Article

A Bayesian Approach to Estimating Seemingly Unrelated Regression for Tree Biomass Model Systems

Longfei Xie ¹, Fengri Li ¹, Lianjun Zhang ², Faris Rafi Almay Widagdo ¹ and Lihu Dong ^{1,*}

¹ Key Laboratory of Sustainable Forest Ecosystem Management-Ministry of Education, School of Forestry, Northeast Forestry University, Harbin 150040, China; xielongfei@nefu.edu.cn (L.X.); fengrili@nefu.edu.cn (F.L.); faris.almay@gmail.com (F.R.A.W.)

² Department of Sustainable Resources Management, College of Environmental Science and Forestry (SUNY-ESF), State University of New York, One Forestry Drive, Syracuse, NY 13210, USA; lizhang@esf.edu

* Correspondence: lihudong@nefu.edu.cn; Tel.: +86-451-82191751

Received: 13 October 2020; Accepted: 2 December 2020; Published: 4 December 2020

Abstract: Accurate estimation of tree biomass is required for accounting for and monitoring forest carbon stocking. Allometric biomass equations constructed by classical statistical methods are widely used to predict tree biomass in forest ecosystems. In this study, a Bayesian approach was proposed and applied to develop two additive biomass model systems: one with tree diameter at breast height as the only predictor and the other with both tree diameter and total height as the predictors for planted Korean larch (*Larix olgensis* Henry) in the Northeast, P.R. China. The seemingly unrelated regression (SUR) was used to fit the simultaneous equations of four tree components (i.e., stem, branch, foliage, and root). The model parameters were estimated by feasible generalized least squares (FGLS) and Bayesian methods using either non-informative priors or informative priors. The results showed that adding tree height to the model systems improved the model fitting and performance for the stem, branch, and foliage biomass models, but much less for the root biomass models. The Bayesian methods on the SUR models produced narrower 95% prediction intervals than did the classical FGLS method, indicating higher computing efficiency and more stable model predictions, especially for small sample sizes. Furthermore, the Bayesian methods with informative priors performed better (smaller values of deviance information criterion (DIC)) than those with the non-informative priors. Therefore, our results demonstrated the advantages of applying the Bayesian methods on the SUR biomass models, not only obtaining better model fitting and predictions, but also offering the assessment and evaluation of the uncertainties for constructing and updating tree biomass models.

Keywords: Bayesian approaches; seemingly unrelated regression; Korean larch; Markov chain Monte Carlo; Gibbs sampler

1. Introduction

In the practice of sustainable forest management, tree biomass is primarily used to calculate the carbon storage and sequestration of forests and further to comprehend climate change and forest health, productivity, and nutrient cycling [1,2]. It is well known that the most appropriate method for estimating tree biomass is to use biomass models developed using the direct measurements of tree biomass (response variable) and predictor variables [3,4]. Over the past decades, hundreds of biomass equations have been developed for different tree species around the world [5–7]. Tree diameter at breast height (DBH) is the commonly used predictor variable, which is considered the readily attainable and most reliable sole predictor [8]. Tree height (H) is also used for developing biomass models to

improve the model performance and explain the potential limitations of intra-species divergence [9,10]. The biomass models with both DBH and H are recognized more stable and reliable for predicting tree biomass [11,12]. Furthermore, various modeling approaches have been explored and applied for developing tree biomass models [13,14].

To date, the biomass of different tree components (i.e., stem, branch, foliage, and root) are usually estimated jointly in order to account for the inherent correlations among those tree components because they are measured within the same sample trees [11]. The additive model systems of either linear or nonlinear functions are used to deal with the correlations between the biomass components from the same tree, which are commonly known as seemingly unrelated regression (SUR) models [15]. Over the past decades, there are different parameter estimation methods proposed for the additive model systems from the perspective of traditional statistics, such as two-step feasible generalized least squares (FGLS), iterative FGLS (IFGLS), maximum likelihood (ML), etc. [16,17]. FGLS has been widely used for parameter estimation of the SUR models [18] and has been considered more favorable in recent years because of its simplicity and flexibility in practical applications [19]. However, it is recognized that those classical methods are not able to access and evaluate the uncertainty of parameter estimates and the reliability of model predictions.

Zellner [20] proposed the usage of Bayesian analyses in modeling practices to take advantage of Bayesian's capability of estimating model parameters using probability statements and, consequently, fulfilling the purposes of uncertainty assessment [21]. In the past, however, the applications of the Bayesian approaches were limited due to the lack of numerical optimization theory and computing power. In addition, Bayesian methods can be readily applied to linear models, but many allometric relationships are nonlinear in nature. Following the widespread applications of the Markov Chain Monte Carlo (MCMC) method and rapid development of computing technology, Griffiths [22] finally implemented Bayesian methods for SUR models in the field of econometrics.

For many years, Bayesian theorem was not accepted and well recognized among researchers, even though it was proposed by Thomas Bayes in the mid-18th century [23]. Scientists began to accept and deepen their understanding of Bayesian modeling philosophy because it has unique benefits in solving some practical problems, which are highly supported by the remarkable escalation in computing power. The fundamental enhancement of applying Bayesian methods lies within their ability to create the full posterior distributions of model parameters; hence, several statistics (e.g., mean, median, quantiles, etc.) can be easily computed from the samples generated during the model fitting processes [24]. In addition, Bayesian methods consider the model parameters as random variables and assume that each parameter has a prior distribution, which then is naturally integrated with the sample information to obtain the posterior probability distribution for the parameter [25]. To date, Bayesian analysis is considered an essential and important element in modeling practices across many scientific fields, such as biology [26], engineering [27], finance [25], genetics [28], etc. In general, Bayesian analysis can be divided into two major categories, namely objective and subjective Bayesian methods, due to the types of prior distributions. The objective Bayesian method often uses a non-informative prior (e.g., Jeffreys prior), while the subjective Bayesian method utilizes humans' perspective or other available information as an informative prior for estimating the model parameters [29]. However, questions still arise among researchers regarding the appropriate determination and selection of the prior distribution in Bayesian methods. In fact, this is one of the most debated issues by researchers, because it influences the computation and results of the posterior distribution [30]. In recent years, the MCMC method has initiated broad prospects for the promotion and application of Bayesian methods so that it has become a reliable tool for handling complex problems in statistics and modeling practices [31].

Bayesian methods have been widely applied in forestry research in recent years, including the studies of tree diameter and height growth, mortality, tree biomass, etc. [1,13,21,24]. Zapata-Cuartas [13] established the aboveground biomass model for the Colombian tropical forest using Bayesian approaches incorporating 132 biomass equations from 72 published articles as the prior information.

Zhang et al. [32] utilized Bayesian approaches to estimate stem, branch, foliage, root, and total tree biomass models with both non-informative and informative priors. They proved that the Bayesian methods with informative priors produced better performances than those with non-informative priors and classical methods. In this sense, Bayesian statistics have the advantage of model performance in the case of fitting individual models. However, the biomass of different tree components is usually estimated jointly, while other studies found in published research were only based on a single biomass model, ignoring the inherent correlations among the biomass of tree components. Although the additive model systems with linear or nonlinear functions have been widely used and become a standard for developing individual tree biomass models [2,8], to the best of our knowledge, there has been no publication in the literature on the applications of Bayesian methods with SUR to jointly estimate the additive biomass model systems of tree components.

Korean larch is one of the most important fast-growing conifer species for timber production in the Northeast, P.R. China. The larch plantations encompass a wide range of geographical coverage, from the northeastern to the northern subalpine areas of China [33]. It is desirable and needed to maximize and maintain the ecological and economic benefits of the Korean larch plantations. Therefore, reliable biomass models with the capability of uncertainty assessment are necessary for estimating, accounting, and monitoring tree biomass and forest carbon stocking. In this study, we used data consisting of 174 destructively sampled Korean larch (*Larix olgensis* Henry) to develop the additive biomass model systems of tree components (i.e., stem, branch, foliage, and root) for planted Korean larch. Two additive model systems were constructed to facilitate the application of different data types; the first one used tree DBH as the only predictor, and the second one used both tree DBH and H as the predictors. For each model system, we compared two parameter estimation methods: classical FGLS (as the benchmark) and Bayesian methods. A comprehensive sample of allometric biomass equations for *Larix* spp. in different locations gathered from published literatures provided sufficient prior information of model parameters.

The objective of this study was to explore the advantages of Bayesian approaches on the SUR models of individual tree biomass. We expected that the SUR model systems using Bayesian methods would produce superior model fitting and predictions, and the Bayesian methods with the informative priors would perform better than non-informative priors. We investigated the impacts of five priors of model parameters on the model fitting and validation results: (1) direct Monte Carlo simulation with Jeffreys invariant prior (DMC); (2) Gibbs sampler using Jeffreys invariant prior (Gs-J); and (3) Gibbs sampler using three priors of multivariate normal distribution, i.e., artificial setting (Gs-MN), self-sampling estimate (Gs-MN1), and other research results on biomass models (Gs-MN2), respectively.

2. Materials and Methods

2.1. Study Sites and Data Collection

The data used in this study were collected from 9 sites of Korean larch plantations of different ages in Heilongjiang Province, northeast China (Figure 1). A total of 38 plots (20 × 30 m or 30 × 30 m in size) were established in July and August of 2007–2016. Different classes of sample trees (dominant, codominant, medium, and suppressed) were felled from each plot (5–7 trees/plot). In the fieldwork, tree variables such as DBH and H were measured and recorded first. Then all live branches in each tree were cut and weighed; the subsamples of branch and foliage were from an average-sized branch in each pseudo-whorl. After that, the stems were each cut into 1 m sections and weighed. A thick disc of about 3 cm was cut at the end of each 1 m section, weighed, and taken to a laboratory for further sampling. The green weight of roots included large (diameter more than 5 cm), medium (diameter between 2 cm to 5 cm), and small (diameter less than 2 cm) roots, and the fine root (diameter less than 5 mm) was not considered. The subsamples of roots were taken from large, medium, and small roots, respectively. In total, 174 sample trees were selected for both above- and below-ground biomass. All subsamples of each component were oven-dried at 80 °C using a blast drying oven, until a constant

weight was reached, and then weighed in the lab. The dry biomass of each tree component was calculated by the product of the green weight and dry/fresh ratio of the corresponding tree component. The tree DBH, H, and dry weight of each tree component are summarized in Table 1.

Figure 1. The distribution of the sampling plots in this study across Heilongjiang province, Northeast China.

Table 1. Summary statistics of the sample trees in this study.

Attributes	Mean	Std	Minimum	Maximum
DBH (cm)	16.3	7.0	2.0	35.7
H (m)	15.4	5.5	3.8	27.0
Root biomass (kg)	31.66	33.83	0.13	203.60
Stem biomass (kg)	106.94	107.50	0.27	510.19
Branch biomass (kg)	11.37	9.05	0.10	42.30
Foliage biomass (kg)	3.56	2.08	0.11	8.73

Note: std is the standard deviation.

