

DOI: 10.5846/stxb201403310601

符利勇,唐守正,张会儒,张则路,曾伟生.东北地区两个主要树种地上生物量通用方程构建.生态学报,2015,35(1):0150-0157.

Fu L Y, Tang S Z, Zhang H R, Zhang Z L, Zeng W S. Generalized above-ground biomass equations for two main species in northeast China. Acta Ecologica Sinica, 2015, 35(1): 0150-0157.

东北地区两个主要树种地上生物量通用方程构建

符利勇¹, 唐守正^{1, *}, 张会儒¹, 张则路², 曾伟生³

1 中国林业科学研究院资源信息研究所, 北京 100091

2 吉林省和龙林业局 133500

3 国家林业局调查规划设计院, 北京 100714

摘要:目前,东北落叶松地上生物量方程主要采用分树种或把不同树种归为一体的方法,但是,既能反映落叶松生物量与自变量的平均关系,又能反映不同树种间生物量差异程度的通用性落叶松生物量方程迄今尚未构建。因此,以东北地区兴安落叶松和长白落叶松地上生物量数据为研究对象,构建一元(自变量为胸径)、二元(自变量为胸径和树高)和三元(自变量为胸径、树高和冠幅)的不同树种生物量通用方程。由于起源和地域的不同,生物量可能会存在一定程度差异,进而,在所构建的不同树种生物量通用方程的基础上,考虑起源和地域的差异,利用哑变量方法构建既能考虑不同树种又能考虑林分起源和不同地域的东北落叶松地上生物量通用方程,并利用加权最小二乘法剔除方程异方差。结果表明:通过哑变量方法构建不同树种生物量模型方法可行;不论是传统的生物量方程,还是只考虑树种或同时考虑树种、起源和地域的通用生物量方程,增加自变量能提高方程预测效果,即,三元生物量方程预测精度最高,二元生物量方程次之,一元生物量方程最低;当同时考虑树种、起源和地域时,方程预测精度最高,只考虑树种的生物量通用方程次之,传统生物量方程最低。因此,如果数据允许,建议构建考虑不同树种、起源和地域的三元生物量方程估计东北地区长白落叶松和兴安落叶松地上生物量。

关键词:兴安落叶松;长白落叶松;地上生物量;通用方程

Generalized above-ground biomass equations for two main species in northeast China

FU Liyong¹, TANG Shouzheng^{1, *}, ZHANG Huiru¹, ZHANG Zelu², ZENG Weisheng³

1 Research Institute of Forest Resources Information Techniques, Chinese Academy of Forestry, Beijing 100091, China

2 Helong Forestry Bureau, Jilin Province, Jilin 133500, China

3 Academy of Forest Inventory and Planting, State Forestry Administration, Beijing 100714, China

Abstract: Individual tree biomass equations have been frequently used in ecological and forestry research over the last 60 years. They represent a powerful tool to understand forest productivity, nutrient cycling, and carbon sequestration, and they are used to estimate other structural and functional characteristics of forest ecosystems. Current attempts to develop above-ground biomass equations for *Larix* forests in northeast China have been mainly focused on only one species or applied to the genus *Larix* as a whole. However, generalized above-ground biomass equations for *Larix* could be used to estimate the average relationship between above-ground biomass and different independent variables and also variations among different *Larix* species. We developed generalized biomass equations for different *Larix* species by using *Larix olgensis* and *Larix gmelinii*. In this study, a total of nine tree variables that were able to predict above-ground biomass in *Larix* species were examined using biomass equations. The results show that *D*, *H*, and *CW* contributed significantly to predict above-ground

基金项目:十二五科技支撑项目(2012BAD22B02);国家自然科学基金青年科学基金项目(31300534)

收稿日期:2014-03-31; 修订日期:2014-11-14

* 通讯作者 Corresponding author. E-mail: stang@caf.ac.cn

biomass. Therefore, three combinations of these variables, including D alone; D and H ; and D , H , and CW , were selected as independent variables to develop univariate, bivariate, and trivariate biomass generalized equations, respectively. The trivariate biomass generalized equations predicted above-ground biomass better than the other two equation types, while the predictive power of the univariate equation was the worse than the rest. Theoretically, the prediction accuracy of trivariate biomass equations could be further increased by adding stands or tree variables; however, including an excessive number of parameters in the biomass equations may hinder computation convergence and reduce the speed required to estimate model parameters. Furthermore, including many stands or tree variables would increase the cost and time required to conduct forest inventories. Therefore, determining the appropriate number of independent variables able to provide the level of accuracy required by forest managers is essential in forest modeling. A parsimonious model with reliable accuracy of prediction has been suggested as a reasonable approach for efficient forest management. For this reason, D , H , and CW were finally selected as independent variables for the generalized biomass equations developed in this study. In general, the biomass of individual trees with the same D would depend on the region studied and the origin of the tree. Thus, the generalized above-ground biomass equations developed for different *Larix* species in northeast China consider this inter-regional variation by using a dummy variable. To reduce heteroskedasticity in the data, we used weighted least square regressions. The results showed that the predictive precision of the biomass equations could be improved by adding predictor variables. Regardless of the traditional biomass equation used, both generalized equations considering only tree species and those considering tree species, tree origin, and region showed the highest prediction power. In addition, the accuracy for predicting above-ground biomass did not differ among univariate, bivariate, and trivariate equations when tree species, tree origin, and region were considered. On the basis of these results, the trivariate generalized biomass equation that considers tree species, tree origin, and region was believed to be the best option for estimating the biomass of *L. olgensis* and *L. gmelinii*.

