ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Environmental Sciences 10 (2011) 215 – 221

Procedia
Environmental Sciences

2011 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011)

A Nonlinear Mixed-Effects Model to Predict Stem Cumulative Biomass of Standing Trees

Yaoxiang Li, Lichun Jiang*, Mingyu liu

Northeast Forestry University, No. 26 Hexing Road, Harbin 150040, China

Abstract

A nonlinear mixed-effects modeling approach was used to model stem cumulative biomass based on logistic model for dahurian larch (*Larix gmelinii* Rupr.) plantations in northeastern China. The NLME procedure in S-Plus is used to fit the mixed-effects models for stem biomass data. The results showed that logistic model with random parameter b_1 could significantly improve the model performance. The fitted mixed effects model was also evaluated using mean error, mean absolute error, mean percent error, and mean absolute percent error. The mixed model was found to predict stem cumulative biomass better than the original model fitted using ordinary least squares based on all errors. The application of mixed stem cumulative biomass model not only showed the mean trends of stem cumulative biomass, but also showed the individual difference based on variance-covariance structure of random parameters.

Selection and/or peer-review under responsibility of Conference ESIAT2011 Organization Committee.

Keywords: mixed models, stem cumulative biomass, random effects, fixed effects, larch

1. Introduction

Dahurian larch (*Larix gmelinii* Rupr.) is one of the most widely planted tree species and also important commercial species in northeastern China. Forest play an important role in the global carbon cycle (Zhou et al., 2002; Tan et al., 2006). Details of succession studies, nutrient cycles, production and competition in vegetal communities require estimation of vegetal biomass and production (Tausch and Tueller, 1998; Salis et al., 2006). Thus estimation of forest biomass have become an important topic in ecosystem studies.

The production of woody tissue by trees is a large component of NPP in forests (Ryan et al., 1997; Woolley et al., 2007). Stem wood and bark production is a large component of tree NPP (40–70%; Runyon et al., 1994; Campbell et al., 2004). Biomass measurements are expensive and time-consuming. Therefore, indirect methods for estimating the time-consuming variables are necessary. One of the most

* Corresponding author. Tel.: +86 451 82191751; fax: +86 451 82191751.
E-mail address: jlichun@yahoo.com

Selection and/or peer-review under responsibility of Conference ESIAT2011 Organization Committee.
doi:10.1016/j.proenv.2011.09.037

common methods is biomass prediction equation that can be used to predict stem biomass on an individual-tree basis.

Stem cumulative biomass varies from one tree to another, and even within the same plot. Mixed effects models provide a tool for analyzing this type of data since this method recognizes and estimates two distinct types of variability: between-individual variability and within-individual variability (Lindstrom and Bates, 1990; Pinheiro and Bates, 1998). In addition, mixed model give an unbiased and efficient estimation of the fixed parameters and explain the variation of stem cumulative biomass from tree to tree. Furthermore, mixed models improve predictive ability if we are able to predict the value of the random parameters for an unsampled location (Vonesh and Chinchilli, 1997). Therefore, the construction of stem cumulative biomass models is performed using mixed models based on existing growth models.

2. Stem biomass samples

The material was collected from dahurian larch plantations located in Wuying forest bureau in Heilongjiang Province, northeastern China. Ten plots were established from ten plantations. Three trees from each plot were selected for biomass measurements. Before the felling, diameter at breast height, the total tree height, height to the base of live crown, and crown diameter were measured for each tree. After the felling, each tree was bucked into sections and wood disks were obtained at 2 m intervals above the tree base. The stem dry weight was determined from the stem volume using the specific gravity of wood. Specific gravity of the stem was determined using wood samples taken at each section. The volume of each sample was determined from the volume of water it displaced when submerged. The basic specific gravity was calculated as oven-dry weight divided by volume.

3. Nonlinear Mixed-Effects Models

The logistic model was selected in this study due to its flexibility and biologically interpretable coefficients. The form of the mixed logistic model is:

$$\begin{cases} y_{ij} = \dfrac{w_1}{1+(w_2)\exp\left(-(w_3)\,x_{ij}\right)} + \varepsilon \\ w_i = A_i\beta + B_i b_i \end{cases} \tag{1}$$

Where y_{ij} is the stem biomass at height j from the i^{th} tree (kg), x_{ij} is the relative height at height j from the i^{th} tree, and ε_{ij} is random error, β is a p-dimensional vector of fixed effects parameters, b_i is a q-dimensional random effects vector associated with the i^{th} individual and $b_i \sim N(0,\varphi)$, φ is a general variance-covariance matrix for the random effects, and A_i and B_i are design matrices for the fixed and random effects respectively. It is further assumed that observations made on different individuals are independent and the ε_i follow a $N(0,\sigma^2 I)$ distribution and are independent of the b_i .

