

利用线性混合模型和哑变量模型方法建立 贵州省通用性生物量方程

曾伟生¹, 唐守正¹, 夏忠胜², 朱松², 罗洪章²

(1. 中国林业科学研究院资源信息研究所,北京 100091; 2. 贵州省森林资源管理站,贵州 贵阳 550001)

摘要:以贵州省人工杉木和马尾松地上生物量数据为例,通过利用线性混合模型和哑变量模型方法,建立了适合不同树种和区域(中心区和一般区)的通用性立木生物量方程,为简化生物量建模工作提供了有效途径。结果表明,相同直径林木的地上生物量估计值随树种、区域的不同存在一定程度的差异,带随机参数的线性混合模型和带特定参数的哑变量模型比总体平均模型的精度高;线性混合模型和哑变量模型方法均同等有效,可推广应用于其它通用性模型(如材积方程)的建立。

关键词:地上生物量;线性混合模型;哑变量模型;通用性方程;杉木;马尾松

中图分类号:S711

文献标识码:A

Using Linear Mixed Model and Dummy Variable Model Approaches to Construct Generalized Single – Tree Biomass Equations in Guizhou

ZENG Wei-sheng¹, TANG Shou-zheng¹, XIA Zhong-sheng², ZHU Song², LUO Hong-zhang²

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

2. Forest Resources Administration Office of Guizhou Province, Guiyang 550001, Guizhou, China)

Abstract: Based on the above-ground biomass data of Chinese fir (*Cunninghamia lanceolata*) and Masson Pine (*Pinus massoniana*) plantations in Guizhou Province, the generalized single-tree biomass equations suitable for different species and regions (central region and other region) were established using linear mixed model and dummy variable model methods, which provided effective approaches to simplify the biomass modeling. The results show that the above-ground biomass estimates of individual trees with the same diameter are different in some extent among different tree species and geographic regions, and linear mixed model with random parameters and dummy model with specific (local) parameters are better than population average model; and linear mixed model and dummy variable model approaches are almost same effective and may be applied to develop other generalized models such as tree volume equations.

Key words: above-ground biomass; linear mixed model; dummy variable model; generalized equation; compatibility; *Cunninghamia lanceolata*; *Pinus massoniana*

由于森林生态系统在调节全球碳平衡、减缓大气温室气体浓度上升等方面具有不可替代的作用,世界各国越来越重视对森林生物量的监测^[1]。开展森林生物量监测,建立适合大尺度范围的通用性立木生物量模型是一项重要的基础工作。近年来,有

不少学者在积极探索建立适合国家、区域乃至全球水平的通用性立木生物量模型^[2-10]。Case & Hall^[9]对加拿大中西部地区 10 个树种的不同尺度(地方的、区域通用的、全国通用的)材积和生物量方程进行比较后发现,随着方程通用性水平的提高,用于局

收稿日期:2011-01-05

基金项目:贵州省林业厅科研项目[2006]25号“贵州省人工杉木马尾松生物量通用模型研究”

作者简介:曾伟生(1966—),男,湖南涟源人,在读博士,教授级高工,主要从事森林资源监测和林业数表研制工作。

部地区的预估误差也相应增大。我国目前正在积极推进全国通用性立木生物量模型的研建工作。在建立通用性立木生物量模型的时候,如何在保证预估精度的前提下尽量简化所建模型,是值得研究的一个问题,而混合模型和哑变量模型方法为解决这一问题提供了可能途径。

混合模型方法是改进参数估计的一种实用统计技术,近20年来在很多研究领域得到了应用,而在林业方面的应用则相对较迟。Lappi & Bailey^[11]阐述了基于Richards模型的非线性混合效应生长曲线的使用,用来在样地和样木两个水平上预估优势木和亚优势木的树高;Gregoire等^[12]以2个固定样地数据为基础,建立了考虑重复测定数据之间的协方差和样地随机效应的线性混合模型;Zhang & Borders^[13]为了提高对美国乔治亚州集约经营的火炬松林立木生物量的估计精度,在建立立木生物量方程时采用了混合模型方法;Calegario等^[14]建立了巴西沿海地区人工桉树林按不同无性系和不同立地的优势高非线性混合生长模型,并认为非线性混合模型方法是一种灵活而精准的方法;Meng等^[15]基于非线性混合模型方法,建立了加拿大艾伯塔省黑松按不同树冠级和不同风速影响下的树高-直径关系模型;Fehrmann等^[16]在建立芬兰的挪威云杉和欧洲赤松立木生物量模型时,对混合模型方法与k-最近邻域方法进行了对比分析。国内近年来关于混合模型在林业方面的研究比较具代表性的有:考虑样地和树冠类型随机影响的林木直径与航片冠幅线性混合模型^[17]、考虑样地及样木之间随机影响的单木断面面积生长混合模型^[18],以及考虑区组间随机效应的杉木林优势木平均高非线性混合模型^[19]。