2.2. Biomass Model

The power function of allometric equations has been commonly used for biomass modeling [10], in which tree biomass is the dependent variable and tree attributes (i.e., DBH and/or H) are the independent or predictor variables [34]. In this study, we applied a likelihood analysis to determine which of the two error structures (additive and multiplicative) was more appropriate to the biomass data of Korean larch [2,35]. It was found that the multiplicative error structure was preferred so that the logarithmic transformation was applied to linearize the power function to construct the linear

seemingly unrelated regression (SUR) models. Two additive model systems were constructed in this study as follows:

$$\text{SURM1} : \begin{cases} \ln W_s = \beta_{s0} + \beta_{s1} \ln DBH + \varepsilon_s \\ \ln W_b = \beta_{b0} + \beta_{b1} \ln DBH + \varepsilon_b \\ \ln W_f = \beta_{f0} + \beta_{f1} \ln DBH + \varepsilon_f \\ \ln W_r = \beta_{r0} + \beta_{r1} \ln DBH + \varepsilon_r \end{cases} \quad (1)$$

$$\text{SURM2} : \begin{cases} \ln W_s = \beta_{s0} + \beta_{s1} \ln DBH + \beta_{s2} \ln H + \varepsilon_s \\ \ln W_b = \beta_{b0} + \beta_{b1} \ln DBH + \beta_{b2} \ln H + \varepsilon_b \\ \ln W_f = \beta_{f0} + \beta_{f1} \ln DBH + \beta_{f2} \ln H + \varepsilon_f \\ \ln W_r = \beta_{r0} + \beta_{r1} \ln DBH + \beta_{r2} \ln H + \varepsilon_r \end{cases} \quad (2)$$

where W_s , W_b , W_f , and W_r are the biomass of stem, branch, foliage and root, respectively; DBH is the tree diameter at the breast height; H is the tree total height; $\ln$ is the natural logarithm; β_{kj} are the model coefficients to be estimated from the data, where the subscript $k = s, b, f$ and r corresponds to the four tree components, and $j = 0, 1, 2$; and ε is the model error term.

A general expression of linear SUR models (in matrix notation) is

$$Y = X\beta + \varepsilon \quad (3)$$

where Y is a vector of the response variable; ε is a vector of the error terms; X is a matrix of the predictors, including $\ln DBH$ for SURM1, and $\ln DBH$ and $\ln H$ for SURM2; β is a vector of the model coefficients; and $\varepsilon \sim N(0, \Omega \otimes I)$, where I is an identity matrix, Ω is the variance–covariance matrix of error term, and $\otimes$ is the tensor product, such that:

$$Y = \begin{pmatrix} \ln W_s \\ \ln W_b \\ \ln W_f \\ \ln W_r \end{pmatrix}, X = \begin{pmatrix} X_s & 0 & 0 & 0 \\ 0 & X_b & 0 & 0 \\ 0 & 0 & X_f & 0 \\ 0 & 0 & 0 & X_r \end{pmatrix}, \beta = \begin{pmatrix} \beta_s \\ \beta_b \\ \beta_f \\ \beta_r \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_s \\ \varepsilon_b \\ \varepsilon_f \\ \varepsilon_r \end{pmatrix}, \Omega = \begin{pmatrix} \sigma_s^2 & \sigma_s \sigma_b & \sigma_s \sigma_f & \sigma_s \sigma_r \\ & \sigma_b^2 & \sigma_b \sigma_f & \sigma_b \sigma_r \\ & & \sigma_f^2 & \sigma_f \sigma_r \\ & & & \sigma_r^2 \end{pmatrix} \quad (4)$$

2.3. Classical Approach of SUR

Generalized least squares (GLS) is an efficient technique for parameter estimation when the error terms between regression models are correlated [2,36]. However, if the covariance Ω of the model errors is unknown, one can get a consistent estimate of Ω using an implementable version of GLS known as the feasible generalized least squares (FGLS). The package `systemfit` in R software 3.5.3 [37,38] was used to obtain the classical estimations of model coefficients and the covariance matrix in this study.

2.4. Bayesian Approaches of SUR

In Bayesian inference, both model coefficients β and covariance Ω are considered random variables with probability distributions that can be calculated using the data D such that

$$\pi(\beta, \Omega | D) = \frac{\pi(\beta, \Omega) L(D | \beta, \Omega)}{\int \pi(\beta, \Omega) f(D | \beta, \Omega) d(\beta, \Omega)} \propto \pi(\beta, \Omega) L(D | \beta, \Omega) \quad (5)$$

where $\pi(\beta, \Omega | D)$ and $\pi(\beta, \Omega)$ are the posterior- and prior-probabilities, respectively, $L(D | \beta, \Omega)$ is the likelihood, and the denominator of Equation (5) is the marginal probability of the data D [13]. Based on Equation (5), algorithms of Bayesian approaches were proposed on the SUR models [39]. To implement the Bayesian estimation, we fit the SUR models using both the Gibbs sampler and direct Monte Carlo (DMC) sampling along with different priors [39] (the detailed calculation processes of the Gibbs sampler and DMC are attached in Appendix A).

We proposed five estimations of the SUR models as follows:

- (1) Direct Monte Carlo simulation approach using Jeffreys invariant prior (namely DMC).
- (2) Gibbs sampler using Jeffreys invariant prior (namely Gs-J).
- (3) Gibbs sampler using a multivariate normal distribution with mean vector and variance defined as 0 and 1000 (the covariance is 0) as priors of model parameters, and a default inverted Wishart distribution [25] as the priors of variance–covariance matrix errors (namely Gs-MN), respectively, which was considered as a non-informative prior.
- (4) Gibbs sampler using a multivariate normal distribution prior, which consisted of parameters estimated by the FGLS method with subsamples (namely Gs-MN1). The percentages of the subsamples were 10%, 20%, 30%, ..., and 90% of the 174 independent trees data, which were sampled without conducting any replacement. Each simulation was repeated 10,000 times.
- (5) Gibbs sampler using a prior of multivariate normal distribution, which consisted of the parameters estimated by using logarithmic functions of different biomass components for *Larix* spp. (namely Gs-MN2). The previous biomass model parameters in the same forms for *Larix* spp. were collected from a normalized tree biomass equation dataset in Luo et al. [10]. A default inverted Wishart distribution proposed by Rossi and Allenby was used as the prior of variance–covariance matrix errors [25]. Even though the biomass functions were sampled from different areas in China, the model parameters were well represented by a multivariate normal distribution with mean vector and variance–covariance matrix after a multivariate normal test ($p > 0.05$).

In this study, the FGLS estimates of the model parameters were used as the starting values for the Gs-J method to reach faster convergence and save computational time. For the Gs-MN, Gs-MN1, and Gs-MN2 methods, the starting values were automatically generated in the R package bayesm when the prior distributions were given [40]. On the setting of options for the Gibbs sampler, the size of the posterior samples was set to 33,000, including 3000 burn-in periods and thinning length of 3 to ensure the performance's stability along with the relatively small autocorrelation of Markov chains constructed. The thinning length of 3 means that we took 1 sample for every 3 samples from the remaining 30,000 samples obtained and subsequently received 10,000 posterior samples. This set of the thinning length is beneficial to reduce the autocorrelations between samples. The convergence of the Gibbs posterior samples was assessed using the Geweke's and the Heidelberger and Welch's convergence diagnostic [41–43]. For the DMC approach, the process was repeated 10,000 times, corresponding to the Gibbs posterior samples. We used the R package bayesm to implement a Gibbs sampler with a prior multivariate normal distribution [40] and package Coda [43] to conduct the convergence diagnostic. However, there were no published packages that could implement Bayesian analysis with Jeffreys's prior, and thus we used custom R code to implement the Bayesian estimation of DMC and Gs-J.

2.5. Model Evaluation and Validation

In this study, the deviance information criterion (DIC) was used to evaluate and compare the model fitting of the Bayesian estimations with different priors [28]. In addition, the coefficient of determination (R^2) and root mean square error (RMSE) were calculated for the description of model performance [44]. The models with higher R^2 and lower DIC and RMSE were preferred.

Model validation is necessary to assess the applicability of the models, because better model fitting does not necessarily indicate better model prediction quality. In this study, we used the leave-one-out cross-validation to test the model prediction performance [45]. Mean absolute error (MAE) and mean absolute percent error (MAE%) were calculated to validate the model prediction performance [2]. The mathematical formula of the DIC, R^2 , RMSE, MAE, and MAE% can be expressed as follows:

$$DIC = \bar{D} + p_D \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (7)$$

$$RMSE = \sqrt{\left(\sum_{i=1}^n (Y_i - \hat{Y}_i) / n \right)^2 + \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 / (n-1)} \quad (8)$$

$$MAE = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_{i,-i}|}{n} \quad (9)$$

$$MAE\% = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_{i,-i}| / Y_i}{n} \times 100 \quad (10)$$

where $\bar{D}$ is the posterior mean of the deviance ($-2 \times \text{Log likelihood}$ of the given data and parameters); p_D is the model complexity, which is summarized by the effective number of parameters; Y_i and $\hat{Y}_i$ are the observations and predictions, respectively; $\hat{Y}_{i,-i}$ are the predictions obtained by using leave-one-out cross-validation; and n is the sample size.

2.6. Anti-Logarithm Correction Factors

When the logarithmic transformation is applied to the power function, the correction factors (CFs) are commonly used to correct the bias produced by the anti-log transformation. Several CFs from previous studies were compared in this study [2,46,47]:

$$CF1 = \exp(S^2/2) \quad (11)$$

$$CF2 = 1 + S^2/2 \quad (12)$$

$$CF3 = \exp\left(\frac{S^2}{2} \left(1 - S^2 \frac{S^2 + 2}{4n} + S^4 \frac{3S^4 + 44S^2 + 84}{96n^2}\right)\right) \quad (13)$$

where S^2 is the mean square error and calculated as $S^2 = \sum (Y_i - \hat{Y}_i)^2 / (n - p)$; n is the sample size; p is the number of model parameters; and Y_i and $\hat{Y}_i$ are the observations and predictions, respectively, taken from Equations (1) and (2). To compare the effectiveness of the three CFs, three measures were used [2,48,49] as follows:

$$B = \frac{CF - 1}{CF} \times 100 \quad (14)$$

$$G = \sqrt{CF - 1} \times 100 \quad (15)$$

$$MPD = \frac{\sum (Y_i - \hat{Y}_i) / Y_i \times 100}{n} \quad (16)$$

where B is percent bias, G is percent standard error, and MPD is mean percentage difference; Y_i and $\hat{Y}_i$ represent anti-log observations and anti-log predictions, respectively, by CFs; and n is the sample size.

2.7. Stability Analysis for Classical and Bayesian Approaches

Simulations were performed using the subsamples of all biomass data to compare the classical SUR models (FGLS) against the Bayesian SUR models. Specifically, we compared the characteristics of the model parameters and prediction performance using different sample sizes. In the simulation process, we sampled 10, 15, 20, 30, 40, 60, 80, 110, 140, and 170 individuals without replacement from the entire dataset (174 trees), and then the procedure was repeated 5000 times. Within each subsample, we estimated the model parameters using both FGLS and Bayesian estimations. The distributions of estimated parameters were analyzed, and corresponding predictions were summarized and used to

calculate the indices of the efficiency of estimation (MAE and MAE%). All simulations were performed using R software 3.5.3 [38].

