-0.386	-0.872	-0.762
$\hat{c}$	0.200	0.185	0.231	0.400	0.242	0.796	0.665

表 4 各不同树高阶的参数  $b$ 、 $c$  值

项 目	8	14	20	26	32	38
$\hat{b}$	-0.062	-0.151	-0.243	-0.267	-0.381	-0.472
$\hat{c}$	-0.106	-0.091	-0.083	0.110	0.235	0.340

表 4 中  $b$ 、 $c$  与  $H$  仍呈线性相关, 经拟合模型为

$$b = 0.0401 - 0.0132H \quad (s = 0.0214, r = 0.99);$$

$$c = -0.269 + 0.015H \quad (s = 0.068, r = 0.93) \quad (4)$$

模型(3)、(4)表明, 参数  $b$ 、 $c$  随  $D$  和  $H$  的变化率大致相同, 为了尽可能减少参数, 提高模型的稳定性, 其动态模型的合理结构确定为

$$f_{1.3} = C_0 \cdot D^{C_1 - C_2(D+H)} \cdot H^{C_3 + C_4(D+H)} \quad (5)$$

将表 1 数据再按 (5) 式建立雪岭云杉二元胸高形数动态模型, 其参数和统计指标值列于表 5。利用模型 (5) 求立木材积, 只需再乘以  $\pi D^2 H / 40\,000$  即可。模型的拟合残差分布如图 2 所示是随机的, 而且与图 1 相比, 显著缩小了残差波动范围。动态模型比常规山本式 (模型 (2)) 的标准差减少达一半之多。动态模型之所以具有显著提高精度的功能, 从图 3 和图 4 的对比中可以明显地看出, 在于其能够通过参数  $b$ 、 $c$  值的动态变化, 追踪样木的  $f_{1.3}$  随  $D$ 、 $H$  的客观变化规律, 而模型 (2) 由于只有一套固定不变的参数, 导致在低和高径阶两端部位严重偏离实测值 (图 3)。

表 5 动态模型 (5) 统计指标

参 数 值	参数变动系数 (%)	标准差和 复相关系数
$c_0 = 0.897\,304\,53$	3.1	$s = 0.008\,628$
$c_1 = 0.056\,158\,2$	32.0	
$c_2 = 0.006\,625\,791$	5.7	$r = 0.982\,3$
$c_3 = -0.291\,712$	7.4	
$c_4 = 0.007\,490\,851$	5.7	

