Journal of Applied Statistics

ISSN: 0266-4763 (Print) 1360-0532 (Online) Journal homepage: www.tandfonline.com/journals/cjas20

A Bayesian growth and yield model for slash pine plantations

Edwin J. Green & William E. Strawderman

To cite this article: Edwin J. Green & William E. Strawderman (1996) A Bayesian growth and yield model for slash pine plantations, Journal of Applied Statistics, 23:2-3, 285-300, DOI: 10.1080/02664769624251

To link to this article: https://doi.org/10.1080/02664769624251

Published online: 02 Aug 2010.

Journal of Applied Statistics, Vol. 23, Nos 2&3, 1996, 285–299

A Bayesian growth and yield model for slash pine plantations

EDWIN J. GREEN & WILLIAM E. STRAWDERMAN, *Rutgers University, New Brunswick, USA*

SUMMARY *We formulate a traditional growth and yield model as a Bayes model. We attempt to introduce as few new assumptions as possible. Zellner's Bayesian method of moments procedure is used, because the published model did not include any distributional assumptions. We generate predictive posterior samples for a number of stand variables using the Gibbs sampler. The means of the samples compare favorably with the predictions from the published model. In addition, our model delivers distributions of outcomes, from which it is easy to establish measures of uncertainty, such as highest posterior density regions.*

1 Introduction

In this paper, we detail the results of a study undertaken to produce a Bayesian version of an existing tree plantation growth and yield model. Our motivation arises from the fact that most, if not all, published growth and yield models deliver only point estimates for a given set of input parameters, i.e. no statement of expected variability is provided. This is correctly viewed as a major limitation of such models. In contrast, a Bayesian model would deliver posterior densities that could then be summarized by any statistic(s) that the user desired.

We chose to develop a Bayesian version of the model published by Zarnoch *et al.* (1991). Their model describes the growth and yield of slash pine plantations in the West Gulf region of the US. Our goal is to put the model in a Bayesian framework and produce marginal posterior distributions of the quantities of interest—not to evaluate the suitability of the original model. We will clearly identify any area where we deviate from the published model. We want to produce

Correspondence: E. J. Green, Department of Natural Resources, Cook College, Rutgers University, PO Box 231, New Brunswick, NJ 08903-0231, USA.

0266-4763/96/020285-15 $6.00

a Bayesian version of the model of Zarnoch *et al.*, while introducing as few new assumptions as possible.

2 Description of original model

The Zarnoch *et al.* model (hereafter simply referred to as the 'published' model) is a probability density function (pdf)-based model. It is based on predicting the distribution of tree diameters in a plantation with specified attributes. The relative frequency function of the diameters is assumed to be adequately characterized by a Weibull density. This model is one of a class of models commonly referred to as 'parameter-recovery' models, because the parameters of the pdf are obtained by predicting functions of diameter, and then solving for the parameters.

Simultaneously, one predicts the number of trees surviving per acre (N) in the plantation. Integration of the Weibull density over diameter classes of interest yields the percentage of trees in those classes. Multiplication by N yields an estimate of the number of trees in those classes. A height model is supplied that predicts the height of a tree, based on its diameter and associated plantation attributes. Finally, a volume equation is given that predicts tree volume as a function of tree diameter and tree height.

For each diameter class of interest, one may predict the height and volume of a tree with, say, the midpoint diameter of the class. Multiplication of the volume of the tree with the midpoint diameter by the number of trees in the class yields the volume in the class. Summing over classes of interest then yields the total volume in the desired interval.

2.1 Equations

Below, we present the equations given in the original model. Note that the equations are of two types: plantation or stand level, and individual tree level.