3. Results

3.1. Fitting the SUR Models

For the Bayesian methods Gs-MN1 on both SURM1 and SURM2, the prior information consisted of the model parameter values estimated by the FGLS method with subsample sizes of 10%, 20%, 30%, ..., and 90% of the 174 trees sampled without replacement. For Gs-MN2, the prior information in the normalized tree biomass equation dataset [10] was used to estimate the SURM1 and SURM2 model parameters. The results of statistical tests showed that these derived parameters values indeed followed a multivariate normal distribution. The specific forms and values were as follows:

Gs-MN1 of SURM2:

$$\bar{\beta} = (-4.0167, 2.7273, -3.3781, 2.7429, -3.5230, 2.0626, -2.9308, 1.4780)^T,$$

$$\bar{A} = \begin{pmatrix} 1760 & 5024 & -801 & -2212 & -43 & -198 & 7 & 126 \\ 5024 & 14,633 & -2220 & -6349 & -163 & -652 & 77 & 545 \\ -801 & -2220 & 2210 & 6090 & 673 & 2017 & -52 & -267 \\ -2212 & -6349 & 6090 & 17,393 & 1985 & 5974 & -258 & -1030 \\ -43 & -163 & 673 & 1985 & 1357 & 3866 & -726 & -2023 \\ -198 & -652 & 2017 & 5974 & 3866 & 11,273 & -2011 & -5821 \\ 7 & 77 & -52 & -258 & -726 & -2011 & 1473 & 4082 \\ 126 & 545 & -267 & -1030 & -2023 & -5821 & 4082 & 11,717 \end{pmatrix},$$

$$v_0 = 4,$$

$$V_0 = \begin{pmatrix} 0.0668 & 0.0270 & -0.0193 & -0.0074 \\ 0.0270 & 0.0547 & -0.0344 & -0.0161 \\ -0.0193 & -0.0344 & 0.1169 & 0.0566 \\ -0.0074 & -0.0161 & 0.0566 & 0.0811 \end{pmatrix};$$

Gs-MN1 of SURM2:

$$\bar{\beta} = (-5.0939, 2.2007, 0.7288, -4.4558, 1.6630, 1.1908, -2.6587, 2.9240, -1.1908, -2.5218, 1.8870, -0.5649)^T,$$

$$\bar{A} = \begin{pmatrix} 1760 & 5045 & 4980 & -638 & -1786 & -1725 & -7 & -71 & -46 & 46 & 196 & 181 \\ 5045 & 14,732 & 14,478 & -1728 & -5067 & -4820 & -58 & -298 & -231 & 159 & 679 & 614 \\ 4980 & 14,478 & 14,268 & -1671 & -487 & -4627 & -38 & -244 & -176 & 162 & 662 & 606 \\ -6380 & -1728 & -1671 & 7547 & 20,716 & 20,647 & 233 & 878 & 783 & -29 & -227 & -190 \\ -1786 & -5067 & -4837 & 20,716 & 58,708 & 57,962 & 837 & 2870 & 2639 & -162 & -835 & -721 \\ -1725 & -4820 & -4627 & 20,647 & 57,962 & 57,478 & 759 & 2684 & 2447 & -1372 & -7545 & -653 \\ -7 & -58 & -387 & 233 & 837 & 759 & 1319 & 3732 & 3668 & -696 & -1939 & -1916 \\ -71 & -298 & -244 & 878 & 2871 & 2685 & 3732 & 10,845 & 10,857 & -1922 & -5550 & -5423 \\ -45 & -231 & -176 & 783 & 2639 & 2447 & 3668 & 10,587 & 10,373 & -1913 & -5455 & -5359 \\ 46 & 159 & 162 & -29 & -162 & -137 & -696 & -1922 & -1913 & 1430 & 3954 & 3826 \\ 196 & 679 & 662 & -227 & -835 & -754 & -1939 & -5550 & -5455 & 3954 & 11,323 & 11,127 \\ 181 & 6149 & 606 & -190 & -721 & -653 & -1916 & -5423 & -5359 & 3926 & 11,128 & 10,990 \end{pmatrix},$$

$$v_0 = 4,$$

$$V_0 = \begin{pmatrix} 0.0568 & 0.0060 & -0.0025 & 0.0005 \\ 0.0060 & 0.0117 & -0.0001 & 0.0002 \\ -0.0025 & -0.0001 & 0.0899 & 0.0438 \\ 0.0005 & 0.0002 & 0.0438 & 0.0752 \end{pmatrix};$$

Gs-MN2 of SURM1:

$$\bar{\beta} = (-4.0167, 2.4730, 0.7288, -2.9983, 2.5546, 4.4397, -2.4903, 4.5258, -2.1122)^T,$$

$$\bar{A} = \begin{pmatrix} 0.7675 & -0.2753 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.2753 & 0.1203 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.8751 & -0.3542 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & -0.3542 & 0.1554 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.1078 & -0.3631 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.3631 & 0.1380 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0443 & -0.3479 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.3479 & 0.1581 \end{pmatrix};$$

Gs-MN2 of SURM2:

$$\bar{\beta} = (-3.8163, 1.7587, 0.8807, -3.3932, 1.8392, 0.9196, -4.6487, 1.7310, -0.8655, -5.0326, 1.5111, 0.7555)^T,$$

$$\bar{A} = \begin{pmatrix} 0.5112 & -0.2052 & -0.5271 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.2052 & 11.5670 & 6.4188 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.5271 & 6.4188 & 43.0293 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.9842 & 1.1609 & -0.1566 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.1609 & 14.2652 & -2.0570 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -0.1566 & -2.0570 & 51.6092 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.2258 & 0.1337 & -0.01687 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.1337 & 3.5121 & -1.2614 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.0169 & -1.2614 & 13.7326 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.1698 & -0.0415 & -0.1603 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.0415 & 2.6080 & 0.5663 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.1603 & 0.5663 & 10.5430 \end{pmatrix};$$

where in $\beta \sim N(\bar{\beta}, A^{-1})$, $\bar{\beta}$ is the prior mean, and A is the prior precision matrix; in $\Omega \sim InvWishart(\nu_0, V_0)$, ν_0 is the degree of freedom parameter for the inverted Wishart prior, and V_0 is the scale parameter for the inverted Wishart prior.

The posterior means, standard deviations, and 95% posterior intervals for the Bayesian methods of both SURM1 and SURM2 are reported in Tables 2 and 3, respectively, in which the posterior draws from the Gibbs sampler successfully passed the convergence test. Furthermore, the parameter estimates and standard errors of FGLS are also listed in Tables 2 and 3. The means of the posterior draws from different sampling were relatively equal, but there was a significant discrepancy in their standard deviations. The tables indicated that DMC produced the lowest standard deviation, followed by Gs-MN1, Gs-J, Gs-MN2, and Gs-MN, as did the 95% posterior intervals. In order to simplify the comparison, Figure 2 visualizes the results of the posterior draws in the form of probability density plots for the Bayesian methods. The same trend was clearly shown for each model parameter (Figure 2). Furthermore, we calculated the correlation matrix of the model errors between the four tree components after the model parameters were obtained. However, those correlation matrices were not significantly different for the similar parameter estimates between the six modeling methods. Hence, we only illustrated the correlation matrix for DMC in Table A1 (see Appendix B).

Table 2. The posterior means, standard deviations (std), and 95% posterior intervals (using the 2.5th and 97.5th percentiles of posterior samples) of the SURM1 (DBH-only) parameter estimates, calculated by using the Bayesian estimation. The parameters means and standard errors (SEs) estimated by FGLS are also listed.

Methods	Stats	β_{r0}	β_{r1}	β_{s0}	β_{s1}	β_{b0}	β_{b1}	β_{f0}	β_{f1}
DMC	Mean	−4.5406	2.7194	−3.3655	2.7386	−3.5304	2.065	−2.9406	1.4810
	Std	0.0277	0.0102	0.0253	0.0093	0.0508	0.0187	0.0418	0.0153
	2.50th	−4.5945	2.6995	−3.4146	2.7201	−3.6301	2.0291	−3.0217	1.4512
	97.50th	−4.4868	2.7391	−3.3157	2.7567	−3.4325	2.1018	−2.8585	1.5109
Gs-J	Mean	−4.5404	2.7192	−3.3651	2.7384	−3.5319	2.0656	−2.9424	1.4819
	Std	0.1079	0.0396	0.0971	0.0356	0.1437	0.0526	0.1192	0.0436
	2.50th	−4.7519	2.642	−3.5549	2.6688	−3.8137	1.96	−3.1756	1.3953
	97.50th	−4.3302	2.7984	−3.1728	2.8092	−3.2475	2.1692	−2.7081	1.5677
Gs-MN	Mean	−4.5397	2.7191	−3.3662	2.739	−3.5277	2.064	−2.9393	1.4806
	Std	0.1346	0.0493	0.1267	0.0465	0.1627	0.0597	0.1433	0.0525
	2.50th	−4.8036	2.6217	−3.6166	2.649	−3.8446	1.947	−3.2192	1.3778
	97.50th	−4.2763	2.8155	−3.1188	2.8302	−3.2162	2.1815	−2.6585	1.5827
Gs-MN1	Mean	−4.545	2.721	−3.3687	2.7397	−3.5304	2.065	−2.938	1.4803
	Std	0.0944	0.0339	0.0779	0.0283	0.1202	0.0433	0.1008	0.0365
	2.50th	−4.7309	2.654	−3.517	2.6838	−3.7696	1.9801	−3.1332	1.4079
	97.50th	−4.359	2.7875	−3.2142	2.7942	−3.2959	2.1517	−2.7393	1.5505
Gs-MN2	Mean	−4.5209	2.7124	−3.3441	2.731	−3.585	2.0848	−2.9947	1.5008
	Std	0.1339	0.0493	0.1238	0.0455	0.1588	0.0582	0.1382	0.0507
	2.50th	−4.7838	2.6145	−3.5829	2.6418	−3.8963	1.9705	−3.2654	1.4017
	97.50th	−4.2546	2.8078	−3.0999	2.8184	−3.2729	2.1992	−2.7241	1.5996
FGLS	Value	−4.5407	2.7195	−3.3655	2.7386	−3.5305	2.0650	−2.9402	1.4810
	SE	0.1067	0.0391	0.0961	0.0353	0.1409	0.0517	0.1173	0.0430

3.2. Model Evaluation and Validation

Table 4 presents the R^2 , RMSE, and DIC of the six modeling methods for the four tree components (i.e., stem, branch, foliage, and root). According to the DICs, Gs-MN1 performed better than other methods either for SURM1 (DBH as the only predictor) or SURM2 (DBH and H as the predictors). For the non-informative prior Bayesian methods, there was not much difference of DIC between DMC and Gs-J, but there was a significant reduction in DIC for Gs-MN. For the informative priors, only Gs-MN1 produced a relatively better performance in most tree components, while Gs-MN2 performed similarly with other non-informative priors. The results also indicated that all models fitted the larch biomass data well, given R^2 all higher than 0.75 in the model (Table 4). The stem biomass models obtained the highest R^2 , while the foliage biomass models obtained the smallest R^2 . Incorporating H as the second predictor (SUM2) improved the model fitting for the stem, branch, and foliage biomass models, while unexpectedly it resulted in a slight reduction in the model R^2 for the root biomass models. In general, the Gs-MN1 method for both SURM1 and SURM2 yielded smaller DIC and relatively similar R^2 and RMSE compared with other modeling methods.

The averages and 95% prediction intervals (PIs) for the DBH-only classical method (FGLS), non-informative (DMC, Gs-J, and Gs-MN), and informative priors (Gs-MN1 and Gs-MN2) Bayesian prediction curves are showed in Figure 3. It is clear that the 95% PIs of the five Bayesian methods were much narrower than the FGLS models across all tree components. The results indicate the Bayesian models are more reliable and stable than FGLS. Overall, DMC evidently produced the narrowest 95% PI among the five Bayesian methods, followed by Gs-MN1. On the other hand, the differences between the other three Bayesian methods (i.e., Gs-J, Gs-MN, Gs-MN2) were indistinguishable.

Table 3. The parameter estimates of SURM2 (DBH-H). The symbols are the same as in Table 2.