Key Words: *Larix olgensis*; *Larix gmelinii*; above-ground biomass; generalized equations

兴安落叶松和长白落叶松是东北地区两个主要落叶松用材树种,以兴安落叶松为例,其面积和蓄积分别占我国寒温带有林地面积和蓄积的 55% 和 70%。而且,由于这两种落叶松具备适应性强,生长速度较快,木材通直、纹理清晰、材质致密等特性,在我国用材树种中占有重要地位^[1]。同时兴安落叶松和长白落叶松也是东北地区水土保持、改善环境和生态建设的两个主要树种。因此,越来越多的学者对其生物量监测日益重视,在生物量方程方面开展了大量工作^[2-7]。例如,程云霄等^[2]通过最小二乘法建立了 3 个主要兴安落叶松林型森林生物量方程,得到的不同林型的生物量经验方程都具有较好的拟合效果;刘志刚等^[3]利用相对生长方程对内蒙古大兴安岭林区 3 个气候区兴安落叶松生物量进行建模,并得到含有胸高直径的线性模型有较好的预测精度;胥辉^[4]以长白落叶松为例,验证树木生物量方程中存在方差非齐性,并提出以方程自身作为权函数来消除方差齐性的参数估计方法;唐守正等^[5]以长白落叶松为研究对象,从解决方程之间的兼容性或可加性出发,对 2 种比例平差方法和 3 种非线性联合估计方法进行了对比分析和评价;孙玉军等^[6]通过基于标准地及森林资源清查资料建立生物量与蓄积量的关系模型,对我国兴安落叶松的主要集中地内蒙古大兴安岭地区的生物量和碳储量进行了估算;付尧等^[7]利用林木生物量相容性理论,以树干生物量作为控制量,利用单株木生物量中各组分之间的代数关系,通过联立独立模型求解得到小兴安岭长白山单株木相容性模型。而且,部分学者把兴安落叶松、长白落叶松和日本落叶松等树种的生物量数据合并在一起构建了东北落叶松生物量方程^[8-10]。在上述研究方法中,不论是分树种,还是把不同树种落叶松归为一体构建的生物量方程,都只能反映不同树种或整个落叶松生物量与自变量的平均关系。然而,既能反映落叶松与自变量的平均关系,又能反映不同树种间生物量差异程度的通用性落叶松生物量方程迄今尚未构建。为此,本研究以兴安落叶松和长白落叶松地上生物量为例,重点解决如何构建包含多个树种的生物量通用性方程。

2 材料与方法

2.1 数据来源

本文所用数据为我国东北落叶松(兴安落叶松和长白落叶松)立木地上生物量实测数据,来自2009年全国连清生物量调查建模项目。共136株样木,采集时间为2009年,采集地点涉及内蒙古、黑龙江、吉林和辽宁4省区。样本单元数的选取大体上按各省资源多少分配,并兼顾天然和人工起源。样木数按2、4、6、8、12、16、20、26、32、38 cm以上共10个径阶均匀分配。每个径阶的样木数按树高级从低到高也尽量均匀分配,在大尺度范围内具有广泛的代表性。全部样木都实测胸径、地径和冠幅,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、树叶称鲜重,并分别抽取样品带回实验室,在85℃恒温下烘干至恒重,根据样品鲜重和干重分别推算出样木各部分干重并汇总得到地上部分干重。在所有样木中,兴安落叶松84株,长白落叶松52株。由于辽宁省只有6株长白落叶松,在分地域构建生物量方程时样本量太少,故剔除。因此将剩下130株落叶松进行研究,其中93株用于建模,37株用于模型检验。建模数据和检验数据统计信息见表1和表2。数据的具体调查和测定方法可以参照文献符利勇等^[11]。

表1 建模和检验数据分布概况

Table 1 Sample trees distribution for modelling and validation data sets by tree species and origin

省份 Provinces	建模数据 Modelling data / 株				检验数据 Validation data / 株			
	兴安落叶松 <i>Larix gmelinii</i>		长白落叶松 <i>Larix olgensis</i>		兴安落叶松 <i>Larix gmelinii</i>		长白落叶松 <i>Larix olgensis</i>	
	天然林 Natural	人工林 Planted	天然林 Natural	人工林 Planted	天然林 Natural	人工林 Planted	天然林 Natural	人工林 Planted
黑龙江	14	9	6	6	8	3	2	2
吉林	0	0	10	11	0	0	6	3
内蒙古	21	16	0	0	9	4	0	0
总共 Total	35	25	16	17	17	7	8	5