2.1.1 Plantation-level models

$$B = \beta_{11} H_{\mathrm{D}}^{\beta_{12}} N^{\beta_{13}} \exp(\beta_{14} A^{-1}) \tag{1}$$

$$D_{0.93} = \beta_{21} H_{\mathrm{D}}^{\beta_{22}} N^{\beta_{23}} \exp(\beta_{24} A^{-1}) \tag{2}$$

$$D_{\min} = \beta_{31} H_{\mathrm{D}}^{\beta_{32}} N^{\beta_{33}} \exp(\beta_{34} A^{-1}) \tag{3}$$

$$N = T^{1-\beta_{41} A} \tag{4}$$

$$H_{\mathrm{D}} = S \exp[\beta_{52}(A^{-0.5} - 0.2)] \tag{5}$$

where B is the basal area (ft^2), H_{D} is the height of dominant and/or co-dominant trees in the plantation (ft), $D_{\min}$ is the minimum diameter in the plantation (in), $D_{0.93}$ is the 93rd percentile of the diameter distribution (in), A is the age of the plantation (years), T is the number of trees planted per acre and S is the site index (height of dominant trees at reference age of 25 years).

In the published model, equation (4) is only one of three possible survival curves. We chose equation (4) because it was the one that related N to T. Equation (5) is derived by fitting the model $\ln H_{\mathrm{D}} = \beta_{51} + \beta_{52} A^{-0.5}$, then imposing the constraint $\ln S = \beta_{51} + \beta_{52}(25)^{-0.5}$ and solving for β_{51} in terms of S and β_{52}.

2.1.2 *Tree-level models*

$$H = \theta_{11} H_{\mathrm{D}}^{\theta_{12}} B^{\theta_{13}} N^{\theta_{14}} \exp(\theta_{15} + \theta_{16} D) \tag{6}$$

$$V = \theta_{21} + \theta_{22} D^2 H + \theta_{23} (D^2 H)^2 \tag{7}$$

where H is the individual tree height (ft), D is the individual tree diameter (in) and V is the individual tree volume (ft^3).

2.1.3 Model operation. The model operates as follows. The user specifies A, S and T. Equation (5) is invoked to predict H_{D}. Then, using equations (1)–(3), predictions for B, $D_{\min}$ and $D_{0.93}$ are obtained. Recall that it is assumed that the diameter distribution follows the Weibull density:

$$f(D) = \frac{c}{b}\left(\frac{D-a}{b}\right)^{c-1} \exp\left(\frac{D-a}{b}\right)^{c}, \qquad D \geq a, a \geq 0, b > 0, c > 0 \tag{8}$$

The three parameters of the Weibull are calculated deterministically from B, $D_{\min}$ and $D_{0.93}$:

$$a = D_{\min}/2 \tag{9}$$

c is the solution to

$$a^2 + 2a(D_{0.93} - a)\frac{\Gamma[1-(1/c)]}{-[\ln(1-p)]^{1/c}} + (D_{0.93} - a)^2 \frac{\Gamma[1+(2/c)]}{-[\ln(1-p)]^{2/c}} - \frac{B}{kN} = 0 \tag{10}$$

and

$$b = \frac{D_{0.93} - a}{-[\ln(1-p)]^{1/c}} \tag{11}$$

where $p = 0.93$ and $k = 0.005\,454$.

The percentage of trees between diameters D_1 and D_2 ($D_2 > D_1$) is found by $F(D_2 | a, b, c) - F(D_1 | a, b, c)$, where F denotes the cumulative distribution function of the Weibull density, which exists in closed form.

The published model also includes a provision for predicting the yield of thinned stands. For simplicity, we will only be concerned with unthinned stands here.

3 Data

The data are described in Zarnoch *et al.* (1991) and interested readers are referred there. We only provide a brief description here. The data come from 507 unthinned plots in slash pine plantations on cut-over lands. The plots varied in size from 0.1 to 0.25 acres. Among the variables measured on each plot were the height of dominants, and the diameter and height of each tree. This resulted in 6326 individual tree observations. The age and number of trees planted were known from plantation records. It is worth mentioning that observations on individual trees from the same plot are clearly not independent. However, they were treated as such in the published model, so we also adopt this assumption, which is a common assumption in growth and yield modelling—and not without merit (Green *et al.*, 1994).

4 Bayes model

4.1 Overview

We generate a joint predictive posterior sample for (B, N, $D_{\min}$, $D_{0.93}$ and H_{D}), given the model inputs A, T and S. Then, for each observation in the posterior sample, we solve for (a, b, c) using equations (9)–(11). We assume that equations (9)–(11) are deterministic, as was done in the published model. Let N_j be the value of N for the jth observation in the posterior sample. We generate N_j independent random numbers (diameters) from the Weibull distribution derived from the jth observation. Then, for each diameter, we stochastically generate a height and a volume. The volumes are summed over the N_j trees to yield an estimate of the total volume (TV). Repeating this for every observation in the posterior sample results in a posterior predictive sample for TV.