哑变量也称虚拟变量,通常取值为0、1或-1。哑变量处理是对定性因子或分类变量进行处理的一种常用方法,统计分析中的各种数量化方法都要涉

及到哑变量处理的问题^[20-22];在各种回归分析和建模实践中也经常用到哑变量模型方法^[23-24]。关于哑变量方法与混合模型方法究竟哪种更合适,Wang等^[25]以建立优势高模型为例进行了对比分析,认为两种方法都适用于建立带有特定参数或地方参数的模型,而且二者的效果基本相当;就拟合效果而言,哑变量方法可能略好;就预估效果而言,混合模型方法更合适。

本文从建立通用性立木生物量模型的实际需求出发,以贵州省的人工杉木(*Cunninghamia lanceolata* (Lamb.) Hook.)和马尾松(*Pinus massoniana* Lamb.)数据为基础,利用哑变量模型和线性混合模型方法,建立适合不同树种和区域(中心区和一般区)的通用性立木地上生物量模型,并将以树种和区域作为特定参数的通用性模型与不分树种和区域的总体平均模型进行比较,分析不同树种、区域对立木生物量估计的影响,为全国和各省森林生物量监测提供参考依据。

1 数据

本文所用数据为贵州省的人工杉木和马尾松立木地上生物量实测数据,共694株样木,采集时间为2007年,采集地点涉及全省杉木和马尾松分布区,具有广泛的代表性。其中,杉木的样木数为399,中心区149株,一般区250株;马尾松的样木数为295,中心区148株,一般区147株。全部样木都实测胸径和冠幅,将样木伐倒后,测量树干长度(树高)和活树冠长度(冠长),分干材、干皮、树枝、树叶称鲜质量,并分别抽取样品带回实验室,在85℃恒温下烘干至恒质量,根据样品鲜质量和干质量分别推算样木各部分干质量并汇总得到地上部分干质量。样木数据的有关统计指标见表1。

表1 贵州省立木地上生物量数据统计指标

| 树种 | 区域 | 样本量 | 胸径/cm | | | | 地上生物量/kg | | | |
|-----|-----|-----|-------|-----|-----|------|----------|---------|-------|---------|
| | | | 平均值 | 标准差 | 最小值 | 最大值 | 平均值 | 标准差 | 最小值 | 最大值 |
| 杉木 | 中心区 | 149 | 15.8 | 6.8 | 4.2 | 36.4 | 73.795 | 77.305 | 1.329 | 529.992 |
| | 一般区 | 250 | 15.1 | 6.7 | 4.1 | 36.3 | 70.364 | 78.726 | 1.619 | 508.733 |
| 马尾松 | 中心区 | 148 | 16.9 | 8.0 | 4.3 | 39.3 | 131.135 | 144.697 | 2.656 | 612.402 |
| | 一般区 | 147 | 15.0 | 8.1 | 4.0 | 44.8 | 105.885 | 137.881 | 3.110 | 808.978 |

2 方法

林木生物量涉及干材、干皮、树枝、树叶、树根等

各个分项,但作为大尺度森林生物量监测,关注的主要是总生物量尤其是地上生物量^[1,4,8-10,26]。立木生物量模型的常用结构形式为:

$$y = \beta_0 x_1^{\beta_1} \cdots x_j^{\beta_j} + \varepsilon \quad (1)$$

式中, y 为林木总生物量或分项生物量, x_j 为反映林木大小的变量, β_j 为模型参数, ε 为误差项。

2.1 基本模型

因为幂函数形式的一元立木生物量方程在全世界得到了广泛应用^[4,8-10,27],本文即以以下一元方程作为基本的生物量模型:

$$M = aD^b(1 + \zeta) \quad (2)$$

式中, M 为地上生物量, a 、 b 为参数, ζ 为相对误差。将(2)式通过对数转换为如下线性形式:

$$y = a_0 + bx + \zeta \quad (3)$$

式中 $y = \ln M$, $x = \ln D$, $a_0 = \ln a$, $\zeta = \ln(1 + \zeta)$ 。根据(3)式的拟合结果,可由下式得到生物量的估计值:

$$\hat{M} = \exp(a_0 + bx) \quad (4)$$

但是由于对数转换存在偏差,还必需对(4)式进行修正,其中最常用的校正因子为 $\exp(S^2/2)$ ^[28-29],这时校正后的生物量估计值为:

$$\hat{M} = \exp(a_0 + S^2/2) \times D^b \quad (5)$$

此外,从实用性角度出发,还可以采用Snowdon^[30]提出的比值估计量校正因子,确保其总体平均偏差为0。

林木地上生物量除了随直径 D 变化以外,还会受到其它一些因素的影响,如树种、区域分布等。这里把只考虑直径 D 这一个变量、具有2个通用参数(或固定参数)的模型(3)称为总体平均模型。下面再来考虑不同树种和不同区域条件下的哑变量模型和混合模型。因为(3)式中的参数 b 变动很小,甚至有人提出可以直接采用统一的固定值^[3,31],因此不同树种和不同区域条件下的局部效应或随机效应主要考虑对参数 a_0 的影响。

树种分为杉木和马尾松2种,代码分别为1和2,样本单元数分别为399和295。区域尽管只分中心区和一般区,但因为2个树种的中心区和一般区在地域上并不等同,因此将其作为4个区域对待,按杉木中心区、杉木一般区、马尾松中心区、马尾松一般区,类型代码分别为1、2、3、4,其样本单元数分别为149、250、148和147。

2.2 哑变量模型

基于(3)式的哑变量模型的基本形式可以表示为:

$$y = a_0 + \sum a_i z_i + bx + \zeta \quad (6)$$

式中, z_i 为哑变量, a_i 为相应的特定参数或局部参数,其它符号同(3)式。为了区别起见,在哑变量模

型中把对应于总体平均模型的参数称为通用参数或全局参数。

哑变量的处理包括以下2种情况:

1)考虑树种:仅含1个哑变量 z_1 ,当树种为杉木时,取 $z_1 = 1$;当树种为马尾松时,取 $z_1 = -1$ 。

2)考虑区域:因为区域与树种并不独立,故相当于同时考虑了区域和树种,包含3个哑变量 z_2 、 z_3 和 z_4 :

当区域类型为1(树种为杉木)时,取 $z_2 = 1$, $z_3 = 0$, $z_4 = 0$;

当区域类型为2(树种为杉木)时,取 $z_2 = 0$, $z_3 = 1$, $z_4 = 0$;

当区域类型为3(树种为马尾松)时,取 $z_2 = 0$, $z_3 = 0$, $z_4 = 1$;

当区域类型为4(树种为马尾松)时,取 $z_2 = -1$, $z_3 = -1$, $z_4 = -1$ 。

由于(6)式属于典型的线性方程,可以直接采用普通最小二乘法求解参数。

2.3 线性混合模型

线性混合模型的基本形式为^[20,22]:

$$y = x\beta + zu + e \quad (7)$$

式中, y 为因变量, β 为固定参数, u 为随机参数, x 为固定参数设计矩阵, z 为随机参数设计矩阵, e 为观测误差。

与(6)式对应的混合模型可以表示为:

$$y = a_0 + \sum u_i z_i + bx + \zeta \quad (8)$$

其中,随机参数 u_i 的数学期望值为0,且相互独立,即:

$$E(u_i) = 0, \text{cov}(u_i, u_j) = 0, i \neq j$$

对应于哑变量模型的前述2种情况,混合模型也作如下考虑:

1)将树种作为随机效应变量;

2)将区域(和树种)作为随机效应变量。

对于(7)或(8)式的线性混合模型,可以利用ForStat软件(V2.1)“统计分析”功能模块中的“线性混合模型”求解参数,其中隐含的参数估计方法为限制性极大似然法^[22]。