表2 建模和检验数据变量概况

Table 2 Summary statistics for modelling and validation data sets

变量 Variable	建模数据 Modelling data				检验数据 Validation data			
	最小值 Min	最大值 Max	平均值 Mean	标准差 SD	最小值 Min	最大值 Max	平均值 Mean	标准差 SD
兴安落叶松 <i>Larix gmelinii</i>								
<i>D/cm</i>	2.10	41.80	17.25	11.86	2.00	38.30	18.69	12.92
<i>H/m</i>	2.85	25.90	12.76	6.62	2.61	23.00	13.16	7.68
<i>CW/m</i>	2.35	15.71	8.03	3.68	2.14	14.50	7.76	4.20
<i>W/kg</i>	1.25	9.55	4.42	2.31	1.14	8.30	4.52	2.28
长白落叶松 <i>Larix olgensis</i>								
<i>D/cm</i>	1.60	44.10	17.13	12.74	2.20	26.80	11.93	8.65
<i>H/m</i>	2.50	28.20	13.13	7.31	3.80	21.50	11.02	6.24
<i>CW/m</i>	1.25	19.90	9.20	5.31	2.00	14.50	7.64	4.70
<i>W/kg</i>	1.00	10.60	4.64	2.59	1.45	7.35	3.60	2.02

2.2 基础方程

国内外对生物量方程进行了大量研究,并得到多种经验方程^[12-17]。对于落叶松,尹艳豹等^[9]从胸径(*D*)、树高(*H*)、冠幅(*CW*)、冠长(*CL*)、*D*²、*CW*²以及反映其交互作用的*H/D*、*D*²*H*、*CW*²*CL*这9个变量中筛选出与落叶松生物量相关性较强的因子构建多元线性和非线性生物量方程。通过对线性和非线性生物量方程综合比较,确定在构建生物量方程时,非线性模型要优于线性模型。对应的一元、二元和多元生物量方程表达式

如下:

$$f(W) = \beta_1 D^{\beta_2} + \varepsilon \quad (1)$$

$$f(W) = \beta_1 D^{\beta_2} H^{\beta_3} + \varepsilon \quad (2)$$

$$f(W) = \beta_1 D^{\beta_2} H^{\beta_3} C W^{\beta_4} + \varepsilon \quad (3)$$

式中, W 为地上生物量, β_1 — β_3 为待估参数, ε 为误差项。本研究将以方程(1)—方程(3)作为基础方程来构建东北兴安落叶松和长白落叶松生物量通用方程。

2.3 通用方程

在方程(1)—方程(3)中为区别落叶松树种,将构造一个哑变量 L , 即 $L = 0$ 为兴安落叶松, 反之 $L = 1$ 为长白落叶松。考虑哑变量 L 对方程中所有参数的影响, 方程(1)—方程(3)写为:

$$f(W) = (\beta_1 + \beta_1^{(L)} L) D^{(\beta_2 + \beta_2^{(L)} L)} + \varepsilon \quad (4)$$

$$f(W) = (\beta_1 + \beta_1^{(L)} L) D^{(\beta_2 + \beta_2^{(L)} L)} H^{(\beta_3 + \beta_3^{(L)} L)} + \varepsilon \quad (5)$$

$$f(W) = (\beta_1 + \beta_1^{(L)} L) D^{(\beta_2 + \beta_2^{(L)} L)} H^{(\beta_3 + \beta_3^{(L)} L)} C W^{(\beta_4 + \beta_4^{(L)} L)} + \varepsilon \quad (6)$$

式中, $\beta_1^{(L)}$, $\beta_2^{(L)}$, $\beta_3^{(L)}$ 和 $\beta_4^{(L)}$ 分别为哑变量 L 作用在参数 β_1 — β_4 上的固定效应。

除此之外, 不同树木起源(天然林和人工林)之间生物量可能存在较大差异^[18]。为了能反映该差异程度的大小, 方程中将构造一个区别林分起源的哑变量 S , 即 $S = 0$ 为人工林, 反之 $S = 1$ 为天然林。由于兴安落叶松和长白落叶松主要分布在黑龙江、吉林和内蒙古等 3 个省或自治区, 由于立地、气候和环境等因素以及森林经营方式等差异, 使得不同地域上相同树种且相同直径大小树木生物量可能差异较大, 因此为提高方程的通用性和实用性, 将构造另一个反映地域的哑变量 $P = (P_1, P_2)$, 即 $P_1 = 1, P_2 = 0$ 为黑龙江省, $P_1 = 0, P_2 = 1$ 为吉林省, 当 $P_1 = 0, P_2 = 0$ 时为内蒙古自治区。假定哑变量 S 和 P 对方程参数 β_1 的影响, 则方程(4)—方程(6)的表达式变为:

$$f(W) = (\beta_1 + \beta_1^{(L)} L + \beta_1^{(S)} S + \beta_1^{(P_1)} P_1 + \beta_1^{(P_2)} P_2) D^{(\beta_2 + \beta_2^{(L)} L)} + \varepsilon \quad (7)$$

$$f(W) = (\beta_1 + \beta_1^{(L)} L + \beta_1^{(S)} S + \beta_1^{(P_1)} P_1 + \beta_1^{(P_2)} P_2) D^{(\beta_2 + \beta_2^{(L)} L)} H^{(\beta_3 + \beta_3^{(L)} L)} + \varepsilon \quad (8)$$

$$f(W) = (\beta_1 + \beta_1^{(L)} L + \beta_1^{(S)} S + \beta_1^{(P_1)} P_1 + \beta_1^{(P_2)} P_2) D^{(\beta_2 + \beta_2^{(L)} L)} H^{(\beta_3 + \beta_3^{(L)} L)} C W^{(\beta_4 + \beta_4^{(L)} L)} + \varepsilon \quad (9)$$

式中, $\beta_1^{(S)}$, $\beta_1^{(P_1)}$, $\beta_1^{(P_2)}$ 分别为哑变量 S 和 P 作用在参数 β_1 上的固定效应。

2.5 异方差

生物量数据普遍存在异方差性, 为消除异方差, 常用方法有对数回归法和加权回归法^[8,10]。本文采用加权最小二乘法求解。首先根据建模数据, 利用非线性最小二乘法计算出方程(1)—方程(10)的参数并对建模数据进行拟合, 得到各方程的残差。把原基础方程本身作为权函数形式^[18]对残差进行拟合。

$$\varepsilon^2 = 1/g(x)^2 \quad (10)$$

2.6 方程评价

利用检验数据并结合指标 \bar{e} 、 σ^2 、 δ 对一元、二元和多元生物量通用方程进行比较和评价, 同时把所构建方程与之相对应的基础方程进行比较。本研究所有计算是在 ForStat2.2 版本上实现。

$$\bar{e} = \sum (W_i - \hat{W}_i) / N \quad (11)$$

$$\sigma^2 = \sum (W_i - \hat{W}_i)^2 / (N - 1) \quad (12)$$

$$\delta = \sqrt{\bar{e}^2 + \sigma^2} \quad (13)$$

式中, $i = 1, \dots, N$, W_i 和 \hat{W}_i 分别为真实值和预测值, N 为总观测数。

2 结果与分析

2.1 一元生物量方程

一元生物量方程(1)、方程(4)和方程(7)的权函数见表 3。表 4 为方程的参数估计值。通过参数值大小

发现,当生物量方程中只有 D 一个自变量时,其对应的参数相对比较稳定,即认为兴安落叶松和长白落叶松的生物量与直径之间具有相同相关性。树种层面, $\beta_1^{(L)}$ 小于 0 表明在相同直径条件下,兴安落叶松的生物量比长白落叶松生物量要多。林分起源方面, $\beta_1^{(S)}$ 大于 0 表明两种落叶松树种满足相同直径条件下天然林的生物量对人工林要多。地域分布上, $\beta_1^{(P_1)}$ 和 $\beta_1^{(P_2)}$ 大于零,并且 $\beta_1^{(P_1)}$ 略大于 $\beta_1^{(P_2)}$ 反映出在相同直径条件下两种落叶松的生物量都满足吉林省最多,黑龙江其次,内蒙古最少。表 5 列出方程(1)、方程(4)和方程(7)对应的评价指标。对于建模数据,3 个方程对应的平均残差 \bar{e} 都较小, R^2 都在 0.95 以上,说明方程拟合效果较好,对于检验数据,尽管 3 个方程的平均残差 \bar{e} 较大,但是残差方差 σ^2 和均方误差 δ 都较小说明方程也有较好的预测精度。总体而言,对于一元生物量方程,考虑树种、林分起源和地域因子的方程(7)预测精度最高、其次是仅考虑树种的方程(4),而传统的方程(1)预测精度最低。为此,把通用生物量方程(7)作为一元兴安落叶松和长白落叶松生物量的最优方程。

表 3 一元、二元和三元生物量方程的权函数

Table 3 Weighted functions (WF) of univariate, bivariate and trivariate biomass equations (BE)

一元生物量方程 Univariate BE	一元权函数 Univariate WF	二元生物量方程 Bivariate BE	二元权函数 Bivariate WF	三元生物量方程 Trivariate BE	三元权函数 Trivariate WF
方程 Model(1)	$g(D) = D^{1.86}$	Model(2)	$g(D, H) = D^{1.54} H^{0.70}$	Model(3)	$g(D, H, CW) = D^{1.18} H^{0.79} CW^{0.39}$
方程 Model(4)	$g(D) = D^{1.75}$	Model(5)	$g(D, H) = D^{1.95} H^{0.35}$	Model(6)	$g(D, H, CW) = D^{0.37} H^{1.69} CW^{0.55}$
方程 Model(7)	$g(D) = D^{1.72}$	Model(8)	$g(D, H) = D^{1.20} H^{1.19}$	Model(9)	$g(D, H, CW) = D^{0.89} H^{0.51} CW^{0.39}$