We used the Gibbs sampler to generate the predictive posterior sample for (B, N, $D_{\min}$, $D_{0.93}$ and H_{D}). Gibbs sampling was popularized by Gelfand and Smith (1990) and has since become a standard method for performing previously intractable or challenging Bayesian analyses. For details on Gibbs sampling, see Gelfand and Smith (1990), Gelfand *et al.* (1990), Smith and Roberts (1993), Geman and Geman (1984), among others. For examples of Gibbs sampling in forestry problems, see Green and Strawderman (1992) and Green *et al.* (1994). To use the Gibbs sampler, one must derive the 'full conditional' distribution for each parameter. In the present context, we need the following distributions (we borrow the Gelfand and Smith (1990) bracket notation to denote pdfs):

$[B \mid \ldots]$
$[N \mid \ldots]$
$[D_{\min} \mid \ldots]$
$[D_{0.93} \mid \ldots]$
$[H_{\mathrm{D}} \mid \ldots]$

where '. . .' is to be understood as 'all the other variables in the model'.

In the published model, no distributional assumptions were presented. In line with our stated goal of introducing as few assumptions as possible, we avoid specifying error distributions for equations (1)–(5), and we also do not impose prior distributions for B, N, $D_{\min}$, $D_{0.93}$ and H_{D}. Instead, we follow Zellner's (1994) Bayesian method of moments/instrumental variable (BMOM/IV) approach. We now summarize the relevant concepts of the BMOM/IV method.

4.2 BMOM/IV

The BMOM/IV method is predicated on conditioning the first two moments of the predictive posterior to be equal to the moments observed in the sample data. Following this, 'maximum entropy' or 'maxent' (for example, see Jaynes, 1982) is used to find the most conservative choice of density that agrees with the side-conditions on the moments.

Zellner presents results for the scalar mean problem and the regression problem. Our interest is in the regression problem. Suppose we entertain the following model:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{u} \tag{12}$$

where $\mathbf{Y}$ is an $n \times 1$ vector of observations, $\mathbf{X}$ is an $n \times p$ design matrix, $\boldsymbol{\alpha}$ is a $p \times 1$

vector of regression coefficients with unknown values, and $\mathbf{u}$ is an $n \times 1$ vector of realized (but unknown) error terms. Pre-multiplication of both sides of equation (12) by $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and taking posterior expectations yields[1]

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{E}\boldsymbol{\alpha} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}\mathbf{u}$$

Now, we impose the conditions on the moments. Specifically, we assume

(1) $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{E}\mathbf{u} = 0$.
(2) $\mathbf{E}(\mathbf{u} - \hat{\mathbf{u}})(\mathbf{u} - \hat{\mathbf{u}})' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\tau^2$, where τ^2 is a positive constant.

Using the above assumptions, Zellner shows the following:

(1) The proper maxent posterior density for α, given τ^2 and the data, is a multivariate normal density with mean $\boldsymbol{\alpha} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ and covariance matrix $(\mathbf{X}'\mathbf{X})^{-1}\tau^2$.
(2) The proper maxent posterior density for τ^2 with $\mathbf{E}\tau^2 = s^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}/(n-p)$ is the exponential density $g_e(\tau^2 | \mathbf{X}, \mathbf{Y}, s^2) = (1/s^2)\exp(-\tau^2/s^2)$, $0 < \tau^2 < \infty$.
(3) If we denote a future observation $\mathbf{y}_f = \mathbf{x}_f\boldsymbol{\alpha} + \mathbf{u}_f$, then the proper maxent predictive density for $\mathbf{y}_f$, given τ^2, $\mathbf{x}_f$, $\mathbf{X}$ and $\mathbf{Y}$, is a normal density $\mathbf{h}_N(\mathbf{y}_f | \tau^2, x_f, \mathbf{X}, \mathbf{Y}) \sim N[\hat{\mathbf{y}}_f, (1 + \mathbf{x}_f(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_f')\tau^2)]$, where $\hat{\mathbf{y}}_f = \mathbf{x}_f\boldsymbol{\alpha}$.