2.4 模型评价

哑变量模型与混合模型之间的对比评价指标采用确定系数 R^2 、残差平方和 SSE 和均方误差 S^2 。其计算公式如下:

$$R^2 = 1 - \sum (y_i - \hat{y}_i)^2 / \sum (y_i - \bar{y})^2 \quad (9)$$

$$SSE = \sum (y_i - \hat{y}_i)^2 \quad (10)$$

$$S^2 = SSE/(n-p) \quad (11)$$

式中, y_i 和 \hat{y}_i 分别为第 i 株样木的实测值和预估值, \bar{y} 为全部样木实测值的平均数, n 为样木总数, p 为参数个数。

哑变量模型或混合模型与总体平均模型之间的差异对比采用 F 统计检验。 F 统计指标按下式计算^[15]:

$$F = \frac{(SSE_{PA} - SSE_{DM})/(df_{PA} - df_{DM})}{SSE_{DM}/df_{DM}} \quad (12)$$

式中, SSE_{PA} 、 df_{PA} 分别为总体平均模型的残差平方和与自由度, SSE_{DM} 、 df_{DM} 分别为哑变量模型或混合模型的残差平方和与自由度。根据 F 值的大小可判定两个模型之间是否有显著差异, 从而对是否需要建立适用于特定树种、区域的立木生物量模型作出评判。

3 结果与分析

利用贵州省 694 株杉木和马尾松样木的地上生

物量数据, 采用 ForStat 软件先对总体平均模型(3)进行拟合, 再对前述 4 种情况下的哑变量模型(6)和混合模型(8)进行拟合, 并对其结果进行分析。

3.1 总体平均模型

贵州省杉木和马尾松地上生物量对数转换形式的总体平均模型为:

$$y = -2.5088 + 2.4128x \quad (R^2 = 0.9265, \\ SSE = 82.7691, S^2 = 0.1196, \\ F = 8724.90, P = 0.0000) \quad (13)$$

其中, 参数 a_0 和 b 的 t 值分别为 -36.28 和 93.41 , 在统计上都极其显著; S^2 为均方误差, F 为模型的统计显著性指标, P 为显著水平。

对应于(5)式的地上生物量估计公式为:

$$\hat{M} = \exp(-2.5088 + 0.0598) \times D^{2.4128}$$

该式即为不分树种的用于全省森林生物量估计的通用性平均模型。

3.2 哑变量模型

前述 2 种情况下的哑变量模型(依次称为模型 1、2, 下同)拟合结果见表 2。

表 2 哑变量模型(6)的拟合结果

| 哑变量模型 | 通用参数 | | 特定参数 | | | | 拟合统计指标 | | | | |
|-------|---------|--------|---------|----------|--------|----------|---------|--------|--------|---------|--------|
| | a_0 | b | a_1 | a_2 | a_3 | a_4 | F | P | R^2 | SSE | S^2 |
| 模型 1 | -2.4769 | 2.4091 | -0.1476 | (0.1476) | / | / | 5377.59 | 0.0000 | 0.9396 | 67.9967 | 0.0984 |
| 模型 2 | -2.4893 | 2.4126 | -0.1771 | -0.1247 | 0.1325 | (0.1693) | 2696.00 | 0.0000 | 0.9399 | 67.6415 | 0.0982 |

注: 通用参数和特定参数全部在 0.01 水平上极显著; 括号中的数值为非独立的特定参数。

从表 2 来看, 模型 1 和模型 2 的各项统计指标差异很小。另外, 通过利用(12)式, 对上述 2 个哑变量模型与总体平均模型进行差异显著性检验, 计算得到的 2 个 F 值分别为 150.12 和 51.36, 均在 0.01 水平上与总体平均模型有显著差异。而用同样的方法对 2 个哑变量模型进行差异显著性检验, 计算得到的 F 值为 1.80, 差异不显著。说明不同树种的生

物量估计值存在显著差异, 应该采用不同的特定参数; 而同一树种不同区域之间无显著差异, 分别区域用不同的特定参数意义不大。

3.3 线性混合模型

前述 2 种情况下的线性混合模型拟合结果见表 3。

表 3 线性混合模型(8)的拟合结果

| 混合模型 | 固定参数 | | 随机参数 | | | | 随机效应分析 | | 拟合统计指标 | | |
|------|---------|--------|---------|---------|--------|--------|---------|-----|--------|---------|--------|
| | a_0 | b | u_1 | u_2 | u_3 | u_4 | F | 显著性 | R^2 | SSE | S^2 |
| 模型 1 | -2.4771 | 2.4091 | -0.1466 | 0.1466 | / | / | 3.9470 | ** | 0.9396 | 67.9974 | 0.0984 |
| 模型 2 | -2.4895 | 2.4126 | -0.1731 | -0.1228 | 0.1299 | 0.1660 | 33.0040 | *** | 0.9399 | 67.6473 | 0.0982 |