Method	Stats	β_{r0}	β_{r1}	β_{r2}	β_{s0}	β_{s1}	β_{s2}	β_{b0}	β_{b1}	β_{b2}	β_{f0}	β_{f1}	β_{f2}
DMC	Mean	−5.0773	2.1886	0.7354	−4.451	1.6653	1.4869	−2.6606	2.9253	−1.1918	−2.5255	1.8912	−0.5683
	Std	0.0329	0.0241	0.0313	0.0079	0.0058	0.0075	0.0522	0.0381	0.0495	0.0501	0.0371	0.0479
	2.50th	−5.1418	2.1407	0.6747	−4.4666	1.6535	1.4717	−2.764	2.8507	−1.2886	−2.6239	1.8189	−0.6635
	97.50th	−5.0132	2.236	0.7973	−4.4353	1.6768	1.5017	−2.5594	2.9998	−1.0945	−2.4265	1.9645	−0.4767
Gs-J	Mean	−5.0756	2.1889	0.7344	−4.4504	1.6651	1.4868	−2.6622	2.924	−1.1896	−2.5266	1.8901	−0.5667
	Std	0.1382	0.1022	0.1318	0.0628	0.0460	0.0598	0.1772	0.1294	0.1684	0.1600	0.1166	0.1511
	2.50th	−5.3515	1.9858	0.4751	−4.573	1.5743	1.369	−3.0063	2.6692	−1.5186	−2.842	1.6595	−0.8624
	97.50th	−4.8045	2.3888	0.9943	−4.3249	1.756	1.6041	−2.31	3.1811	−0.8569	−2.2116	2.1221	−0.2764
Gs-MN	Mean	−5.075	2.1876	0.7354	−4.4496	1.6642	1.4874	−2.6599	2.9266	−1.1934	−2.5245	1.8934	−0.5708
	Std	0.1798	0.1312	0.1694	0.1321	0.096	0.1245	0.2062	0.1515	0.1964	0.1925	0.1432	0.1846
	2.50th	−5.4231	1.9314	0.3979	−4.7091	1.4758	1.2382	−3.0647	2.6302	−1.5807	−2.9043	1.6134	−0.9359
	97.50th	−4.7202	2.4472	1.0676	−4.1913	1.8536	1.7293	−2.2546	3.2223	−0.8115	−2.147	2.1742	−0.2104
Gs-MN1	Mean	−5.0835	2.1907	0.7353	−4.4515	1.6653	1.487	−2.661	2.9246	−1.1908	−2.5249	1.8902	−0.5672
	Std	0.1148	0.0867	0.11	0.0536	0.0395	0.0516	0.1418	0.1088	0.1372	0.1327	0.0986	0.1271
	2.50th	−5.308	2.0222	0.517	−4.5562	1.5889	1.386	−2.9347	2.7129	−1.457	−2.7884	1.6964	−0.8195
	97.50th	−4.8589	2.3608	0.9504	−4.3455	1.742	1.5856	−2.3849	3.1356	−0.924	−2.2669	2.0862	−0.3151
Gs-MN2	Mean	−5.081	2.1356	0.7902	−4.2369	1.869	1.2016	−3.342	2.2326	−0.2395	−2.9663	1.461	0.0299
	Std	0.1524	0.0971	0.1190	0.1182	0.0756	0.0941	0.2009	0.1313	0.167	0.1898	0.1296	0.1653
	2.50th	−5.3795	1.9463	0.559	−4.4704	1.7196	1.0191	−3.7383	1.9677	−0.5602	−3.3426	1.2026	−0.2873
	97.50th	−4.7805	2.3257	1.0196	−4.0046	2.0177	1.3876	−2.952	2.4869	0.0949	−2.5913	1.7106	0.3584
FGLS	Mean	−5.0777	2.1885	0.7355	−4.4511	1.6652	1.4870	−2.6605	2.9251	−1.1915	−2.5252	1.8913	−0.5684
	Se	0.1370	0.1010	0.1307	0.0620	0.0457	0.0592	0.1722	0.1270	0.1643	0.1575	0.1161	0.1503

SURM1

SURM2

Figure 2. Posterior distributions of the model parameters estimated by Bayesian statistics. DMC represents the direct Monte Carlo with Jeffreys' invariant prior; Gs-J is the Gibbs sampler with Jeffreys' invariant prior; and Gs-MN represents the Gibbs sampler with a prior of multivariate normal distribution by 0 vector, with the corresponding variance of 1000 and the covariance of 0. The Gs-MN1 and Gs-Mn2 is the Gibbs sampler with an informative prior of multivariate normal distribution by different means, variances, and covariances.

Table 4. The model fitting statistics of the six modeling methods for the four tree biomass components.

Method	DIC	R ²				RMSE (kg)			
		W _r	W _s	W _b	W _f	W _r	W _s	W _b	W _f
SURM1									
DMC	−1894.25	0.9604	0.9385	0.8201	0.7625	6.7588	26.6776	3.8632	1.0172
Gs-J	−1890.65	0.9604	0.9385	0.8200	0.7619	6.7630	26.6763	3.8655	1.0188
Gs-MN	−719.567	0.9604	0.9385	0.8204	0.7626	6.7582	26.6846	3.8601	1.0171
Gs-MN1	−1939.09	0.9606	0.9384	0.8200	0.7625	6.7453	26.6914	3.8643	1.0175
Gs-MN2	−719.14	0.9598	0.9389	0.8152	0.7547	6.8180	26.5947	3.9231	1.0353
FGLS	-	0.9605	0.9385	0.8201	0.7625	6.7552	26.6785	3.8636	1.0175
SURM2									
DMC	−3004.07	0.9575	0.9840	0.8478	0.7692	6.9882	13.6087	3.5441	1.0022
Gs-J	−2993.10	0.9575	0.9840	0.8476	0.7691	6.9882	13.6063	3.5481	1.0023
Gs-MN	−1274.25	0.9574	0.9840	0.8478	0.7688	7.0005	13.6026	3.5440	1.0030
Gs-MN1	−3046.41	0.9577	0.9840	0.8478	0.7689	6.9736	13.6096	3.5448	1.0028
Gs-MN2	−1163.85	0.9551	0.9824	0.8368	0.7600	7.1937	14.3112	3.6719	1.0231
FGLS	-	0.9575	0.9840	0.8478	0.7691	6.9885	13.6088	3.5448	1.0023

Figure 3. The 95% prediction interval of W_r , W_s , W_b , and W_f predictions from the six approaches.

Leave-one-out cross-validation was used to validate the log-transformed SUR models for the six methods (Table 5). According to MAE and MAE%, the model prediction bias varied across different tree components. The stem and branch models consistently produced the smallest and the largest MAE%, respectively, across all methods for both SURM1 and SURM2, while Gs-MN1 yielded the best prediction for both SUM1 and SUM2 for most tree components (Table 5).

Table 5. The model validation statistics of six modeling methods for the four tree biomass components.

Method	W_r		W_s		W_b		W_f	
	MAE (kg)	MAE% (%)	MAE (kg)	MAE% (%)	MAE (kg)	MAE% (%)	MAE (kg)	MAE% (%)
SURM1								
DMC	4.2952	20.2532	17.4227	19.6344	2.5116	29.0780	0.7296	23.5099
Gs-J	4.2864	20.2547	17.4221	19.6383	2.5145	29.1252	0.7299	23.5313
Gs-MN	4.0020	20.6052	17.3580	20.1764	2.5846	30.1746	0.7438	24.4397
Gs-MN1	3.9844	20.5028	17.2887	20.0967	2.5735	30.0245	0.7408	24.3333
Gs-MN2	4.0425	20.7010	17.3555	20.1893	2.6148	30.1029	0.7532	24.3886
FGLS	4.0163	20.5847	17.3484	20.1571	2.5791	30.0815	0.7425	24.3796
SURM2								
DMC	4.3199	18.5282	8.0317	8.5308	2.3198	25.2545	0.7465	23.4813
Gs-J	4.3205	18.5531	8.0323	8.5320	2.3232	25.3086	0.7472	23.5177
Gs-MN	4.3194	19.2793	7.9945	8.5978	2.3673	26.1265	0.7555	24.3034
Gs-MN1	4.2878	19.1338	7.9593	8.5451	2.3530	25.9476	0.7511	24.1500
Gs-MN2	4.2764	19.2940	8.0025	8.6385	2.4473	25.9014	0.7967	24.3156
FGLS	4.3155	19.2164	7.9995	8.5863	2.3615	26.0279	0.7542	24.2290

3.3. Comparison of Correction Factors on Anti-Log Transformation

In order to further analyze the effects of applying different correction factors, the percent bias (B, %) and percent standard error (G, %) of three correction factors (CF1, CF2, and CF3) for both SURM1 and SURM2 were calculated using the Gs-MN1 method (Table 5). In addition, the mean percentage difference (MPD, %) of CF0, CF1, CF2, and CF3 was also listed in Table 6 (CF0 = 1, without applying any correction factor). Surprisingly, SURM2 always had smaller values of correction factors (B, G, and MPD) across the four tree components than those of SURM1, indicating that adding H into the biomass equations indeed improved the model fitting and performance. However, the values of the correction factors, B, and G were relatively similar for the CF1, CF2, and CF3, revealing that the three correction factors were indistinguishable for either SURM1 or SURM2 in this study. Further, the mean percentage difference of the uncorrected prediction (MPD0) was the lowest value (Table 6), so that the correction factor might not be necessary for the anti-log transformation in this study.

3.4. Stability Analysis in Repeated Trials between Various Sizes of Data and Methods

In this study, the parameter estimates obtained by FGLS and Bayesian methods were similar using the entire biomass data (Tables 2 and 3). However, the two methods produced the parameter estimates with different ranges when the sample size decreased, in which the FGLS models had a wider range of the parameter estimates than those of the Bayesian models on SURM1 (Figure 4), in which the five models (except Gs-J) were selected to illustrate the stability of the parameter estimation methods. Figure A1 (see Appendix B) shows the trend of SURM2, which is similar to that of SURM1. It was evident that the ranges of the parameter estimates decreased as the subsample size increased. The distributions of the parameter estimates were relatively identical for each subsample size in DMC and FGLS, in which both methods estimated the parameters outside the confidence intervals obtained with the entire data (174 trees). For the other non-informative priors, the simulation results of Gs-MN were slightly different from both DMC and FGLS. The Bayesian methods with the informative priors produced less variation or fluctuation for the parameter estimates with narrower ranges than those with the non-informative priors (Figures 4 and 5). For the small subsamples of Gs-MN2, the means of the parameter estimates clearly deviated from the means of the parameter estimates obtained with the entire data, while Gs-MN1 performed well in each subsample size. The results indicated that using an inaccurate informative prior will produce incorrect results.

Table 6. The percent bias (B, %) and percent standard error (G, %) of the three correction factors (CF1, CF2, and CF3) with the mean percentage difference (MPD, %) of the CF0, CF1, CF2, and CF3 calculated by the error terms obtained using the Gs-MN1. MPD0 was calculated without using any correction factor (CF0), in which CF0 = 1.

Components	CF1	B1	G1	CF2	B2	G2	CF3	B3	G3	MPD0	MPD1	MPD2	MPD3
SURM1													
W_r	1.0342	3.3069	18.4932	1.0336	3.2508	18.3303	1.0342	3.3069	18.4932	19.9470	20.3233	20.3114	20.3232
W_s	1.0277	2.6953	16.6433	1.0273	2.6575	16.5227	1.0277	2.6953	16.6433	19.3369	19.9458	19.9364	19.9457
W_b	1.0604	5.6960	24.5764	1.0586	5.5356	24.2074	1.0604	5.6960	24.5764	27.4698	29.6666	29.5909	29.6656
W_f	1.0415	3.9846	20.3715	1.0406	3.9016	20.1494	1.0415	3.9846	20.3715	23.0224	24.0755	24.0451	24.0752
SURM2													
W_r	1.0289	2.8088	17.0000	1.0285	2.7710	16.8819	1.0289	2.8088	17.0000	18.0167	18.8596	18.8457	18.8595
W_s	1.0059	0.5865	7.6811	1.0058	0.5767	7.6158	1.0059	0.5865	7.6811	8.3797	8.4365	8.4363	8.4365
W_b	1.0461	4.4068	21.4709	1.0451	4.3154	21.2368	1.0461	4.4068	21.4709	24.0403	25.5298	25.4932	25.5293
W_f	1.0384	3.6980	19.5959	1.0377	3.6330	19.4165	1.0384	3.6980	19.5959	22.8702	23.7830	23.7636	23.7827

Figure 4. Simulation results for the parameters estimation of SURM1 using the DMC, Gs-MN, Gs-MN1, Gs-MN2, and FGLS. The dashed and solid horizontal lines represent the 95% confidence intervals and means of the simulation using the entire dataset (174 trees), respectively.