4.3 Modifications to published model

We make two substantive modifications to the published model. First, we linearize equations (1)–(6) by taking logarithms. This permits the use of the BMOM/IV method. The second modification arises from our reluctance to accept the implicit assumption in the published model that B, $D_{\min}$ and $D_{0.93}$ are independent. We 'build in' interdependence among these variables, by altering equations (1)–(3) in a hierarchical fashion to give

$$\ln B = \ln \beta_{11} + \beta_{12} \ln H_D + \beta_{13} \ln N + \beta_{14} A^{-1} \quad (13)$$

$$\ln D_{0.93} = \ln \beta_{21} + \beta_{22} \ln H_D + \beta_{23} \ln N + \beta_{24} A^{-1} + \beta_{25} \ln B \quad (14)$$

$$\ln D_{\min} = \ln \beta_{31} + \beta_{32} \ln H_D + \beta_{33} \ln N + \beta_{34} A^{-1} + \beta_{35} \ln B + \beta_{36} \ln D_{0.93} \quad (15)$$

The forms of equations (13)–(15) were selected after brief experimentation. The PRESS statistic (see, for example, Green, 1983) was used to evaluate the model forms. However, a more thorough investigation might result in different model forms. As stated earlier, our intent is to modify the published model as little as possible, while putting it in a Bayesian framework.

In the published model, no error terms were included, although standard errors from the fitting process were reported. We will explicitly include error terms in the model. Specifically, we will assume that the right-hand side of each model should be augmented (after linearization) by '$+\mathbf{u}_i$', $i = 1, 2, \ldots, 5$. The list of full conditionals from which we must generate samples is then augmented by the following five distributions: $[\sigma_i^2 | \ldots]$, $i = 1, 2, \ldots, 5$, where σ_i^2 is the error variance for model i.

4.4 Derivation of full conditional distributions

Let $\mathbf{Y}_1$, $\mathbf{Y}_2$, $\mathbf{Y}_3$, $\mathbf{Y}_4$ and $\mathbf{Y}_5$ be $n \times 1$ vectors containing the $n = 507$ observations on $\ln B$, $\ln D_{\min}$, $\ln D_{0.93}$, $[(\ln T - \ln N)/\ln T]$ and $\ln H_D$, respectively. For brevity, let $Q = (\ln T - \ln N)/\ln T$. Furthermore, let $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, $\mathbf{X}_4$ and $\mathbf{X}_5$ be matrices with dimensions $(n \times 4)$, $(n \times 5)$, $(n \times 6)$, $(n \times 1)$ and $(n \times 2)$, respectively, with a column of ones for the intercept (except for $\mathbf{X}_4$), and observations on the appropriate stand attributes in the remaining columns. Finally, let $\hat{\beta}_i$, $i = 1, 2, \ldots, 5$, be the least-squares estimates of β_i, and let s_i^2, $i = 1, 2, \ldots, 5$, be the usual mean square errors. Then, using the BMOM/IV method, and exploiting the hierarchical structure of equations (11)–(13) as in Wakefield *et al.* (1994), we find the full conditional predictive distributions for $\ln B$, $\ln D_{0.93}$, $\ln D_{\min}$ and Q to be (up to proportionality)

$$[\ln B \mid \ln D_{\min}, \ln D_{0.93}, \sigma^2, \mathbf{x}^*, \hat{\beta}) \propto (\Pi_{i=1}^3 \delta_i)^{-1} \exp\{-[\tfrac{1}{2}(\ln B - \mathbf{x}_1^* \hat{\beta}_1)^2/\delta_1^2] - [\tfrac{1}{2}(\ln D_{0.93} - \mathbf{x}_2^* \hat{\beta}_2)^2/\delta_2^2] - [\tfrac{1}{2}(\ln D_{\min} - \mathbf{x}_3^* \hat{\beta}_3)^2/\delta_3^2]\} \tag{16}$$