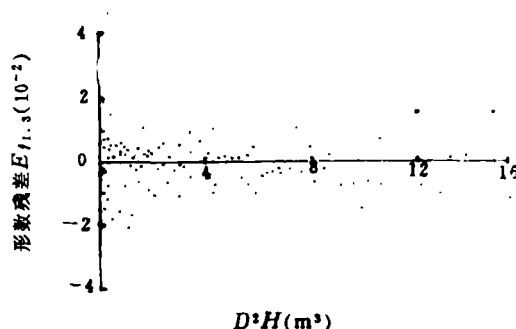


图 2 动态模型 (5) 残差分布

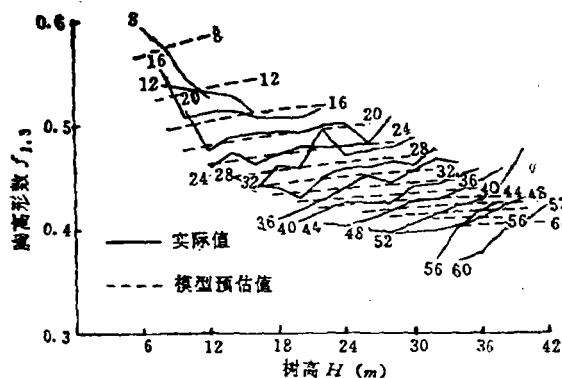


图3 模型(2)不同径阶  $f_{1.3}$  随  $H$  变化规律  
与实际变化值

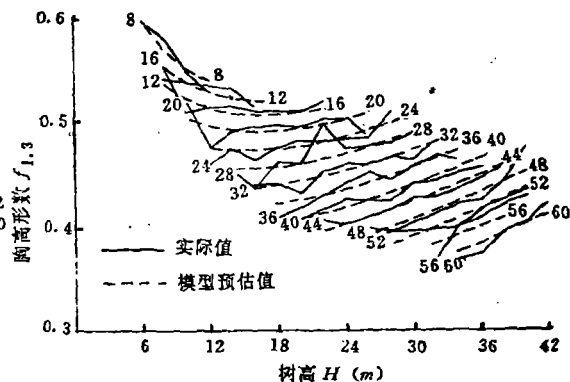


图4 模型(5)不同径阶  $f_{1.3}$  随  $H$  变化规律  
与实际变化值

### 3 动态模型与分段建模的比较研究

将表1数据按径阶8、12、16~20、24~40、44 cm 以上分成5段以(2)式建模。各分段模型间不强加处理连接点的限制条件。样本整体的拟合残差,由各分段模型对所属范围内的样本点实测值的预估所产生的残差组成,并以此与动态模型的预估残差对比。分段建模结果列于表6。

表6 各分段模型统计指标

分 段	参 数 值	参数变动系数(%)	标准差和复相关系数
8	$a = 0.989533$	6.7	$s = 0.0087383$
	$b = -0.092717$	9.4	$r = 0.956$
	$c = -0.173776$	17.2	
12	$a = 0.899750$	6.4	$s = 0.007238$
	$b = -0.151419$	4.8	$r = 0.727$
	$c = -0.063205$	39.3	
16~20	$a = 0.882838$	15.6	$s = 0.012508$
	$b = -0.161684$	34.8	$r = 0.675$
	$c = -0.0333087$	56.8	
24~40	$a = 0.825203$	5.5	$s = 0.00867$
	$b = -0.290254$	6.1	$r = 0.982$
	$c = 0.1290466$	9.6	
44~60	$a = 0.917201$	16.1	$s = 0.0103688$
	$b = -0.483004$	9.6	$r = 0.897$
	$c = 0.313041$	10.4	

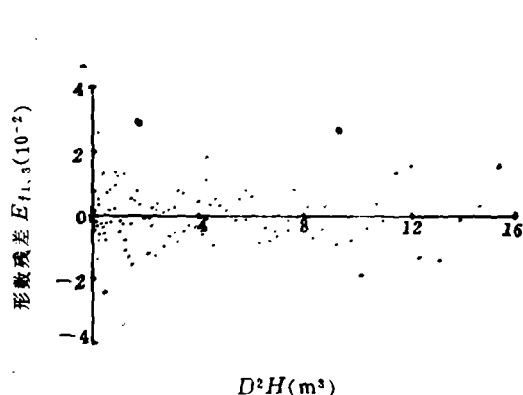


图5 分段模型残差分布

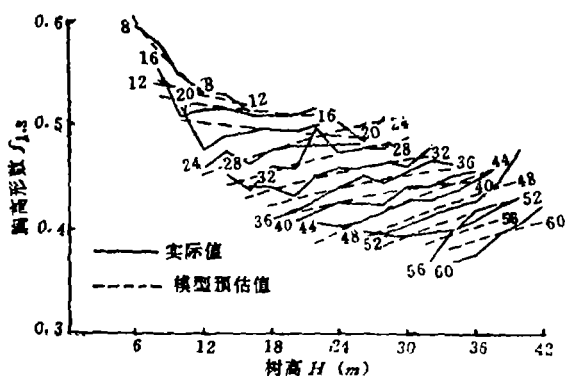
图6 分段模型径阶  $f_{1.3}$  随  $H$  变化规律与实际变化值

图5为分段拟合对整个样本实测点的预估残差分布,与图1相比,缩小了残差的波动范围,显著提高了精度;但与图2相比,其残差波动范围要大,精度提高效果不如动态模型好。

图6显示了分段建模不如动态模型对样本信息的利用充分,是精度较低的主要原因,同时还形象地揭示了分5段建模呈现出5种不同趋势的曲线簇,导致相邻模型间交叉的矛盾,解决这一矛盾虽可采取整体回归分析的方法,但由于必须使连接点处两相邻曲线相切而必然损失拟合精度,并可能导致模型的稳定性差。