The MAEs and MAE% were calculated using SURM1 produced in the simulations and reapplied to the entire data (Figure 5); the simulation result of SURM2 was highly similar to that of SURM1, as shown in Figure A2 (see Appendix B). The summary of MAEs and MAE% for all methods and tree components across ten different sample sizes are given in Table A2 (see Appendix B). SURM2 had a higher prediction efficiency than SURM1 from the values of the MAEs and MAE%. Further, the prediction bias of Gs-MN1 was most stable, indicating that a good informative prior has the advantages of using a small number of data. For Gs-MN2, the prediction bias calculated using the parameter estimates obtained from the small samples deviated significantly from those calculated using all data, although the ranges of MAE and MAE% were smaller than both Bayesian models with non-informative priors and FGLS.

A

B

Figure 5. The MAE (A) and MAE% (B) of the SURM1 using different sample sizes. Solid horizontal line represents the value of MAE and MAE% calculated using the entire dataset (174 trees).

4. Discussion

4.1. Biomass SUR Models

Tree biomass is the basis for estimating net primary production and carbon of forest ecosystems, especially when the global carbon emissions and carbon sinks have gradually become the emergent topics [50,51]. Usually, allometric equations are used to develop tree biomass models for specific species and areas [8,14]. In our research, two additive model systems (i.e., SURM1 and SURM2) were constructed to estimate tree biomass of planted Korean larch by both classical and Bayesian approaches. The main objective of this study was to explore the advantages of Bayesian approaches on the SUR biomass models of the four tree components (i.e., stem, branch, foliage, and root). As we expected, our results demonstrated that applying the Bayesian methods on the SUR model systems not only obtained better model fitting and predictions, but also provided the assessment and evaluation of the uncertainties for constructing and updating tree biomass models.

For computational efficiency, we compared the computing time for both FGLS and Bayesian approaches on the SURM1 estimations by a computer of a 3.2 GHz Intel (R) Core (TM) i7—8700 CPU and RAM of 32.0 GB. The result showed that it took about 0.46 s to obtain the posterior samples of parameters for the Bayesian approaches (i.e., Gs-MN, Gs-MN1 and Gs-MN2) when the sample size was 33,000 (including 3000 burn-in periods and 3 thinning length) using the R package bayesm, while the

classical FGLS took only 0.03 s to get the parameter estimates using the R package `systemfit`. It was clear that the computing time of both approaches was small enough to be ignored. However, the DMC and Gs-J took more time (about 15 min) to estimate the models of the same setup described above, which may be caused by the limitation of our customized R codes, and the computation would have been faster if the codes had been optimized.

The published research revealed that the model parameters of biomass equations can be described by a multivariate normal distribution. In other words, the variability in these model parameters represented different biotic and abiotic factors influencing its own allometric biomass relationships for specific species and locations [13]. With the Bayesian approaches, the model parameters of new locations can be estimated and updated from the prior information, which was different from the calibration of the random-effect parameters in mixed-effect models [14], because the random-effect parameters were fixed for a specific level. Compared to the classical approaches, the Bayesian approaches produced narrower 95% prediction intervals so that the Bayesian statistics have the advantages of reducing uncertainty and obtaining more stable predictions (Figure 3) [13,32], and the good prior information improved the accuracy of biomass predictions (Table 5). However, the Bayesian approaches sometimes performed worse if the prior information was not good, like Gs-MN2.

Our results of stability analysis showed the evidence, as we expected, that the Bayesian approaches with informative priors provided more stable parameter estimates and reduced the uncertainty of parameters for small sample sizes, similar to the findings of Zapatacuartas et al. [13] (also see more discussion in the Section 4.2). The smaller fluctuations of both MAE and MAE% for the Bayesian methods with the informative priors indicated that the Bayesian methods had higher estimation efficiency than the classical method, especially for small sample sizes ($N = 10, 20, 30$). However, the difference in the estimation efficiency between FGLS and Bayesian methods decreased as the sample size increased. Overall, the sample size reflected the amount of the prior data information, of which the values of the posterior parameters mainly depended [28,30]. Increasing sample size continuously corrected the prior distribution until it significantly improved the posterior distribution. This phenomenon can be considered as one of the reasons that the estimations of non-informative Bayesian and FGLS were identical.

Considering the inherent correlations between tree components (additivity or compatibility) in developing biomass equations has been recognized as an essential characteristic of the biomass model systems [2,4]. Numerous SUR models in the literature commonly used one or even more constraints in the simultaneous model fitting process [18,52]. Zhao et al. [53], Dong et al. [14], and Widagdo et al. [54] conducted a comprehensive simulation and comparison of applying zero, one, and three restriction(s) in the SUR model systems. All of these researches confirmed that applying SUR without using any constraint, as proposed by Affleck and Diéguez-Aranda [55], gave slightly better model predictions than those using one or more constraints. Therefore, we decided to use no constraint to account for the correlations between tree components. Our results indicated that the correlation coefficients ranged from 0.1062 to 0.5836 across the tree components for SURM1, while they were slightly smaller (from 0.0026 to 0.5432) for SURM2. Therefore, it was appropriate to jointly fit the simultaneous biomass equations.

The sole DBH predictor of the SUR model (SURM1) produced relatively large variations for the tree components, and the inclusion of H as the second predictor variable in stem, branch, and foliage biomass models (SURM2) yielded a relatively greater enhancement on the model performance for both model fitting and validation. Our results were compatible with previous publications [8,56]. Further analyses showed that the root biomass model's performance using DBH as the sole predictor was better than those of two predictors (DBH and H), although the parameter estimates of H were statistically significant. These findings indicated that the benefit of adding H as the additional predictor on root biomass models was relatively small [56,57]. Overall, SURM2 apparently produced better parameter estimation and model prediction than SURM1 for three of the four tree components biomass.

4.2. Non-Informative vs. Informative Priors in Bayesian Methods

Both DMC and Gs-J used the Jeffreys non-informative priors but obtained slightly different results for model fitting. The DIC values indicated that DMC was better than Gs-J, with smaller standard deviations (Tables 2 and 3, and Figure 2). The Gs-J method needed to check on the posterior samples' convergence, while DMC did not. Zellner [39] reported one of the advantages of DMC, in which DMC might generate independent samples within a weak autocorrelation. Hence, we also evaluated the posterior samples' autocorrelation of the parameter β_{r0} (Table A3 in Appendix B). The autocorrelation of various methods at every non-zero lag was very small, indicating that the autocorrelation in the posterior samples rarely existed. Thus, it may not be apparent that the DMC's autocorrelation was less than those of other methods based on the Gibbs sampler, which was slightly different from Zellner's inference [39]. As described in Section 2.4, the thinning length used in the Gibbs sampler's posterior samples was set to 3 in order to weaken the autocorrelation, but this step was absent in Zellner [39].

The DMC method seemed to be more efficient in the computation process according to the real size of the generated posterior samples (DMC method was 10,000 and Gibbs was 33,000), similar to the inferences in other studies [58]. However, the computational efficiency may not become the future's primary indicator in evaluating the quality of the estimation method, since computing capacity has been exponentially increased over the years. Furthermore, the Gs-MN method also used a non-informative prior of multivariate normal distribution with mean vector 0 and variance 1000 (covariance is 0), in which its standard deviations of the model parameters were larger than Gs-J. This phenomenon may be caused by the variances of model parameters in the priors being set to 1000. When the variances were limited, there was still an upper and lower limit, although the distribution range of the prior's parameters was relatively wide. Thus, this can be categorized as a pseudo-non-informative prior with remarkably insufficient information.

In our study, we determined two types of informative priors according to the samples' sources to further implement the Gibbs sampler-based Bayesian analysis. The first type was derived from the results of our data's repeated self-sampling (Gs-MN1), while the other was acquired from the literature (Gs-MN2). The self-sampling informative prior gave a more accurate value of not only parameters' means and variances, but also fitting and validation performance compared with the one based on the literature's informative prior. These findings confirmed that accurate informative priors would bring a greater improvement in the model performances. Overall, the Gs-MN1 method produced the smallest DIC value, followed sequentially by DMC, Gs-J, Gs-MN, and Gs-MN2. However, our results also suggested that, at least by using the data obtained from this research, the literature's informative priors may not be suitable for biomass modeling in the study area. In future research, further research on the determination of prior information is still needed to improve biomass model performance.

4.3. Anti-Logarithm Correction Factor

Logarithmic transformation in biomass modeling is used to simply linearize the nonlinear relationships between biomass and predictors [59,60]. It is important to make a decision on whether or not a bias correction is needed. Usually, it is determined by the size of the error term [61]. In this study, three evaluated correction factors (CFs) were all small (ranging from 1.0058 to 1.0604). For the statistics of CFs, the B and G were relatively similar, suggesting that the difference between the CFs was not significant. Furthermore, we also compared the mean percentage difference (MPD) of the four types of CFs (CF0, CF1, CF2, and CF3). The results showed that MPD0 (calculated by CF0) was the smallest among others, indicating that the anti-log correction may not be necessary for our biomass data. In general, the CFs' inferences obtained in this study were consistent with previous research [61,62].

5. Conclusions

In this study, we applied Bayesian methods to construct two sets of the seemingly unrelated regression models (i.e., SURM1 and SURM2) with no constraint to take account of the inherent correlations between four tree components (i.e., stem, branch, foliage, and root) collected from the same tree. The classical method FGLS was also used to estimate the model parameters of the SUR models as comparison. The Bayesian methods with the informative priors not only produced better model fitting, but also higher prediction accuracy and computing efficiency. In addition, the Bayesian methods had more stable predictions for small sample sizes, which could be used to estimate new model parameters in new locations, as we suggested. However, finding good informative priors for the model parameters was crucial and necessary to improve the model fitting and prediction accuracy and efficiency. On the other hand, the anti-logarithm correction factors on the model predictions may not be necessary for our developed models, as the model error terms were relatively small.

Overall, applying the Bayesian methods on the SUR models of tree biomass showed good benefits compared to the classical modeling method, especially when the amount of available data was relatively small. These newly-developed SUR models based on Bayesian methods can be used to accurately estimate the tree biomass of planted Korean larch in northeast China and to estimate carbon storage and further understand the energy matter distribution of larch plantations.

Author Contributions: L.D. and F.L. provided the data, conceived the ideas, and designed methodology; L.X. analyzed the data and wrote the manuscript. F.R.A.W. and L.Z. helped in analyzing the data and writing the paper. All authors contributed critically to the data collection and manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the National Natural Science Foundation of China (No. 31971649), the National Key R&D Program of China (No. 2017YFD0600402), Provincial Funding for National Key R&D Program of China in Heilongjiang Province (No. GX18B041), the Fundamental Research Funds for the Central Universities (No. 2572019CP08) and the Heilongjiang Touyan Innovation Team Program (Technology Development Team for High-Efficient Silviculture of Forest Resources).