$$[\ln D_{0.93} \mid \ln D_{\min}, B, \sigma^2, \mathbf{x}^*, \hat{\beta}) \propto (\Pi_{i=2}^3 \delta_i)^{-1} \exp\{-[\tfrac{1}{2}(\ln D_{0.93} - \mathbf{x}_2^* \hat{\beta}_2)^2/\delta_2^2] - [\tfrac{1}{2}(\ln D_{\min} - \mathbf{x}_3^* \hat{\beta}_3)^2/\delta_3^2]\} \tag{17}$$

$$[\ln D_{\min} \mid (\ln D_{\min}, B, \sigma^2, \mathbf{x}^*, \hat{\beta})] \propto \frac{1}{\delta_3} \exp\left[-\tfrac{1}{2}(\ln D_{\min} - \mathbf{x}_3^* \hat{\beta}_3)^2/\delta_3^2\right] \tag{18}$$

$$[Q \mid \ln B, \ln D_{0.93}, \ln D_{\min}, A, \sigma^2, \mathbf{x}^*, \hat{\beta}) \propto (\Pi_{i=1}^4 \delta_i)^{-1} \exp\{-[\tfrac{1}{2}(Q - \mathbf{x}_4^* \hat{\beta}_4)^2/\delta_4^2] - [\tfrac{1}{2}(\ln B - \mathbf{x}_1^* \hat{\beta}_1)^2/\delta_1^2] - \tfrac{1}{2}(\ln D_{0.93} - \mathbf{x}_2^* \hat{\beta}_2)^2/\delta_2^2] - \tfrac{1}{2}(\ln D_{\min} - \mathbf{x}_3^* \hat{\beta}_3)^2/\delta_3^2]\} \tag{19}$$

where $\mathbf{x}_i^*$ represents the level of the independent variables for the new observation (i.e. $\mathbf{x}_1^* = (1, \ln H_{\mathrm{D}}, \ln N, A^{-1})$, $\mathbf{x}_2^* = (1, \ln H_{\mathrm{D}}, \ln N, A^{-1}, \ln B)$, $\mathbf{x}_3^* = (1, \ln H_{\mathrm{D}}, \ln N, A^{-1}, \ln B, \ln D_{0.93})$ and $\mathbf{x}_4^* = A$); $\sigma^2 = (\sigma_i^2, i = 1, 2, \ldots, 5)$; $\mathbf{x}^* = (\mathbf{x}_1^{*\prime}, \mathbf{x}_2^{*\prime}, \mathbf{x}_3^{*\prime}, \mathbf{x}_4^{*\prime}, \mathbf{x}_5^{*\prime})'$; $\hat{\beta} = (\hat{\beta}_1', \hat{\beta}_2', \hat{\beta}_3', \hat{\beta}_4', \hat{\beta}_5')'$; and $\delta_i^2 = \sigma_i^2[1 + \mathbf{x}_i^* (\mathbf{X}_i'\mathbf{X}_i)^{-1} \mathbf{x}_i^{*\prime}]$, $i = 1, 2, \ldots, 4$.

There is a slight complication with the full conditional for $\ln H_{\mathrm{D}}$. As previously mentioned, we do not fit equation (5), but, instead, the linear model $\ln H_{\mathrm{D}} = \beta_{51} + \beta_{52} A^{-0.5}$. The intercept is removed by imposing the constraint $\ln S = \beta_{51} + \beta_{52}(25^{-0.5})$. Hence, $\ln H_{\mathrm{D}}$ is a function of the unknown coefficient β_{52}, and we have $\mathbf{E}(\ln H_{\mathrm{D}}) = \ln S + (A^{-0.5} - 0.2)\mathbf{E}(\beta_{52})$ and $\mathrm{Var}(\ln H_{\mathrm{D}} = (A^{-0.5} - 0.2)^2 \mathrm{Var}(\beta_{52})$. By applying the BMOM/IV method, we know that the conditional posterior distribution of β_{52} is normal.[2] Hence, we include the full conditional distribution for β_{52} in the Gibbs sampler. For each generated value of β_{52}, we can determine a value for $\ln H_{\mathrm{D}}$, by simply multiplying the generated value by $(A^{-0.5} - 0.2)$ and adding $\ln S$. This is proper, because we know that, if the Gibbs sampler is used to produce a marginal posterior sample of a given variable, then a marginal posterior sample of a function of that variable is available by simply computing the appropriate functional values for each observation in the original marginal posterior sample (Gelfand *et al.*, 1990). The full conditional for β_{52} is (up to proportionality)