4. Results and Discussion

Individual fits approach was first used to determine parameter effect either as mixed or purely fixed. Confidence intervals were obtained on the parameters in logistic model showed in eq. (1) based on individual fits using *nlsList* function in S-Plus (Pinheiro and Bates, 2000). Fig. 1 gives the approximate 95% confidence intervals for three parameters of b_1, b_2, and b_3 in eq. (1) for each tree. It was noticed that, for parameter b_1, the confidence intervals of the thirty trees showed more among-tree variability. Comparatively, confidence intervals for parameter b_2 and b_3 showed little variability from tree to tree. Therefore, random parameter b_1 was included into logistic mixed model.

Scatter plots of residuals were constructed for mixed-effects and fixed-effects models (Fig. 2-3). The scatter plots showed that the mixed-effects model showed more homogeneous residual variance and no systematic pattern in the variation of the residuals. It indicated that mixed-effects model significantly improve the model performances compared to fixed-effects model.

The performance of the mixed and fixed models was visualized by displaying the fitted and observed values in the same tree (Fig. 4). Both the fixed-effects model with random effects set to zero and mixed-effects model are compared. The mixed-effects model more closely followed the trend of actual values for most trees and indicated that mixed effects model described the cumulative stem biomass of larch well.

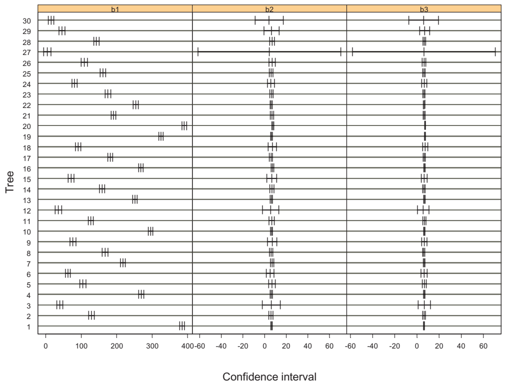

Fig. 1. Ninety-five percent confidence interval on the logistic model parameters for each tree