Acknowledgments: The authors would like to thank the faculty and students of the Department of Forest Management, Northeast Forestry University (NEFU), China, who provided and collected the data for this study. And the authors also thank two anonymous reviewers for suggestions and comments that improved the contents and quality of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Classical Approach of SUR

A m -dimensional linear SUR model system can be expressed as follows:

$$y_i = X_i\beta_i + \varepsilon_i \quad (\text{A1})$$

where $E[\varepsilon_i\varepsilon_j'] = \begin{cases} \omega_{ij}I & i \neq j \\ \omega_i^2I & i = j \end{cases}$, i and j are 1, 2, 3, ..., m . A general expression of linear SUR models (in matrix notation) is

$$Y = X\beta + \varepsilon \quad (\text{A2})$$

FGLS for estimating SUR model system:

Step 1. Fit the m linear functions of the SUR models individually and calculate the error terms ε_i for each function, where $i = 1, 2, \dots, m$;

Step 2. Compute the variance-covariance matrix $\hat{\Omega}$ of ε_i obtained in step 1 by Equation (A3):

$$\hat{\Omega} = \frac{1}{n} \sum \varepsilon_i\varepsilon_j' \quad (\text{A3})$$

Step 3. Estimate the model coefficients by GLS as follows:

$$\hat{\beta} = (X^T(\hat{\Omega}^{-1} \otimes I)X)^{-1}X^T(\hat{\Omega}^{-1} \otimes I)Y \quad (\text{A4})$$

where upper right T is a transpose of the matrix, and Ω is the variance–covariance matrix.

Bayesian approaches of SUR

Based on Equation (A2), the likelihood function by giving the data D is

$$L(D|\beta, \Omega) = \frac{1}{(2\pi)^{nm/2}|\Omega|^{n/2}} \exp\left[-\frac{1}{2}(Y - X\beta)^T(\Omega^{-1} \otimes I)(Y - X\beta)\right] \quad (\text{A5})$$

Gibbs Sampler on SUR models

1. A simple prior of β and Ω was expressed as follows:

$$\begin{cases} \pi(\beta, \Omega|D) = \pi(\beta)\pi(\Omega) \\ \beta \sim N(\bar{\beta}, A^{-1}) \\ \Omega \sim InvWishart(v_0, V_0) \end{cases} \quad (\text{A6})$$

where $N(\bar{\beta}, A^{-1})$ denotes the multivariate normal distribution, where $\bar{\beta}$ is the prior mean vector of $\sum p_i \times 1$, and p_i is the number of parameters in the i_{th} single model of Equation (A1); A is the prior precision matrix of $\sum p_j \times \sum p_j$; and $InvWishart(v_0, V_0)$ denotes the inverse Wishart distribution, where both v_0 and V_0 are the degree of freedom and $m \times m$ prior precision matrix for the inverse Wishart prior. The normal prior of β is conjugated with the conditional likelihood of Equation (A5), and the posterior of β by giving Ω is normal:

$$\pi(\beta|\Omega, D) = N(\hat{\beta}, \hat{\Omega}_\beta) \quad (\text{A7})$$

where $\hat{\beta} = (X^T(\Omega^{-1} \otimes I)X + A)^{-1}(X^T(\Omega^{-1} \otimes I)Y + A\bar{\beta})$ and $\hat{\Omega}_\beta = (X^T(\Omega^{-1} \otimes I)X + A)^{-1}$. Meanwhile, the inverted Wishart prior is a conditional conjugate and is used to obtain the posterior distribution with the following prior:

$$\pi(\Omega|\beta, D) = InvWishart(v_0 + n, V_0 + R) \quad (\text{A8})$$

where R is the $m \times m$ matrix of (r_{ij}) , and $r_{ij} = (y_i - X_i\beta_i)^T(y_j - X_j\beta_j)$ in Equation (A1).

2. A non-informative prior that is widely used, namely Jeffreys invariant prior is expressed as follows:

$$\pi(\beta, \Omega|D) = (\beta)\pi(\Omega) \propto |\Omega|^{-\frac{m+1}{2}}. \quad (\text{A9})$$

The conditional posteriors are [39]

$$\pi(\beta|\Omega, D) = N(\hat{\beta}, \hat{\Omega}_\beta) \quad (\text{A10})$$

$$\pi(\Omega|\beta, D) = InvWishart(R, n) \quad (\text{A11})$$

where $\hat{\beta} = (X^T(\Omega^{-1} \otimes I)X)^{-1}X^T(\Omega^{-1} \otimes I)Y$, and $\hat{\Omega}_\beta = (X^T(\Omega^{-1} \otimes I)X)^{-1}$.

According to the conditional posteriors on Equations (A7), (A8), (A10), and (A11), a Markov Chain Monte Carlo (MCMC) method with a Gibbs sampler can be used to obtain posterior samples of the SUR model parameters. The steps of the Gibbs sampler on SUR models are summarized as follows:

- Step 1. Give the starting values of Ω , namely Ω_0 , the starting values were obtained from prior distributions generally;

- Step 2. Draw β_1 from Equation (A7) or (A10);
 Step 3. Draw Ω_1 from Equation (A8) or (A11);
 Step 4. Repeat both step 2 and step 3 to draw β_j and Ω_j for N times, $j = 1, 2, 3, \dots, N$,

where N is the sample size that could be manually set.

DMC for estimating SUR model system

The SUR models (Equation (A1)) are reformulated as follows:

$$\begin{cases} Y_1 = X_1\beta_1 + e_1 = Z_1b_1 + e_1 \\ Y_2 = X_2\beta_2 + \rho_{21}e_1 + e_2 = Z_2b_2 + e_2 \\ \dots \\ Y_m = X_m\beta_m + \sum_{j=1}^{m-1} \rho_{mj}e_j + e_m = Z_mb_m + e_m \end{cases} \quad (\text{A12})$$

with $e_1 = \varepsilon_1$, $e_i \sim N(0, \Sigma)$, and $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2)$. There is no correlation detected between the error terms e_i and e_j of Equation (A12). The Z_i is the design matrix that including predictors, and b_i is the vector of model parameters, where $i=1, 2, 3$ and 4. Thus, the likelihood function of the data D becomes

$$L(D|b, \Sigma) = \prod_{j=1}^m \frac{1}{(2\pi\sigma_j^2)^{n/2}} \exp \left[-\frac{(Y_j - Z_jb_j)^T (Y_j - Z_jb_j)}{2\sigma_j^2} \right]. \quad (\text{A13})$$

Herein, we set the Jeffreys invariant prior as the prior:

$$\pi(b, \Sigma|D) = \pi(b)\pi(\Sigma) \propto \prod_{j=1}^m (\sigma_j)^{-1} \quad (\text{A14})$$

then the conditional normal inverse-gamma posterior is

$$\pi(b_j | b_{j-1}, \dots, b_1, \sigma_j^2, D) = N((Z_j^T Z_j)^{-1} Z_j^T Y_j, \sigma_j^2 (Z_j^T Z_j)^{-1}) \quad (\text{A15})$$

$$\pi(\sigma_j^2 | b_{j-1}, \dots, b_1, D) = IG(\frac{1}{2}(Y_j - Z_jb_j)^T (Y_j - Z_jb_j), \frac{1}{2}(n - p_j - j + 1)) \quad (\text{A16})$$

The sampling procedure of m -dimensional SUR models can be determined using the following steps:

- Step 1. Set $j = 1$ to generate $\sigma_1^{2(k)}$ ($k = 1, 2, 3, \dots, N$), where N is the samples size. Draw $b_1^{(k)}$ from the Equation (A15).
 Step 2. $j = j + 1$ is set as the increment, and then $\sigma_j^{2(k)}$ is drawn by using Equation (A16) to generate the $b_j^{(k)}$ in Equation (A15).
 Step 3. Repeat the step 2 sequentially until $j = m$ is reached.
 Step 4. Transform $\{b, \Sigma\}$ into to $\{\beta, \Omega\}$.

The more detailed information of DMC approach could be found in the Zellner's inference [39].

Predicting with Bayesian method

Given the posterior samples of model parameters, the predictive density of model's dependent variables can be approximated as follows:

$$\int f(y_{new}|x_{new}, \beta, \Sigma) g(\beta, \Sigma|D) d\beta d\Sigma \approx \frac{1}{n} \sum_{k=1}^n f(y_{new}|x_{new}, \beta^k, \Sigma^k) \quad (\text{A17})$$

Appendix B

Figure A1. Simulation results for the parameter estimations of SURM2 using the DMC, Gs-MN, Gs-MN1, Gs-MN2, and FGLS. The dashed and solid horizontal lines represent the 95% confidence intervals and means of the simulation using the entire dataset (174 trees), respectively.

A

B

Figure A2. The MAE (A) and MAE% (B) of the SURM2 using different sample sizes. The solid line represents the value of MAE and MAE% calculated by all dataset (174 trees).

Table A1. The correlation matrices of SURM1 and SURM2 fitted by DMC method.

Model		Correlation Matrix			
SURM1	W_s	1.0000	-0.4292	-0.2428	0.4487
	W_b	-0.4292	1.0000	0.5836	-0.2208
	W_f	-0.2428	0.5836	1.0000	-0.1062
	W_r	0.4487	-0.2208	-0.1062	1.0000
SUMR2	W_s	1.0000	0.0026	0.0084	0.2311
	W_b	0.0026	1.0000	0.5342	-0.0362
	W_f	0.0084	0.5342	1.0000	0.0042
	W_r	0.2311	-0.0362	0.0042	1.0000

Table A2. The average of the mean absolute error ($\overline{\text{MAE}}$, kg) and average of the mean absolute percent error ($\overline{\text{MAE\%}}$) of the six modeling methods with different sample sizes in the simulations for the four tree biomass components.