从表 3 结果可知, 对参数 a_0 的随机效应, 杉木为负, 马尾松为正(二者刚好是正负相抵的), 说明相同直径林木的地上生物量, 马尾松要大于杉木, 而且这种随机效应在统计上的差异是非常显著的。从区域来看, 相同直径林木的地上生物量, 2 个树种都

是中心区略小于一般区, 但不存在显著差异。参数随树种和区域的总体变化规律与哑变量模型是完全一致的。从 3 项统计指标 R^2 、 SSE 和 S^2 来看, 混合模型与哑变量模型之间的差异都很小。

另外, 通过利用(12)式, 对上述 2 个混合模型与

总体平均模型进行差异显著性检验,计算得到的2个F值分别为153.54和52.48,均在0.01水平上与总体平均模型有显著差异。而用同样的方法对2个混合模型进行差异显著性检验,计算得到的F值为1.78,差异不显著。检验结论与哑变量模型完全相同。

3.4 两类模型的对比分析

哑变量模型中特定参数的大小通常是以某一类型的响应值为0值作为基准来处理 and 估计的,本研究为了更好地与混合模型进行对照,将通常的(1,0)赋值方式调整为(1,-1)赋值方式。从而在哑变量模型中,各个类型的特定参数响应值之和也等于0,与混合模型完全类似。如对于模型1,在哑变量模型中杉木和马尾松这2个树种的特定参数估计值分别为-0.1476和0.1476,而在混合模型中其随机参数估计值分别为-0.1466和0.1466。二者不但反映了同样的规律(马尾松大于杉木),而且在程度上也差异很小。模型2也存在完全类似的情况。也就是说,在本研究中,混合模型所反映出来的各个类型之间的差异与哑变量模型几乎一致。另外,不论是模型1还是模型2,哑变量模型和混合模型中的通用参数或固定参数b取5位有效数时均完全相同,这也进一步印证了b值的稳定性。

为了反映哑变量模型和混合模型与总体平均模

型之间的差异,以及不同类型(树种、区域)模型之间的相容性,表4列出了前述2种情况下不同类型模型对地上生物量(对数值)的估计结果对比,其中模型0指总体平均模型(13)。

从表4可以看出,尽管总体平均模型对总体的估计值是无偏的(总体相对误差为0),但对不同树种、不同区域的估计值,会产生将近±5%的误差。通过拟合2种情况下的哑变量模型和混合模型,可以显著减少对各个类型的估计误差,而且两类模型的估计效果几乎相当,其中模型1的相对误差均在±1%以内,模型2则在±0.1%以内。哑变量模型2相当于为树种和区域的4种组合建立了4个特定参数的模型,且每一个特定参数都作为固定参数对待,从而4种类型的估计值及按树种的合计均不存在误差;而混合模型2把树种和区域的4种组合的影响视为随机效应,每一个特定参数都作为随机参数对待,从而4种类型的估计值及按树种的合计均还存在约±0.1%的误差,但这基本上可忽略不计。

另外,不论是哑变量模型还是混合模型,分别树种、区域进行估计后的合计值与总体平均模型的估计值结果是一样的。也就是说,汇总各树种或区域用特定模型得到的估计值,其全省尺度的总估计值与用单一的总体平均模型得到的估计值是相等的,从而保证了不同模型之间的相容性。