2.2 二元生物量方程

二元生物量方程(2)、方程(5)和方程(8)的权函数见表 3。3 个方程的参数估计值见表 4。从表 4 中可得知,与一元生物量方程类似,3 个方程中的 β_2 较为稳定,表明即使同时考虑 D 和 H 对兴安落叶松和长白落叶松生物量影响,两种树种生物量与直径有相同的相关性。 β_1 大于 β_2 说明直径对两个树种生物量影响强度比树高对生物量的影响要大。在树种层面,与一元生物量方程相类似,方程(5)和方程(8)中 $\beta_3^{(L)}$ 都小于 0 表明在相同直径条件下,兴安落叶松的生物量比长白落叶松生物量要多。林分起源方面,与一元生物量方程相同($\beta_1^{(S)}$ 大于 0),表明两个落叶松树种满足相同直径条件下天然林的生物量对人工林要多。在地域分布上,与一元生物量方程相同,即 $\beta_1^{(P_1)}$ 小于 $\beta_1^{(P_2)}$ 表明在相同直径条件下两种落叶松的生物量都满足吉林省最多,黑龙江其次,内蒙古最少。表 5 列出方程(2)、方程(5)和方程(8)对应的评价指标。对于建模数据,方程(8)对应的 3 个评价指标都最小,对于检验数据,除 \bar{e} 外(方程(2)对应的 $\bar{e} = 3.36$ 最小),方程(8)对应的指标同样比

表 4 方程(1)—方程(9)参数估计值

Table 4 Parameter estimates for equation (1)—equation (9)

参数 Parameters	方程(1) Equation (1)	方程(2) Equation (2)	方程(3) Equation (3)	方程(4) Equation (4)	方程(5) Equation (5)	方程(6) Equation (6)	方程(7) Equation (7)	方程(8) Equation (8)	方程(9) Equation (9)
β_1	0.1135	0.1028	0.0968	0.1132	0.1053	0.1061	0.1067	0.1031	0.0920
β_2	2.3875	2.1507	1.8893	2.3738	2.1020	1.8749	2.3753	2.1352	1.7450
β_3	—	0.2975	0.4391	—	0.3267	0.4120	—	0.2930	0.4767
β_4	—	—	0.2800	—	—	0.2753	—	—	0.4520
$\beta_1^{(L)}$	—	—	—	-0.0030	0.0427	0.0116	-0.0177	0.0376	-0.0369
$\beta_1^{(S)}$	—	—	—	—	—	—	0.0018	0.0021	0.0043
$\beta_1^{(P_1)}$	—	—	—	—	—	—	0.0140	0.0063	0.0148
$\beta_1^{(P_2)}$	—	—	—	—	—	—	0.0197	0.0069	0.0202
$\beta_2^{(L)}$	—	—	—	0.0481	0.3045	0.3186	0.0547	0.3462	0.4735
$\beta_3^{(L)}$	—	—	—	—	-0.4197	-0.3197	—	-0.4630	-0.1281
$\beta_4^{(L)}$	—	—	—	—	—	-0.0571	—	—	-0.4390

方程(5)和方程(8)要小。总体而言,与一元生物量方程相类似,考虑树种、林分起源和地域因子的二元生物量方程(7)预测精度最高,其次是仅考虑树种的生物量方程(4),而传统的生物量方程(1)预测精度最低。为此,把通用生物量方程(8)作为二元兴安落叶松和长白落叶松生物量的最优方程。

2.3 三元生物量方程

三元生物量方程(3)、方程(6)和方程(9)的权函数见表3。3个方程的参数估计值见表4。从表4中可得知,当同时考虑 D 、 H 和 CW 对兴安落叶松和长白落叶松生物量影响时,3个方程中的 β_2 也较为稳定。 $\beta_2 > \beta_3 > \beta_4$ 表明直径对两个树种生物量影响强度最大,其次是树高,树冠最弱。在树种层面,方程(6)中 $\beta_3^{(L)}$ 和 $\beta_4^{(L)}$ 小于0,以及方程(9)中 $\beta_1^{(L)}$ 、 $\beta_3^{(L)}$ 和 $\beta_4^{(L)}$ 小于0,都表明在相同直径条件下,兴安落叶松的生物量比长白落叶松生物量要多。林分起源方面,与一元和二元生物量方程相类似,即 $\beta_1^{(S)}$ 大于0表明两种落叶松树种满足相同直径条件下天然林的生物量对人工林要多。在在地域分布上,与一元和二元生物量方程相同,即 $\beta_1^{(P_1)}$ 小于 $\beta_1^{(P_2)}$ 表明在相同直径条件下两种落叶松的生物量都满足吉林省最多,黑龙江其次,内蒙古最少。表5列出方程(3)、方程(6)和方程(9)对应的评价指标。对于建模数据,方程(9)对应的3个评价指标都最小,尤其是 \bar{e} 和 δ ,分别比方程(3)减少98.79%和26.23%,比方程(6)减少99.07%和16.64%。对于检验数据,方程(3)的3个指标最小,其次是方程(9),而方程(6)的评价指标最大。总之,在利用 D 、 H 和 CW 构建生物量方程时,3个方程的预测精度都很高,例如在建模数据中,方程(3)、方程(6)和方程(9)对应的决定指数都在0.97以上。总体而言,方程(9)的拟合效果最好,因此把通用生物量方程(9)作为三元兴安落叶松和长白落叶松生物量的最优方程。