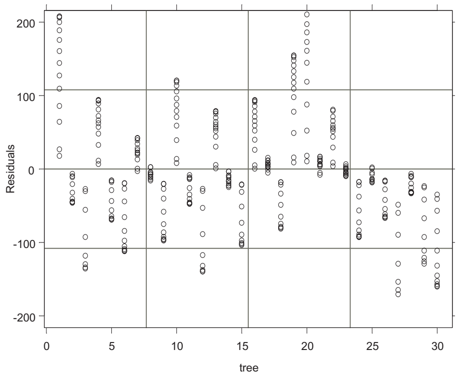

Fig. 2. Scatter plot of residuals for fixed model

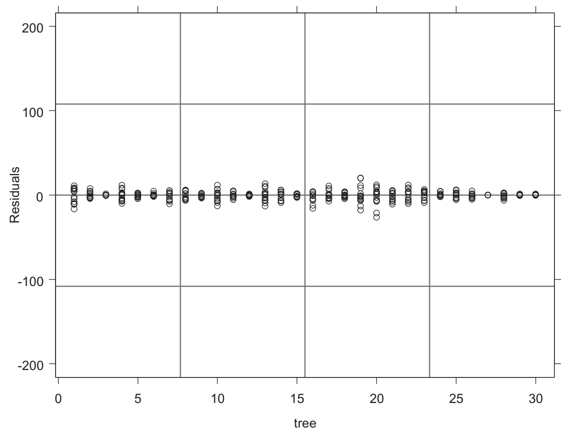

Fig. 3. Scatter plot of residuals for mixed model

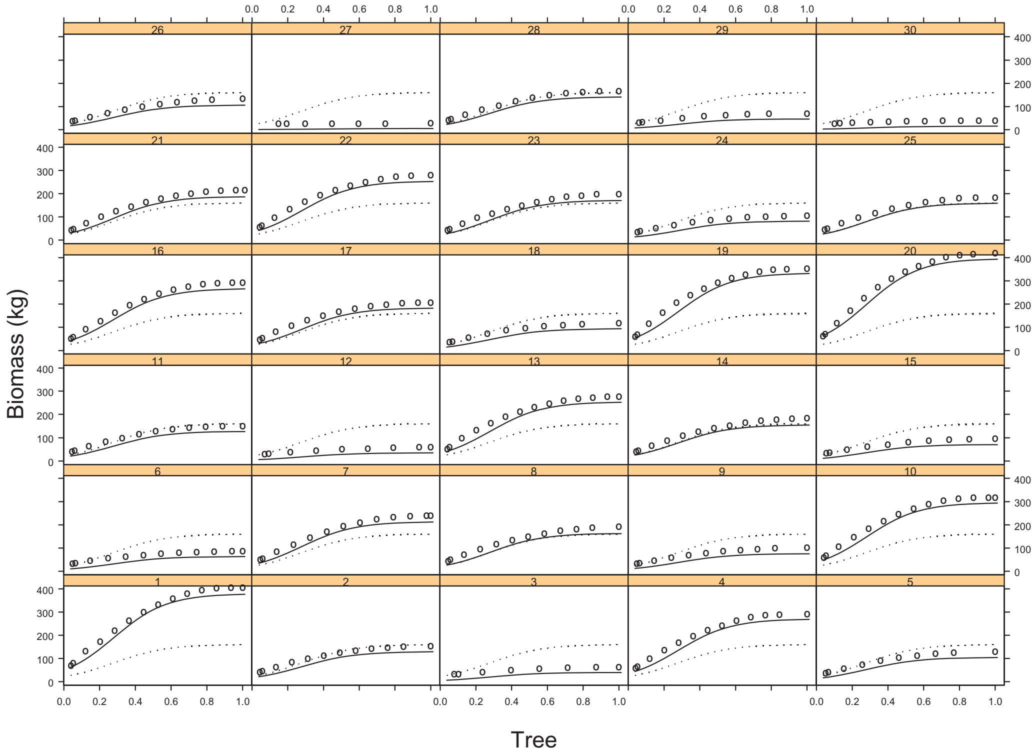

Fig. 4. Comparison of fixed-effects and mixed-effects prediction (circle is actual biomass, solid lines are values from mixed-effects model, dot lines are values from fixed-effects model)

To demonstrate the predictive ability of the mixed and fixed models, Mean error (ME), Mean absolute error (MAE), Mean percent error ($M\%E$), and Mean absolute percent error ($MA\%E$) were used for comparisons as following:

$$ME = \sum_{i=1}^{n} \left(\frac{y_i - \hat{y}_i}{n} \right) \tag{2}$$

$$MAE = \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{n} \right| \tag{3}$$

$$M\%E = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i - \hat{y}_i}{y_i} \right) \times 100\% \tag{4}$$

$$MA\%E = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \tag{5}$$

ME, MAE, $M\%E$, and $MA\%E$ were calculated for each tree (Table 1). Ranges of mean errors and mean absolute errors were [-1.9575, 0.5354] and [0.1881, 8.3406] for mixed model and [-116.5673, 144.0735] and [4.8330, 144.0735] for fixed model, respectively. The mixed model also showed lower mean percent errors and mean absolute percent errors than the fixed model for all test trees.

Table 1. Mean error, mean percent error, mean absolute error, and mean absolute percent error for fixed and mixed models

Tree	Mixed model				Fixed model			
	ME /kg	$M\%E$	MAE /kg	$MA\%E$	ME /kg	$M\%E$	MAE /kg	$MA\%E$
1	-0.8332	-0.0290	7.9828	0.0649	137.2137	0.5270	137.2137	0.5270
2	0.2505	-0.0091	3.0857	0.0637	-29.6683	-0.3580	29.6683	0.3580
3	0.0117	0.0003	0.4017	0.0241	-90.3942	-3.3583	90.3942	3.3583
4	0.1384	-0.0173	5.2161	0.0622	62.0331	0.3445	62.0331	0.3445
5	-0.3504	-0.0374	2.0003	0.0685	-44.7159	-0.7251	44.7159	0.7251
6	0.4564	0.0214	1.4426	0.0477	-75.5147	-1.6827	75.5147	1.6827
7	-1.2275	-0.0478	4.6187	0.0680	25.0770	0.1452	25.5321	0.1626
8	-0.2998	-0.0258	4.3119	0.0701	-7.8341	-0.0909	8.6619	0.1050
9	-0.4597	-0.0546	1.6355	0.0774	-64.4859	-1.4150	64.4859	1.4150
10	-0.3274	-0.0269	5.2285	0.0626	82.1833	0.3950	82.1833	0.3950
11	-0.1814	-0.0291	2.2369	0.0650	-31.7878	-0.4027	31.7878	0.4027
12	-0.0312	-0.0010	0.8154	0.0390	-96.0505	-4.0574	96.0505	4.0574
13	0.0934	-0.0312	5.9487	0.0839	53.5473	0.2933	53.5473	0.2933
14	-0.9900	-0.0632	4.4470	0.0939	-14.6540	-0.1940	14.6540	0.1940
15	-0.2988	-0.0361	1.5702	0.0666	-69.1544	-1.5538	69.1544	1.5538
16	-1.3981	-0.0588	4.9139	0.0795	61.0103	0.3098	61.0103	0.3098
17	0.3774	-0.0118	3.8454	0.0638	5.9298	0.0393	6.8502	0.0768
18	-0.2092	-0.0420	2.1794	0.0838	-51.9490	-0.9277	51.9490	0.9277
19	0.5354	-0.0314	8.3406	0.0900	105.9494	0.4623	105.9494	0.4623
20	-1.9575	-0.0742	8.1453	0.1015	144.0735	0.5283	144.0735	0.5283
21	-0.9747	-0.0570	4.7902	0.0878	7.8174	0.0218	9.7175	0.1079
22	0.0941	-0.0203	6.1682	0.0700	52.6177	0.3011	52.6177	0.3011