表4 不同类型(树种、区域)模型的估计结果对比

| 模型 | 类型 | 实测值 | 模型估计值 | | | 相对误差/% | | |
|-------|------|----------|----------|----------|----------|--------|-------|-------|
| | | | 模型0 | 模型1 | 模型2 | 模型0 | 模型1 | 模型2 |
| 哑变量模型 | 杉木 | 1 473.70 | 1 523.75 | 1 473.70 | 1 473.70 | 3.40 | 0.00 | 0.00 |
| | 马尾松 | 1 186.11 | 1 136.06 | 1 186.11 | 1 186.11 | -4.22 | 0.00 | 0.00 |
| | 区域1 | 555.44 | 579.03 | 560.32 | 555.44 | 4.25 | 0.88 | 0.00 |
| | 区域2 | 918.26 | 944.72 | 913.38 | 918.26 | 2.88 | -0.53 | 0.00 |
| | 区域3 | 618.66 | 596.28 | 621.34 | 618.66 | -3.62 | 0.43 | 0.00 |
| | 区域4 | 567.45 | 539.79 | 564.77 | 567.45 | -4.88 | -0.47 | 0.00 |
| | (合计) | 2 659.81 | 2 659.81 | 2 659.81 | 2 659.81 | 0.00 | 0.00 | 0.00 |
| 混合模型 | 杉木 | 1 473.70 | 1 523.75 | 1 474.03 | 1 474.65 | 3.40 | 0.02 | 0.06 |
| | 马尾松 | 1 186.11 | 1 136.06 | 1 185.78 | 1 185.16 | -4.22 | -0.03 | -0.08 |
| | 区域1 | 555.44 | 579.03 | 560.45 | 556.00 | 4.25 | 0.90 | 0.10 |
| | 区域2 | 918.26 | 944.72 | 913.59 | 918.66 | 2.88 | -0.51 | 0.04 |
| | 区域3 | 618.66 | 596.28 | 621.18 | 618.24 | -3.62 | 0.41 | -0.07 |
| | 区域4 | 567.45 | 539.79 | 564.60 | 566.91 | -4.88 | -0.50 | -0.09 |
| | (合计) | 2 659.81 | 2 659.81 | 2 659.81 | 2 659.81 | 0.00 | 0.00 | 0.00 |

注:表中划方框数字所在行代表的类型就是模型设置特定参数或随机参数所考虑的类型。

4 结论与讨论

本文以贵州省人工杉木和马尾松地上生物量数据为例,通过利用哑变量模型和线性混合模型方法,建立了适合不同树种和区域的全省通用性立木生物量模型,为简化生物量建模进行了有益尝试。从模型反映的结果看,相同直径林木的地上生物量估计值随树种、区域的不同存在一定程度的差异。从树种看,马尾松的地上生物量要明显大于杉木;从区域看,一般区的林木地上生物量要略大于中心区,但差异不明显。

另外,究竟应该选择哑变量模型还是混合模型方法,在生物计量学和统计学领域一直以来都还存在着争论^[25]。Wang 等^[25]认为,从实用性角度出发,一般可以根据划分类型的数量及每一类型的样本量来决定:如果类型数量少(如 10 个以下),则选择哑变量模型可能较好;如果类型数量多,而且每个类型包含的样本量又少,则一般推荐混合模型;如果每个类型的样本量都较大,则选择哪个模型都无关紧要了。针对本文的例子,因为划分的类型少,而且每个类型的样本量都很大,因此,哑变量模型和线性混合模型的预估结果几乎等同,这也印证了上述观点。

总之,混合模型和哑变量模型方法为简化立木生物量建模工作提供了有效途径,其它通用性模型(如材积模型)的建立也可以采用这一方法。

参考文献:

- [1] Tomppo E, Gschwantner T, Lawrence M, *et al.* National Forest Inventories: Pathways for Common Reporting [M]. Springer, New York, 2010
- [2] Hansen M. Volume and biomass estimation in FIA: National consistency vs. regional accuracy [R]. GTR NC-230, North Central Research Station, Forest Service USDA, 2002: 109-120
- [3] Chojnacky D C. Allometric scaling theory applied to FIA biomass estimation [R]. GTR NC-230, North Central Research Station, Forest Service USDA, 2002: 96-102
- [4] Jenkins J C, Chojnacky D C, Heath L S, *et al.* National-scale biomass estimators for United States tree species [J]. *For Sci*, 2003, 49(1): 12-35
- [5] Snorrason A, Einarsson S F. Single-tree biomass and stem volume functions for eleven tree species used in Icelandic forestry [J]. *Icelandic Agric Sci*, 2006, 19: 15-24
- [6] Vallet P, Dhôte J-F, Le Moguédec G, *et al.* Development of total aboveground volume equations for seven important forest tree species in France [J]. *Forest Ecology and Management*, 2006, 229: 98-110
- [7] Repola J, Ojansuu R, Kukkola M. Biomass functions for Scots pine, Norway spruce and birch in Finland [R/OL]. Working Papers of the Finnish Forest Research Institute 53, 2007, http://www.metla.fi/julkaisut/working_papers/2007/mwp053.htm
- [8] Muukkonen P. Generalized allometric volume and biomass equations for some tree species in Europe [J]. *Eur J Forest Res*, 2007, 126: 157-166
- [9] Case B, Hall R J. Assessing prediction errors of generalized tree biomass and volume equations for the boreal forest region of west-central Canada [J]. *Can J For Res*, 2008, 38: 878-889
- [10] Návár J. Allometric equations for tree species and carbon stocks for forests of northwestern Mexico [J]. *Forest Ecology and Management*, 2009, 257: 427-434
- [11] Lappi J, Bailey R L. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods [J]. *For Sci*, 1988, 38(2): 409-429
- [12] Gregoire T G, Schabenberger O, Barrett J P. Linear modeling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements [J]. *Can J For Res*, 1995, 25: 137-156
- [13] Zhang Y J, Borders B E. Using a system mixed-effects modeling method to estimate tree compartment biomass for intensively managed loblolly pines—an allometric approach [J]. *Forest Ecology and Management*, 2004, 194: 145-157
- [14] Calegario N, Daniels R F, Maestri R, *et al.* Modeling dominant height growth based on nonlinear mixed-effects model: a clonal *Eucalyptus* plantation case study [J]. *Forest Ecology and Management*, 2005, 204: 11-20
- [15] Meng S X, Huang S, Lieffers V J, *et al.* Wind speed and crown class influence the height-diameter relationship of lodgepole pine: Nonlinear mixed effects modeling [J]. *Forest Ecology and Management*, 2008, 256: 570-577
- [16] Fehrmann L, Lehtonen A, Kleinn C, *et al.* Comparison of linear and mixed-effect regression models and a k-nearest neighbor approach for estimation of single-tree biomass [J]. *Can J For Res*, 2008, 38: 1-9
- [17] 郎璞玫. 航空像片冠幅与地面直径的线性混合模型 [J]. *林业科学*, 2008, 44(3): 41-44
- [18] 雷相东, 李永慈, 向 玮. 基于混合模型的单木断面面积生长模型 [J]. *林业科学*, 2009, 45(1): 74-80
- [19] 李春明, 张会儒. 利用非线性混合模型模拟杉木林优势木平均高 [J]. *林业科学*, 2010, 46(3): 89-95
- [20] 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
- [21] 李丽霞, 郜艳晖, 张 瑛. 哑变量在统计分析中的应用 [J]. *数理医药学杂志*, 2006, 19(1): 51-52
- [22] 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, 2008
- [23] 李希菲, 洪玲霞. 用哑变量法求算立地指数曲线族的研究 [J]. *林业科学研究*, 1997, 10(2): 215-219
- [24] 李 河, 麦劲壮, 肖 敏, 等. 哑变量在 Logistic 回归模型中的应用 [J]. *循证医学*, 2008, 8(1): 42-45

- [25] Wang M, Borders B E, Zhao D. An empirical comparison of two subject-specific approaches to dominant heights modeling, the dummy variable method and the mixed model method [J]. *Forest Ecology and Management*, 2008, 255: 2659 – 2669
- [26] FAO. Global forest resources assessment 2005: progress towards sustainable forest management [R]. Rome: FAO Forestry Paper 147, Food and Agriculture Organization of the United Nations, 2006
- [27] Ter-Mikaelian M T, Korzukhin M D. Biomass equations for sixty-five North American tree species [J]. *Forest Ecology and Management*, 1997, 97: 1 – 24
- [28] Baskerville G L. Use of logarithmic regression in the estimation of plant biomass [J]. *Can J For Res*, 1972, 2: 49 – 53
- [29] Flewelling J W, Pienaar L V. Multiplicative regression with lognormal errors [J]. *For Sci*, 1981, 27(2): 281 – 289
- [30] Snowdon P. A ratio estimator for bias correction in logarithmic regressions [J]. *Can J For Res*, 1991, 21(5): 720 – 724
- [31] West G B, Brown J H, Enquist B J. A general model for the structure and allometry of plant vascular systems [J]. *Nature*, 1999, 400: 664 – 667