#### 4 动态模型与几个多项式模型比较研究

选作对比研究的4个多项式模型是

$$V = C_0 + C_1 D^2 H + C_2 D^3 H + C_3 D^2 H \lg H \quad (6)$$

$$V = C_0 + C_1 D^2 + C_2 D^2 H + C_3 H^2 + C_4 D H^2 \quad (7)$$

$$V = C_0 + C_1 D + C_2 D^2 + C_3 D H + C_4 D^2 H \quad (8)$$

$$V = C_0 + C_1 D^2 H + C_2 D^3 H + C_3 D^2 + C_4 D^2 H \lg D \quad (9)$$

为了消除异方差影响并便于与动态模型比较,将上述方程变换为相应的形数模型

$$f_{1.3} = C_0 D^{-2} H^{-1} + C_1 + C_2 D + C_3 \lg H \quad (6')$$

$$f_{1.3} = C_0 D^{-2} H^{-1} + C_1 H^{-1} + C_2 + C_3 D^{-2} H + C_4 D^{-1} H \quad (7')$$

$$f_{1.3} = C_0 D^{-2} H^{-1} + C_1 D^{-1} H^{-1} + C_2 H^{-1} + C_3 D^{-1} + C_4 \quad (8')$$

$$f_{1.3} = C_0 D^{-2} H^{-1} + C_1 + C_2 D + C_3 H^{-1} + C_4 \lg D \quad (9')$$

利用(6'~9')式分别拟合表1样本数据的结果列于表7。可以看出,其剩余标准差分别为动态模型的1.79、1.35、2.27和1.96倍,远不如动态模型的精度高。其中模型(7')的剩余标准差虽相对较小,但参数  $C_0$  的变动系数高达227.8%,此外,模型(9')  $C_4$  的变动系数为83.6%,均远大于50%,这类含不稳定参数的病态模型是不能用于预估的。研究表明,利用多项式逼近的原理建立二元材积模型,当要求参数达到稳定标准时(变动系数不大于50%),对精度的提高是很有限的。

表7 多项式模型的统计指标

模 型 号	参 数 值	参数变动系数(%)	标准差和复相关系数
(6')	$C_0 = 41.145\ 715\ 482$	13.4	$s = 0.015\ 437\ 3$ $r = 0.94216$
	$C_1 = 0.443\ 344\ 949$	4.2	
	$C_2 = -0.003\ 220\ 867$	5.8	
	$C_3 = 0.037\ 308\ 546\ 8$	19.1	
(7')	$C_0 = -2.913\ 242\ 879$	227.8	$s = 0.011\ 674\ 8$ $r = 0.967\ 34$
	$C_1 = 2.007\ 238\ 722$	4.6	
	$C_2 = 0.211\ 591\ 118$	4.7	
	$C_3 = -1.094\ 447\ 545$	11.9	
(8')	$C_4 = 0.236\ 728\ 336$	5.6	$s = 0.019\ 587\ 4$ $r = 0.905\ 07$
	$C_0 = -316.989\ 359\ 131$	17.6	
	$C_1 = 54.555\ 213\ 869$	20.7	
	$C_2 = -1.922\ 969\ 403$	24.7	
(9')	$C_3 = 1.067\ 025\ 168$	26.8	$s = 0.016\ 950\ 8$ $r = 0.929\ 44$
	$C_4 = 0.432\ 488\ 351$	3.2	
	$C_0 = 28.048\ 387\ 853$	40.6	
	$C_1 = 0.625\ 814\ 071$	9.2	
(9')	$C_2 = -0.001\ 927\ 0396$	36.1	$s = 0.016\ 950\ 8$ $r = 0.929\ 44$
	$C_3 = -0.278\ 692\ 024$	49.0	
	$C_4 = -0.027\ 551\ 184$	83.6	

## 5 分析与结论

(1) 以山本式为基本框架的二元立木材积动态模型, 由于其参数能追踪样本所提供的形数随  $D$ 、 $H$  的不同变化率而改变, 充分利用了样本信息, 从而能将精度提高到一个新的水平。 $D$ 、 $H$  变幅越大, 干形变化率差异越显著的树种, 精度提高的效果愈显著。研究表明, 从剩余标准差、复相关系数、参数多少及其稳定性和可解释性以及残差分布的随机性等指标综合评价, 它优于已发表各类二元立木材积模型(其中含分段拟合模型)。