23	-0.6279	-0.0456	4.6162	0.0874	-2.9209	-0.0595	4.8330	0.0885
24	0.2544	-0.0010	1.7754	0.0548	-61.2894	-1.1219	61.2894	1.1219
25	-0.0843	-0.0173	3.3735	0.0587	-10.5913	-0.1109	11.0430	0.1192
26	-0.6641	-0.0532	2.4683	0.0788	-45.3030	-0.7237	45.3030	0.7237
27	-0.0199	0.0185	0.1881	0.0686	-116.5673	-35.6084	116.5673	35.6084
28	-0.6130	-0.0396	2.4766	0.0616	-22.9584	-0.2770	22.9584	0.2770
29	0.0362	0.0006	0.4985	0.0231	-81.9174	-2.7212	81.9174	2.7212
30	0.1739	0.0460	0.4932	0.0683	-112.2232	-9.8846	112.2232	9.8846

5. Conclusions

In this study, a nonlinear mixed-effects cumulative stem biomass model was developed for dahurian larch in northeastern China. Parameter effects were determined using individual fit. Nonlinear mixed-effects modeling techniques were used to estimate fixed and random effects parameters for logistic model. The results showed that logistic model with random parameter b_1 was found to be best based on confidence intervals of three parameters of b_1, b_2, and b_3. The mixed model showed lower mean error, mean absolute error, mean percent error, and mean absolute percent error than the fixed model for all trees. The mixed-effects model provided better model fitting and more precise cumulative stem biomass estimations than the fixed-effects model. The application of mixed cumulative stem biomass developed not only showed the mean trends of stem biomass, but also showed the individual difference based on random parameters and variance-covariance structure.

Acknowledgements

This work was supported by the grant from the NSFC (30972363), Special Fund for Forestry-Scientific Research in the Public Interest (201004026), China and Heilongjiang Province Postdoctoral Science Foundation (200902362, LBH-Z08271), and Scientific Research Foundation for Returned Scholars, Heilongjiang Province (LC2009C08) and State Education Ministry.

References

[1] Zhou, G., Wang, Y., Jiang, Y., and Yang, Z. Estimating biomass and net primary production from forest inventory data: a case study of China's *larix* forests. For Ecol Manage, 2002; 169, 149-157.

[2] Tan, K., Piao, S., Peng, C., and Fang, J. Satellite-based estimation of biomass carbon stocks for northeast China's forests between 1982 and 1999. For Ecol Manage, 2006; 240, 114-121.

[3] Tausch, R.J., and Tueller, P.T. Comparison of regression methods for predicting singleleaf pinyon phytomass, Great Basin Nat, 1988; 48, 39–45.

[4] Salis, S.M., Assis, M.A., Mattos, P.P., and Piao, A.C.S. Estimating the aboveground biomass and wood volume of savanna woodlands in Brazil's Pantanal wetlands based on allometric correlations. For Ecol Manage, 2006; 228, 61-68.

[5] Ryan, M.G., Binkley, D., and Fownes, J.H. Age related decline in forest productivity: pattern and process. Adv Ecol Res. 1997; 27, 213–262.

[6] Woolley, T.J., Harmon M.E., and O'Connell, K.B. Estimating annual bole biomass production using uncertainty analysis. For Ecol Manage, 2007; 253, 202-210.

[7] Runyon, J., Waring, R.H., Goward, S.N., and Welles, J.M. Environmental limitations on net primary production and light-use efficiency across the Oregon transect. Ecol Appl, 1994; 4, 226–237.

[8] Campbell, J., Sun, O., and Law, B.E. Disturbance and net ecosystem production across three climatically distinct forest landscapes, Global Biogeochem. Cycles, 2004; 18, 1–11.

[9] Lindstrom, M.J., and Bates, D.M. Nonlinear mixed effects models for repeated measures data, Biometrics, 1990; 46, 673-687.

[10] Pinheiro, J.C., and Bates, D.M. Model building for nonlinear mixed effects models, 1998; Tech. rep, Dep of Stat, Univ of Wisconsin.

[11] Vonesh, E.F., and Chinchilli, V.M. Linear and nonlinear models for the analysis of repeated measurements. 1997; Marcel Dekker, New York; 1999.

[12] Pinheiro, J.C., and Bates, D.M. Mixed-effects models in S and S-PLUS., Springer, New York; 2000.