表5 方程评价指标

Table 5 Equation evaluation indexes

方程 Equation	建模数据 Modelling data				检验数据 Validation data		
	\bar{e}	σ^2	δ	R^2	\bar{e}	σ^2	δ
方程 Equation(1)	0.5847	3011.8334	54.8833	0.9512	1.4057	785.3531	28.0594
方程 Equation(2)	3.6402	2673.9237	51.8380	0.9571	3.36	460.1889	21.71431
方程 Equation(3)	1.9174	1844.6836	42.9926	0.9701	4.2333	280.5467	17.2762
方程 Equation(4)	-0.0240	2398.9183	48.9788	0.9612	6.0808	590.0516	25.0405
方程 Equation(5)	5.8043	2065.6592	45.8187	0.9677	9.8428	478.9464	23.9964
方程 Equation(6)	2.4965	1441.4730	38.0487	0.9766	9.6828	405.2733	22.3390
方程 Equation(7)	-0.3320	1970.9774	44.3969	0.9682	4.2819	783.5254	28.3171
方程 Equation(8)	3.5557	1886.1775	43.5755	0.9695	7.5391	460.3530	22.7418
方程 Equation(9)	-0.0232	1005.9013	31.7160	0.9837	7.3271	333.3708	19.6738

2.4 一元、二元和三元生物量方程比较

对于传统生物量方程(方程(1)(一元)、方程(2)(二元)和方程(3)(三元)),考虑树种的生物量方程(方程(4)(一元)、方程(5)(二元)和方程(6)(三元))以及考同时考虑树种、起源和地域的生物量方程(方程(7)(一元)、方程(8)(二元)和方程(9)(三元)),不管是建模数据还是检验数据,都满足三元生物量方程预测精度最高、其次是二元生物量方程,而一元生物量方程最低。尤其是均方误差 δ ,三元生物量方程比一元和二元生物量方程明显要小。

根据前面分析得知,不管是一元、二元还是三元生物量方程,当同时考虑树种、起源和地域时,方程拟合效果最好,并结合本小节分析结果最终确定方程(9),即同时考虑树种、起源和地域的三元生物量方程,拟合精度最高(表5)。因此把方程(9)作为长白山和兴安落叶松地上生物量通用方程。图1方程(9)对建模数据拟合的残差和 D 、 H 和 CW 之间关系的分布图。从图中得知,不管是 D 、 H 还是 CW ,残差分布趋势都平行于横轴,即把原基础方程本身作为权函数能明显剔除方程异方差。从图1中的残差分布QQ图得知残差分布服从正态分布。因此进一步说明方程(9)拟合效果较好。