(2) 动态模型的结构通式为:  $V = C_0 D^{f_1(D,H)} \cdot H^{f_2(D,H)}$  或转换为形数通式:  $f_{1,3} = C_0' D^{f_1'(D,H)} \cdot H^{f_2'(D,H)}$ 。但其具体结构式, 会因树种的不同而异。例如本研究所确定的杉木结构式为

$$f_{1,3} = C_0 D^{C_1 - C_2(D+2H)} \cdot H^{C_3 + C_4(D+2H)} \quad (10)$$

经样本资料分析, 参数  $b$ 、 $c$  随  $D$  的变化率大致相当于随  $H$  变化率的 2 倍而不同于雪岭云杉其变化率大致相同的规律。因此, 动态模型具体结构式的确定, 必须以样本所提供的  $f_{1,3}$  随  $D$ 、 $H$  的客观变化规律为依据, 使模型中的每一个参数具有可解释性, 这样才能保证模型有尽可能少的参数和稳定的预测效果。否则, 脱离样本的具体规律, 仅依多项式逼近的数学原理确定模型结构, 同样会导致模型病态而失败。例如本研究曾以(11)式对表 1 数据建模, 其结果如表 8。

$$f_{1,3} = C_0 D^{(C_1 + C_2 H + C_3 D + C_4 DH)} \cdot H^{(C_5 + C_6 D + C_7 H + C_8 DH)} \quad (11)$$

将表 8 与动态模型(5)的建模结果(表 5)进行比较, 便充分暴露了模型(11)盲目采用多项式逼近原理确定动态模型结构的弊端, 模型(11)比模型(5)增加了 4 个参数, 不但未提高精

表 8 模型(11)统计指标

参 数 值	参数变动系数 (%)	标 准 差 和 复相关系数
$C_0 = 0.886\ 112\ 325$	8.1	$s = 0.608\ 681\ 37$
$C_1 = -0.059\ 684\ 566$	270.1	
$C_2 = 0.020\ 468\ 331$	151.6	
$C_3 = 0.001\ 540\ 237$	766.0	
$C_4 = 0.000\ 287\ 016$	111.4	$r = 0.982\ 077$
$C_5 = -0.168\ 705\ 232$	83.1	
$C_6 = -0.012\ 114\ 496$	207.4	
$C_7 = -0.008\ 166\ 253$	207.9	
$C_8 = -0.000\ 290\ 891$	114.3	

度,反而使 9 个参数中 8 个参数的变动系数大大超过 50%,其中  $C_3$  的变动系数高达 766%,导致模型极不稳定。

(3) 在变量间关系可以依据样本所提供的信息予以明确的条件下,应尽可能避免盲目采用纯实验性的多项式模型,当必须采用多项式建模时,应控制模型不含变动系数大于 50% 的参数,保证模型有稳定的预测效果。

(4) 二元材积方差是与  $D^4H^2$  成比例的,建模中应采取消除异方差影响的措施。将材积转化为胸高形数的变量转换方法,既可起

到消除异方差影响的作用,又便于进行干形变化率分析,是一个一举两得的有效措施。

(5) 统计指标  $r$ 、 $s$  不能作为评价模型优劣的唯一依据,因为两者均不能揭示模型的残差是否有偏和模型是否稳定。就二元立木材积模型而言,还应增加残差的随机分布,参数的稳定性和可解释性三项指标。

(6) 本研究只涉及到二元立木材积模型的选择和拟合问题,对于影响二元材积模型质量的样本构成及建模后的检验问题,有待另文探讨。

### 参 考 文 献

- 1 Alder D. Forest volume estimation and yield prediction. Rome FAO of UN, 1980, 1~2.
- 2 李卓立. 经济计量学. 北京: 清华大学出版社, 1987.

## A Study on Standard Volume Dynamic Model

Luo Qibang Ning Hui He Dongbei Jiang Jusheng

Wu Zhide Zeng Weisheng Zhang Shuguang

(The Central South Design Institute of Forest Inventory and Planning Ministry of Forestry)

**Abstract** This paper took *Picea schrenkiana* and *Cunninghamia lanceolata* as study objects and aimed at improving the precision of the standard volume model. The study provided a dynamic model and its fitting method which can use the sample information completely. The model not only can improve the precision significantly but also can explain the objective rules. The model has stable and reliable characteristics for application.

**Key words** standard volume dynamic model unequal variance parameter stability error distribution diagram