Methods	Components	Indices	Sample Sizes									
			10	15	20	30	40	60	80	110	140	170
SUM1												
DMC	W_r	$\overline{\text{MAE}}$	4.89	4.62	4.50	4.42	4.38	4.37	4.37	4.36	4.36	4.37
		$\overline{\text{MAE\%}}$	22.08	21.23	20.86	20.50	20.32	20.16	20.07	20.01	19.98	19.97
	W_s	$\overline{\text{MAE}}$	19.36	18.48	18.14	17.80	17.63	17.47	17.39	17.31	17.28	17.27
		$\overline{\text{MAE\%}}$	21.16	20.41	20.09	19.79	19.64	19.50	19.43	19.38	19.36	19.34
	W_b	$\overline{\text{MAE}}$	2.80	2.66	2.59	2.53	2.49	2.46	2.45	2.43	2.43	2.42
		$\overline{\text{MAE\%}}$	31.94	30.16	29.35	28.59	28.25	27.90	27.74	27.60	27.51	27.47
	W_f	$\overline{\text{MAE}}$	0.81	0.77	0.76	0.75	0.74	0.73	0.73	0.73	0.72	0.72
		$\overline{\text{MAE\%}}$	26.62	25.12	24.50	23.91	23.64	23.35	23.23	23.13	23.06	23.02
Gs-MN	W_r	$\overline{\text{MAE}}$	4.90	4.63	4.51	4.42	4.39	4.37	4.37	4.36	4.36	4.36
		$\overline{\text{MAE\%}}$	22.09	21.24	20.87	20.50	20.33	20.15	20.08	20.02	19.98	19.97
	W_s	$\overline{\text{MAE}}$	19.33	18.47	18.14	17.80	17.63	17.47	17.39	17.32	17.28	17.27
		$\overline{\text{MAE\%}}$	21.14	20.41	20.09	19.79	19.64	19.50	19.43	19.38	19.35	19.34
	W_b	$\overline{\text{MAE}}$	2.80	2.66	2.59	2.53	2.49	2.46	2.45	2.43	2.43	2.42
		$\overline{\text{MAE\%}}$	31.97	30.18	29.36	28.59	28.25	27.89	27.73	27.59	27.52	27.47
	W_f	$\overline{\text{MAE}}$	0.81	0.77	0.76	0.75	0.74	0.73	0.73	0.73	0.72	0.72
		$\overline{\text{MAE\%}}$	26.63	25.13	24.51	23.92	23.64	23.37	23.24	23.13	23.06	23.02
Gs-MN1	W_r	$\overline{\text{MAE}}$	4.30	4.31	4.32	4.33	4.33	4.34	4.35	4.35	4.35	4.36
		$\overline{\text{MAE\%}}$	19.89	19.92	19.95	19.97	19.98	19.97	19.96	19.95	19.95	19.95
	W_s	$\overline{\text{MAE}}$	17.28	17.28	17.28	17.29	17.29	17.28	17.28	17.27	17.27	17.26
		$\overline{\text{MAE\%}}$	19.37	19.37	19.38	19.38	19.37	19.36	19.36	19.35	19.34	19.34
	W_b	$\overline{\text{MAE}}$	2.43	2.43	2.43	2.43	2.43	2.43	2.43	2.42	2.42	2.42
		$\overline{\text{MAE\%}}$	27.65	27.65	27.65	27.62	27.61	27.59	27.57	27.53	27.49	27.47
	W_f	$\overline{\text{MAE}}$	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72
		$\overline{\text{MAE\%}}$	23.17	23.17	23.17	23.15	23.14	23.12	23.10	23.07	23.04	23.02
Gs-MN2	W_r	$\overline{\text{MAE}}$	6.28	5.73	5.37	4.98	4.78	4.60	4.53	4.47	4.44	4.43
		$\overline{\text{MAE\%}}$	22.72	22.10	21.66	21.10	20.79	20.46	20.29	20.17	20.10	20.07
	W_s	$\overline{\text{MAE}}$	19.76	18.91	18.45	17.99	17.78	17.57	17.47	17.37	17.33	17.30
		$\overline{\text{MAE\%}}$	19.54	19.51	19.51	19.49	19.48	19.42	19.39	19.36	19.34	19.33
	W_b	$\overline{\text{MAE}}$	3.60	3.19	2.96	2.73	2.62	2.53	2.49	2.46	2.45	2.44
		$\overline{\text{MAE\%}}$	34.22	31.79	30.50	29.18	28.58	28.05	27.82	27.64	27.53	27.48
	W_f	$\overline{\text{MAE}}$	1.08	0.96	0.89	0.81	0.78	0.75	0.74	0.73	0.73	0.73
		$\overline{\text{MAE\%}}$	29.76	27.16	25.62	24.23	23.69	23.32	23.19	23.10	23.03	22.99
FGLS	W_r	$\overline{\text{MAE}}$	4.89	4.62	4.50	4.42	4.38	4.36	4.37	4.36	4.36	4.36
		$\overline{\text{MAE\%}}$	22.08	21.23	20.86	20.50	20.32	20.15	20.08	20.02	19.98	19.97
	W_s	$\overline{\text{MAE}}$	19.36	18.48	18.14	17.80	17.63	17.47	17.39	17.32	17.28	17.27
		$\overline{\text{MAE\%}}$	21.16	20.41	20.09	19.79	19.64	19.50	19.43	19.38	19.35	19.34
	W_b	$\overline{\text{MAE}}$	2.80	2.66	2.59	2.53	2.49	2.46	2.45	2.43	2.43	2.42
		$\overline{\text{MAE\%}}$	31.94	30.16	29.35	28.59	28.25	27.89	27.73	27.59	27.52	27.47
	W_f	$\overline{\text{MAE}}$	0.81	0.77	0.76	0.75	0.74	0.73	0.73	0.73	0.72	0.72
		$\overline{\text{MAE\%}}$	26.61	25.12	24.50	23.91	23.64	23.37	23.24	23.13	23.06	23.02
SURM2												
DMC	W_r	$\overline{\text{MAE}}$	5.18	4.76	4.61	4.46	4.40	4.34	4.31	4.28	4.27	4.27
		$\overline{\text{MAE\%}}$	21.67	20.14	19.55	18.97	18.71	18.41	18.26	18.15	18.09	18.04
	W_s	$\overline{\text{MAE}}$	9.49	8.88	8.60	8.33	8.20	8.08	8.02	7.96	7.93	7.91
		$\overline{\text{MAE\%}}$	10.05	9.39	9.10	8.83	8.70	8.57	8.50	8.45	8.41	8.38
	W_b	$\overline{\text{MAE}}$	2.72	2.53	2.45	2.37	2.33	2.29	2.27	2.26	2.26	2.25
		$\overline{\text{MAE\%}}$	29.39	27.21	26.34	25.38	24.97	24.56	24.38	24.23	24.11	24.05
	W_f	$\overline{\text{MAE}}$	0.86	0.80	0.78	0.76	0.75	0.74	0.74	0.73	0.73	0.73
		$\overline{\text{MAE\%}}$	27.61	25.49	24.63	23.80	23.47	23.15	23.02	22.94	22.88	22.86

Table A2. Cont.

Methods	Components	Indices	Sample Sizes									
			10	15	20	30	40	60	80	110	140	170
Gs-MN	W_r	$\overline{\text{MAE}}$	5.16	4.76	4.61	4.46	4.40	4.34	4.31	4.28	4.27	4.27
		$\overline{\text{MAE}\%}$	21.62	20.13	19.55	18.97	18.71	18.41	18.26	18.15	18.09	18.04
	W_s	$\overline{\text{MAE}}$	9.49	8.88	8.60	8.33	8.20	8.08	8.02	7.96	7.93	7.91
		$\overline{\text{MAE}\%}$	10.05	9.39	9.11	8.83	8.70	8.57	8.51	8.45	8.41	8.38
	W_b	$\overline{\text{MAE}}$	2.70	2.53	2.45	2.37	2.33	2.29	2.27	2.26	2.26	2.25
		$\overline{\text{MAE}\%}$	29.34	27.21	26.34	25.38	24.97	24.56	24.38	24.23	24.11	24.05
	W_f	$\overline{\text{MAE}}$	0.85	0.80	0.78	0.76	0.75	0.74	0.74	0.73	0.73	0.73
		$\overline{\text{MAE}\%}$	27.57	25.49	24.63	23.80	23.47	23.15	23.02	22.94	22.88	22.86
	W_r	$\overline{\text{MAE}}$	4.24	4.25	4.25	4.26	4.26	4.26	4.26	4.26	4.26	4.26
		$\overline{\text{MAE}\%}$	18.07	18.09	18.11	18.12	18.13	18.11	18.09	18.07	18.05	18.02
	W_s	$\overline{\text{MAE}}$	7.96	7.96	7.96	7.96	7.96	7.95	7.94	7.93	7.92	7.91
		$\overline{\text{MAE}\%}$	8.44	8.45	8.45	8.45	8.45	8.44	8.43	8.41	8.40	8.38
Gs-MN1	W_b	$\overline{\text{MAE}}$	2.26	2.26	2.26	2.26	2.26	2.26	2.26	2.26	2.25	2.25
		$\overline{\text{MAE}\%}$	24.30	24.30	24.31	24.28	24.26	24.22	24.19	24.14	24.08	24.05
	W_f	$\overline{\text{MAE}}$	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73
		$\overline{\text{MAE}\%}$	22.98	22.98	22.98	22.97	22.96	22.94	22.92	22.90	22.88	22.87
	W_r	$\overline{\text{MAE}}$	6.45	6.04	5.67	5.18	4.92	4.67	4.55	4.47	4.41	4.38
		$\overline{\text{MAE}\%}$	24.50	22.72	21.52	20.10	19.41	18.82	18.57	18.39	18.29	18.21
	W_s	$\overline{\text{MAE}}$	13.68	12.59	11.71	10.59	10.02	9.48	9.17	8.85	8.62	8.46
		$\overline{\text{MAE}\%}$	14.08	12.97	12.14	11.15	10.69	10.25	9.98	9.65	9.38	9.17
	W_b	$\overline{\text{MAE}}$	3.47	3.30	3.19	3.05	2.95	2.78	2.65	2.49	2.38	2.32
		$\overline{\text{MAE}\%}$	36.87	35.73	35.04	33.87	32.94	31.31	29.94	28.24	26.97	26.11
	W_f	$\overline{\text{MAE}}$	1.08	0.97	0.91	0.85	0.82	0.78	0.76	0.75	0.73	0.72
		$\overline{\text{MAE}\%}$	30.93	28.73	27.60	26.49	25.86	25.04	24.47	23.84	23.40	23.10
FGLS	W_r	$\overline{\text{MAE}}$	5.18	4.76	4.61	4.46	4.40	4.34	4.31	4.28	4.27	4.27
		$\overline{\text{MAE}\%}$	21.67	20.14	19.55	18.97	18.71	18.41	18.26	18.15	18.09	18.04
	W_s	$\overline{\text{MAE}}$	9.49	8.88	8.60	8.33	8.20	8.08	8.02	7.96	7.93	7.91
		$\overline{\text{MAE}\%}$	10.05	9.39	9.10	8.83	8.70	8.57	8.50	8.45	8.41	8.38
	W_b	$\overline{\text{MAE}}$	2.72	2.53	2.45	2.37	2.33	2.29	2.27	2.26	2.26	2.25
		$\overline{\text{MAE}\%}$	29.39	27.21	26.34	25.38	24.97	24.56	24.38	24.23	24.11	24.05
	W_f	$\overline{\text{MAE}}$	0.86	0.80	0.78	0.76	0.75	0.74	0.74	0.73	0.73	0.73
		$\overline{\text{MAE}\%}$	27.61	25.49	24.63	23.80	23.47	23.15	23.02	22.94	22.88	22.86

Table A3. The autocorrelation function of the posterior samples for the parameter β_{r0} of the Bayesian methods calculated using 10,000 samples. The max lag was set to 50.

Lag	DMC	Gs-J	Gs-MN	Gs-MN1	Gs-MN2
SURM1					
0	1.0000	1.0000	1.0000	1.0000	1.0000
1	−0.0005	0.0120	−0.0060	0.0033	0.0062
5	−0.0108	0.0059	0.0296	0.0078	0.0002
10	0.0149	−0.0011	−0.0043	−0.0180	−0.0011
50	−0.0020	−0.0011	0.0075	0.0015	0.0160
SURM2					
0	1.0000	1.0000	1.0000	1.0000	1.0000
1	0.0022	−0.0039	0.0051	0.0007	−0.0040
5	−0.0104	0.0001	0.0037	−0.0048	−0.0181
10	0.0033	−0.0087	−0.0048	0.0021	−0.0057
50	−0.0018	0.0014	0.0043	−0.0108	−0.0129

where the subscript *new* means the new data that need to be predicted, and x_{new} is the matrix of new independent variables; β and Σ are the parameters and variance–covariance matrix of errors, respectively; and n is the size of the posterior samples.