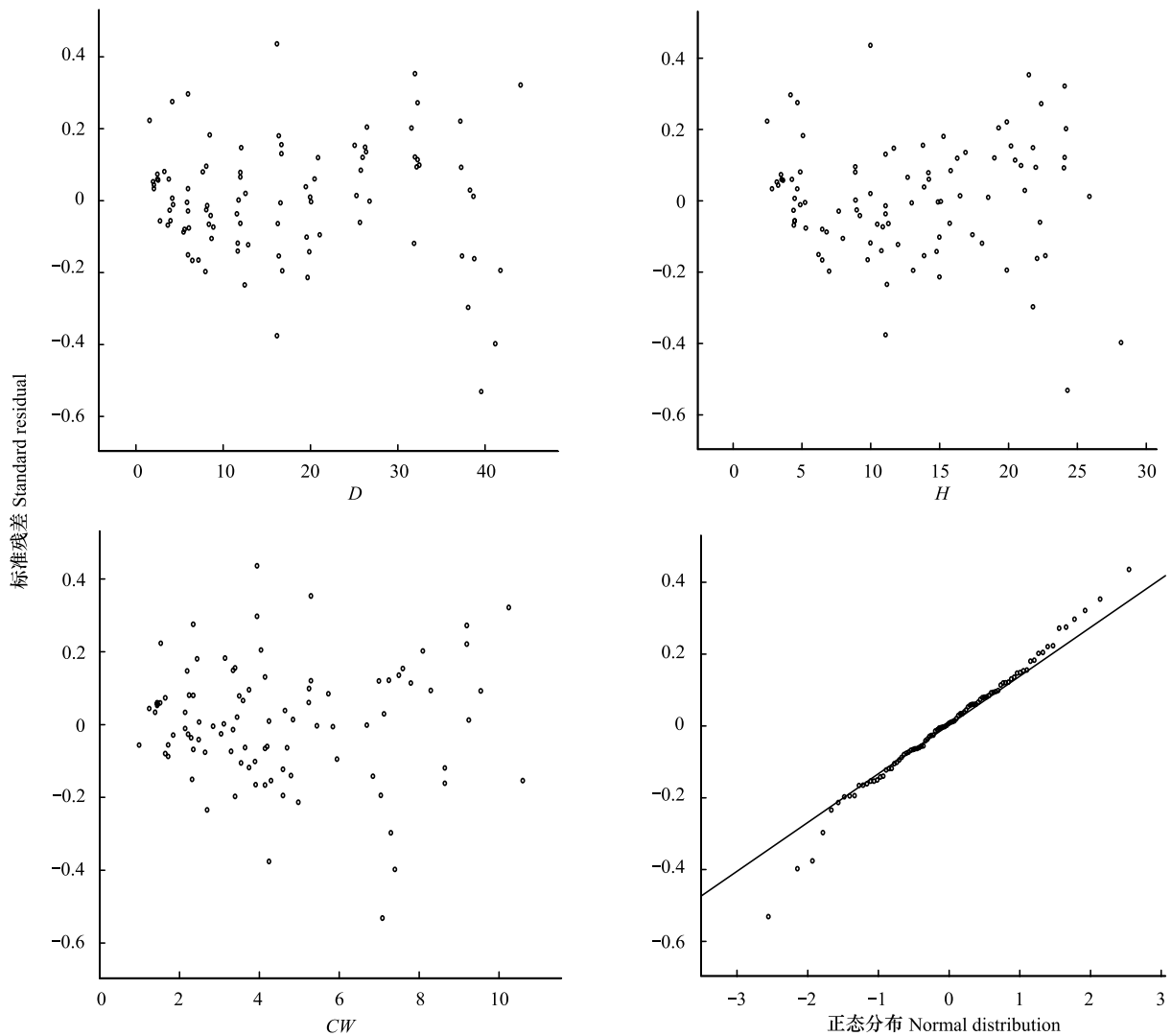


图1 方程(9)的残差分布和QQ图

Fig.1 Residual distribution and QQ plot for equation (9)

3 结论与讨论

本文以东北地区长白落叶松和兴安落叶松生物量数据为研究对象,构建一元、二元和三元生物量方程。通过对9个变量进行筛选,最后确定直径、树高、和冠幅与生物量相关性较强,这与尹艳豹等^[9]对东北落叶松地上生物量、符利勇等^[18]对南方马尾松地上生物量等研究结论一致。从决定系数单纯而言,一元、二元和三元生物量方程对应的决定系数都较高,并且它们之间相差不是很大(表5),这也与尹艳豹等^[9]对东北落叶松地上生物量建模时,对应的一元、二元和三元生物量方程的决定系数分别为0.9591,0.9726,和0.9624完全相似。但从均方差误差方面而言,不管是建模数据还是检验数据,增加自变量能明显提高方程预测效果。

符利勇等^[18]在分析南方马尾松时发现对于相同直径大小的树木,天然林地上生物量比人工林地上生物量要大。在本研究中,落叶松也有相似的结论。因此在构建生物量方程时,考虑林分不同起源能提高方程预测精度。本文是把林分起源差异当作固定效应,通过构造哑变量对方程参数化。而曾伟生等^[8]和Fu等^[19]把林分起源差异作为随机效应,利用线性混合效应模型对马尾松生物量进行建模,结果发现,考虑随机效应的生物量方程比总体平均模型的预测精度高。同时Wang等^[20]和曾伟生等^[8]研究发现,当每个类型(天然林和人工林)的样本量都较大时(大于10),选择哑变量和混合模型方法其所构建的方程预测结构几乎相同。由于本

文研究重点是从如何构建多树种、多起源和多地域的通用性生物量方程,与此同时,建模数据中两个树种天然林和人工林数量也较多。因此,本研究只把林分起源因子作为哑变量进行建模。在后续的研究中,将尝试把林分起源作为随机因子利用非线性混合效应模型方法建模。

曾伟生等^[8]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,结果表明,相同直径林木的地上生物量估计值随树种和区域的不同存在一定程度的差异。为此,在本研究中,把树种和起源在生物量方程中通过构建哑变量区分开,通过方程中哑变量参数值大小区分不同树种间、不同地域间以及不同林分起源的差异,故方程通用性较强。与传统的生物量方程相比,该方法除解决了方程通用性问题,而且还能进一步提高预测精度(表5)。因此,在生物建模中建议使用该方法。本研究只分析了2个树种,对于多个树种(大于3个),方法与本文类似,此处哑变量构建方法见文献 Zeng 等^[16]。

本文最终确定方程(9),即同时考虑树种、起源和地域的三元生物量方程作为东北区长白落叶松和兴安落叶松地上生物量通用方程。然而现实中,树高和冠幅的测定工作量大,耗时耗力,并且由于样地地形险要或林分结构复杂等因素使得测量的数据存在较大的度量误差等问题,因此使得方程适用时要求较高。为此,应用时,如果偏重方程的适用性,建议使用同时考虑树种、起源和地域的一元生物量通用方程,例如方程(3)。如果强调方程的预测精度,则建议应用三元生物量通用方程,例如方程(9)。由于 H 和 CW 不可避免含有度量误差,因此可以尝试在当前方程基础上进一步考虑度量误差提高方程预测精度,即利用度量误差理论构建多元生物量方程。

参考文献 (References):

- [1] 张俊. 兴安落叶松人工林群落结构、生物量与碳储量研究[D]. 北京:北京林业大学, 2008.
- [2] 程云霄, 李忠孝. 兴安落叶松三个主要林型森林生物量的初步研究. 内蒙古林业调查设计, 1989, (4): 29-39.
- [3] 刘志刚, 马钦彦, 潘向丽. 兴安落叶松天然林生物量及生产力的研究. 植物生态学报, 1994, 18(4): 328-337.
- [4] 胥辉. 生物量模型方差非齐性研究. 西南林学院学报, 1999, 19(2): 73-77.
- [5] 唐守正, 张会儒, 胥辉. 相容性生物量模型的建立及其估计方法研究. 林业科学, 2000, 36(S1): 19-27.
- [6] 孙玉军, 张俊, 韩爱惠, 王雪军, 王新杰. 兴安落叶松(*Larix gmelini*)幼中龄林的生物量与碳汇功能. 生态学报, 2007, 27(5): 1756-1762.
- [7] 付尧, 马炜, 王新杰, 孙玉军. 小兴安岭长白落叶松相容性生物量模型的构建. 东北林业大学学报, 2012, 39(7): 42-45.
- [8] 曾伟生, 唐守正, 夏忠胜, 朱松, 罗洪章. 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. 林业科学研究, 2011, 24(3): 285-291.
- [9] 尹艳豹, 曾伟生, 唐守正. 中国东北落叶松立木生物量模型的研建. 东北林业大学学报, 2010, 38(9): 23-26, 92-92.
- [10] 曾伟生, 唐守正. 东北落叶松和南方马尾松地下生物量模型研建. 北京林业大学学报, 2011, 33(2): 1-6.
- [11] 符利勇, 曾伟生, 唐守正. 利用混合模型分析地域对国内马尾松生物量的影响. 生态学报, 2011, 31(19): 5797-5808.
- [12] Parresol B R. Additivity of nonlinear biomass equations. Canadian Journal of Forest Research, 2001, 31(5): 865-878.
- [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. Forest Ecology and Management, 2004, 194(1/3): 145-157.
- [14] Bi H, Turner J, Lambert M J. Additive biomass equations for native eucalypt forest trees of temperate Australia. Trees, 2004, 18(4): 467-479.
- [15] Fehrmann L, Lehtonen A, Kleinn C, Tomppo E. Comparison of linear and mixed-effect regression models and a k -nearest neighbour approach for estimation of single-tree biomass. Canadian Journal of Forest Research, 2008, 38(1): 1-9.
- [16] Zeng W S, Zhang H R, Tang S Z. Using the dummy variable model approach to construct compatible single-tree biomass equations at different scales—a case study for Masson Pine (*Pinus massoniana*) in southern China. Canadian Journal of Forest Research, 2011, 41(7): 1547-1554.
- [17] Fu L Y, Zeng W S, Tang S Z, Sharma R P, Li H K. Using linear mixed model and dummy variable model approaches to construct compatible single-tree biomass equations at different scales— a case study for masson pine in Southern China. Journal of Forest Science, 2012, 58(3): 101-115.
- [18] 符利勇, 雷渊才, 孙伟, 唐守正, 曾伟生. 不同林分起源的相容性生物量模型构建. 生态学报, 2014, 34(6): 1461-1470.
- [19] Fu L Y, Zeng W S, Zhang H R, Wang G X, Lei Y C, Tang S Z. Generic linear mixed-effects individual-tree biomass models for *Pinus massoniana* in Southern China. Southern Forests: A Journal of Forest Science, 2014, 76(1): 47-56.
- [20] Wang M L, Borders B E, Zhao D H. An empirical comparison of two subject—specific approaches to dominant heights modeling, the dummy variable method and the mixed model method. Forest Ecology and Management, 2008, 255(7): 2659-2669.