References

- Clark, J.; Murphy, G. Estimating forest biomass components with hemispherical photography for Douglas-fir stands in northwest Oregon. *Can. J. For. Res.* **2011**, *41*, 1060–1074. [\[CrossRef\]](#)
- Dong, L.; Zhang, L.; Li, F. A compatible system of biomass equations for three conifer species in Northeast, China. *For. Ecol. Manag.* **2014**, *329*, 306–317. [\[CrossRef\]](#)
- Weiskittel, A.R.; MacFarlane, D.W.; Radtke, P.J.; Affleck, D.L.; Hailemariam, T.; Woodall, C.W.; Westfall, J.A.; Coulston, J.W. A call to improve methods for estimating tree biomass for regional and national assessments. *J. For.* **2015**, *113*, 414–424. [\[CrossRef\]](#)
- Zhao, D.; Kane, M.; Markewitz, D.; Teskey, R.; Clutter, M. Additive tree biomass equations for midrotation Loblolly pine plantations. *For. Sci.* **2015**, *61*, 613–623. [\[CrossRef\]](#)
- Ter-Mikaelian, M.T.; Korzukhin, M.D. Biomass equations for sixty-five north American tree species. *For. Ecol. Manag.* **1997**, *97*, 1–24. [\[CrossRef\]](#)
- Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. National-scale biomass estimators for united states tree species. *For. Sci.* **2003**, *49*, 12–35.
- Zianis, D.; Muukkonen, P.; Makipaa, R.; Mencuccini, M. Biomass and stem volume equations for tree species in Europe. *Silva Fenn. Monogr.* **2005**, *4*, 63.
- Wang, C. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. *For. Ecol. Manag.* **2006**, *222*, 9–16. [\[CrossRef\]](#)
- Chave, J.; Rejoumechain, M.; Burquez, A.; Chidumayo, E.N.; Colgan, M.S.; Delitti, W.B.C.; Duque, A.; Eid, T.; Fearnside, P.M.; Goodman, R.C. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **2014**, *20*, 3177–3190. [\[CrossRef\]](#)
- Luo, Y.; Wang, X.; Ouyang, Z.; Lu, F.; Feng, L.; Tao, J. A review of biomass equations for China's tree species. *Earth Syst. Sci. Data* **2020**, *12*, 21–40. [\[CrossRef\]](#)
- Bi, H.; Murphy, S.; Volkova, L.; Weston, C.J.; Fairman, T.A.; Li, Y.; Law, R.; Norris, J.; Lei, X.; Caccamo, G. Additive biomass equations based on complete weighing of sample trees for open eucalypt forest species in south-eastern Australia. *For. Ecol. Manag.* **2015**, *349*, 106–121. [\[CrossRef\]](#)
- Kralicek, K.; Huy, B.; Poudel, K.P.; Temesgen, H.; Salas, C. Simultaneous estimation of above- and below-ground biomass in tropical forests of Viet Nam. *For. Ecol. Manag.* **2017**, *390*, 147–156. [\[CrossRef\]](#)
- Zapatacuartas, M.; Sierra, C.A.; Alleman, L. Probability distribution of allometric coefficients and Bayesian estimation of aboveground tree biomass. *For. Ecol. Manag.* **2012**, *277*, 173–179. [\[CrossRef\]](#)
- Dong, L.; Zhang, Y.; Xie, L.; Li, F. Comparison of tree biomass modeling approaches for larch (*Larix olgensis* Henry) trees in Northeast China. *Forests* **2020**, *11*, 202. [\[CrossRef\]](#)
- Parresol, B.R. Assessing tree and stand biomass: A review with examples and critical comparisons. *For. Sci.* **1999**, *45*, 573–593.
- Zellner, A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J. Am. Stat. Assoc.* **1962**, *57*, 348–368. [\[CrossRef\]](#)
- Mehtatalo, L.; Lappi, J. *Biometry for Forestry and Environmental Data: With Examples in R*; Chapman and Hall; CRC Press: New York, NY, USA, 2020; p. 426.
- Dong, L.; Zhang, L.; Li, F. Additive biomass equations based on different dendrometric variables for two dominant species (*Larix gmelini* Rupr. and *Betula platyphylla* Suk.) in natural forests in the Eastern Daxing'an Mountains, Northeast China. *Forests* **2018**, *9*, 261. [\[CrossRef\]](#)
- Parresol, B.R. Additivity of nonlinear biomass equations. *Can. J. For. Res.* **2001**, *31*, 865–878. [\[CrossRef\]](#)
- Zellner, A. *An Introduction to Bayesian Inference in Econometrics*; Wiley: New York, NY, USA, 1971; p. 448.
- Lu, L.; Wang, H.; Chhin, S.; Duan, A.; Zhang, J.; Zhang, X. A Bayesian model averaging approach for modelling tree mortality in relation to site, competition and climatic factors for Chinese fir plantations. *For. Ecol. Manag.* **2019**, *440*, 169–177. [\[CrossRef\]](#)
- Griffiths, W.E. *Bayesian Inference in the Seemingly Unrelated Regressions Model*; Department of Economics, The University of Melbourne: Melbourne, Australia, 2003; p. 520.
- Bayes, T. An essay towards solving a problem in the doctrine of chances. *M.D. Comput.* **1991**, *8*, 157. [\[CrossRef\]](#)
- Li, R.; Stewart, B.; Weiskittel, A.R. A Bayesian approach for modelling non-linear longitudinal/hierarchical data with random effects in forestry. *Forestry* **2012**, *85*, 17–25. [\[CrossRef\]](#)

25. Rossi, P.E.; Allenby, G.M. *Bayesian Statistics and Marketing*; John Wiley & Sons, Ltd.: Chichester, UK, 2005; p. 368.
26. Huelsenbeck, J.P.; Ronquist, F.; Nielsen, R.; Bollback, J.P. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **2001**, *294*, 2310–2314. [[CrossRef](#)] [[PubMed](#)]
27. Portinale, L.; Raiteri, D.C.; Montani, S. Supporting reliability engineers in exploiting the power of dynamic Bayesian networks. *Int. J. Approx. Reason.* **2010**, *51*, 179–195. [[CrossRef](#)]
28. Reich, B.J.; Ghosh, S.K. *Bayesian Statistical Methods*; CRC Press: Boca Raton, FL, USA, 2019; p. 275.
29. Berger, J.O. Bayesian analysis: A look at today and thoughts of tomorrow. *J. Am. Stat. Assoc.* **2000**, *95*, 1269–1276. [[CrossRef](#)]
30. Samuel, K.; Wu, X. *Modern Bayesian Statistics*; China Statistics Press: Beijing, China, 2000; p. 244.
31. Van Ravenzwaaij, D.; Cassey, P.; Brown, S.D. A simple introduction to Markov Chain Monte-Carlo sampling. *Psychon. Bull. Rev.* **2018**, *25*, 143–154. [[CrossRef](#)]
32. Zhang, X.; Duan, A.; Zhang, J. Tree biomass estimation of Chinese fir (*Cunninghamia lanceolata*) based on Bayesian method. *PLoS ONE* **2013**, *8*, e79868. [[CrossRef](#)]
33. *State Forestry and Grassland Administration The Ninth Forest Resource Survey Report (2014–2018)*; China Forestry Press: Beijing, China, 2018; p. 451.
34. Zeng, W.; Duo, H.; Lei, X.; Chen, X.; Wang, X.; Pu, Y.; Zou, W. Individual tree biomass equations and growth models sensitive to climate variables for *Larix* spp. in China. *Eur. J. For. Res.* **2017**, *136*, 233–249. [[CrossRef](#)]
35. Xiao, X.; White, E.P.; Hooten, M.B.; Durham, S.L. On the use of log-transformation vs. nonlinear regression for analyzing biological power laws. *Ecology* **2011**, *92*, 1887–1894. [[CrossRef](#)]
36. SAS Institute Inc. *SAS/ETS®14.1 User's Guide*; SAS Institute Inc.: Cary, NC, USA, 2015; p. 4100.
37. Henningsen, A.; Hamann, J.D. Systemfit: A package for estimating systems of simultaneous equations in R. *J. Stat. Softw.* **2007**, *23*, 1–40. [[CrossRef](#)]
38. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
39. Zellner, A.; Ando, T. A direct Monte Carlo approach for Bayesian analysis of the seemingly unrelated regression model. *J. Econom.* **2010**, *159*, 33–45. [[CrossRef](#)]
40. Rossi, P. Bayesm: Bayesian Inference for Marketing/Micro-Econometrics; R package version 3.1-4. 2019. Available online: <https://CRAN.R-project.org/package=bayesm> (accessed on 20 April 2020).
41. Heidelberger, P.; Welch, P.D. Simulation run length control in the presence of an initial transient. *Oper. Res.* **1983**, *31*, 1109–1144. [[CrossRef](#)]
42. Gewke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4*; Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M., Eds.; Clarendon Press: Oxford, UK, 1992; pp. 169–193.
43. Plummer, M.; Best, N.; Cowles, K.; Vines, K. Coda: Convergence diagnosis and output analysis for MCMC. *R News* **2006**, *6*, 7–11.
44. Fu, L.; Sharma, R.P.; Hao, K.; Tang, S. A generalized interregional nonlinear mixed-effects crown width model for prince rupperecht larch in northern China. *For. Ecol. Manag.* **2017**, *389*, 364–373. [[CrossRef](#)]
45. Cawley, G.C.; Talbot, N.L.C. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recognit.* **2003**, *36*, 2585–2592. [[CrossRef](#)]
46. Finney, D.J. On the distribution of a variate whose logarithm is normally distributed. *J. R. Statist. Soc. B* **1941**, *7*, 155–161. [[CrossRef](#)]
47. Baskerville, G.L. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* **1972**, *4*, 149. [[CrossRef](#)]
48. Wiant, H.V.; Harner, E.J. Percent bias and standard error in logarithmic regression. *For. Sci.* **1979**, *25*, 167–168.
49. Yandle, D.O.; Wiant, H.V. Estimation of plant biomass based on the allometric equation. *Can. J. For. Res.* **1981**, *11*, 833–834. [[CrossRef](#)]
50. Bond-Lamberty, B.; Wang, C.; Gower, S.T. Net primary production and net ecosystem production of a boreal black spruce wildfire chronosequence. *Glob. Chang. Biol.* **2004**, *10*, 473–487. [[CrossRef](#)]
51. Pregitzer, K.S.; Euskirchen, E.S. Carbon cycling and storage in world forests: Biome patterns related to forest age. *Glob. Chang. Biol.* **2004**, *10*, 2052–2077. [[CrossRef](#)]
52. Wang, X.; Zhao, D.; Liu, G.; Yang, C.; Teskey, R.O. Additive tree biomass equations for *Betula platyphylla* Suk. plantations in Northeast China. *Ann. For. Sci.* **2018**, *75*, 60. [[CrossRef](#)]

53. Zhao, D.; Westfall, J.A.; Coulston, J.W.; Lynch, T.B.; Bullock, B.P.; Montes, C.R. Additive biomass equations for slash pine trees: Comparing three modeling approaches. *Can. J. For. Res.* **2019**, *49*, 27–40. [\[CrossRef\]](#)
54. Widagdo, F.R.A.; Li, F.; Zhang, L.; Dong, L. Aggregated biomass model systems and carbon concentration variations for tree carbon quantification of natural *Mongolian Oak* in Northeast China. *Forests* **2020**, *11*, 397. [\[CrossRef\]](#)
55. Affleck, D.L.R.; Dieguez-Aranda, U. Additive nonlinear biomass equations: A likelihood-based approach. *For. Sci.* **2016**, *62*, 129–140. [\[CrossRef\]](#)
56. Dong, L.; Zhang, L.; Li, F. Developing additive systems of biomass equations for nine hardwood species in Northeast China. *Trees-Struct. Funct.* **2015**, *29*, 1149–1163. [\[CrossRef\]](#)
57. Kusmana, C.; Hidayat, T.; Tiryana, T.; Rusdiana, O. Istomo Allometric models for above- and below-ground biomass of *Sonneratia* spp. *Glob. Ecol. Conserv.* **2018**, *15*, 10.
58. Ando, T. Bayesian variable selection for the seemingly unrelated regression models with a large number of predictors. *J. Jpn. Stat. Soc.* **2012**, *41*, 187–203. [\[CrossRef\]](#)
59. Tang, S. Bias correction in logarithmic regression and comparison with weighted regression for non-linear models. *For. Res.* **2011**, *24*, 137–143.
60. Mascaro, J.; Litton, C.M.; Hughes, R.F.; Uowolo, A.; Schnitzer, S.A. Is logarithmic transformation necessary in allometry? Ten, one-hundred, one-thousand-times yes. *Biol. J. Linn. Soc.* **2014**, *111*, 230–233. [\[CrossRef\]](#)
61. Madgwick, H.A.I.; Satoo, T. On estimating the aboveground weights of tree stands. *Ecology* **1975**, *56*, 1446–1450. [\[CrossRef\]](#)
62. Zianis, D.; Xanthopoulos, G.; Kalabokidis, K.; Kazakis, G.; Ghosn, D.; Roussou, O. Allometric equations for aboveground biomass estimation by size class for *Pinus brutia* Ten. trees growing in North and South Aegean Islands, Greece. *Eur. J. For. Res.* **2011**, *130*, 145–160. [\[CrossRef\]](#)

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).