$$[\beta_{52} \mid \ln B \mid \ln D_{0.93}, \ln D_{\min}, A, \sigma^2, \mathbf{x}^*, \hat{\beta}) \propto (\Pi_{i=1}^5 \delta_i)^{-1} \exp\{-[\tfrac{1}{2}(\beta_{52} - \hat{\beta}_{52})^2/\delta_5^2] - [\tfrac{1}{2}(Q - \mathbf{x}_4^* \hat{\beta}_4)^2/\delta_4^2] - [\tfrac{1}{2}(\ln B - \mathbf{x}_1^* \hat{\beta}_1)^2/\delta_1^2] - [\tfrac{1}{2}(\ln D_{0.93} - \mathbf{x}_2^* \hat{\beta}_2)^2/\delta_2^2] - [\tfrac{1}{2}(\ln D_{\min} - \mathbf{x}_3^* \hat{\beta}_3)^2/\delta_3^2]\} \tag{20}$$

where $\delta_5^2 = c_{22}\sigma_5^2$, and c_{22} is the (2, 2) element in $(\mathbf{X}_5'\mathbf{X}_5)^{-1}$.

We also need the full conditionals for the error variances. These are (up to proportionality)

$$[\sigma_1^2 \mid \ln B, \ln D_{0.93}, \ln D_{\min}, A, \mathbf{x}^\star, \hat{\beta}, s_1^2] \propto (1/\delta_1)(1/s_1^2)\exp\{(-\sigma_1^2/s_1^2) + [-\tfrac{1}{2}(\ln B - \mathbf{x}_1^\star \hat{\beta}_1)^2/\delta_1^2]\} \tag{21}$$

$$[\sigma_2^2 \mid \ln B, \ln D_{0.93}, \ln D_{\min}, A, \mathbf{x}^\star, \hat{\boldsymbol{\beta}}, s_2^2] \propto (1/\delta_2)(1/s_2^2)\exp\{(-\sigma_2^2/s_2^2) + [-\tfrac{1}{2}(\ln D_{0.93} - \mathbf{x}_2^\star \hat{\boldsymbol{\beta}}_2)^2/\delta_2^2]\} \tag{22}$$

$$[\sigma_3^2 \mid \ln B, \ln D_{0.93}, \ln D_{\min}, A, \mathbf{x}^\star, \hat{\boldsymbol{\beta}}, s_3^2] \propto (1/\delta_3)(1/s_3^2)\exp\{(-\sigma_3^2/s_3^2) + [-\tfrac{1}{2}(\ln D_{\min} - \mathbf{x}_3^\star \hat{\boldsymbol{\beta}}_3)^2/\delta_3^2]\} \tag{23}$$

$$[\sigma_4^2 \mid \ln B, \ln D_{0.93}, \ln D_{\min}, A, \mathbf{x}^\star, \hat{\boldsymbol{\beta}}, s_4^2] \propto (1/\delta_4)(1/s_4^2)\exp\{(-\sigma_4^2/s_4^2) + [-\tfrac{1}{2}(Q - \mathbf{x}_4^\star \hat{\boldsymbol{\beta}}_4)^2/\delta_4^2]\} \tag{24}$$

$$[\sigma_5^2 \mid \ln B, \ln D_{0.93}, \ln D_{\min}, A, \mathbf{x}^\star, \hat{\boldsymbol{\beta}}, s_5^2] \propto (1/\delta_5)(1/s_5^2)\exp\{(-\sigma_5^2/s_5^2) + [-\tfrac{1}{2}(\beta_{52} - \hat{\beta}_{52})^2/\delta_5^2]\} \tag{25}$$

where s_i^2, $i = 1, 2, \ldots, 5$, is the usual least-squares mean square error (MSE) for the ith model.

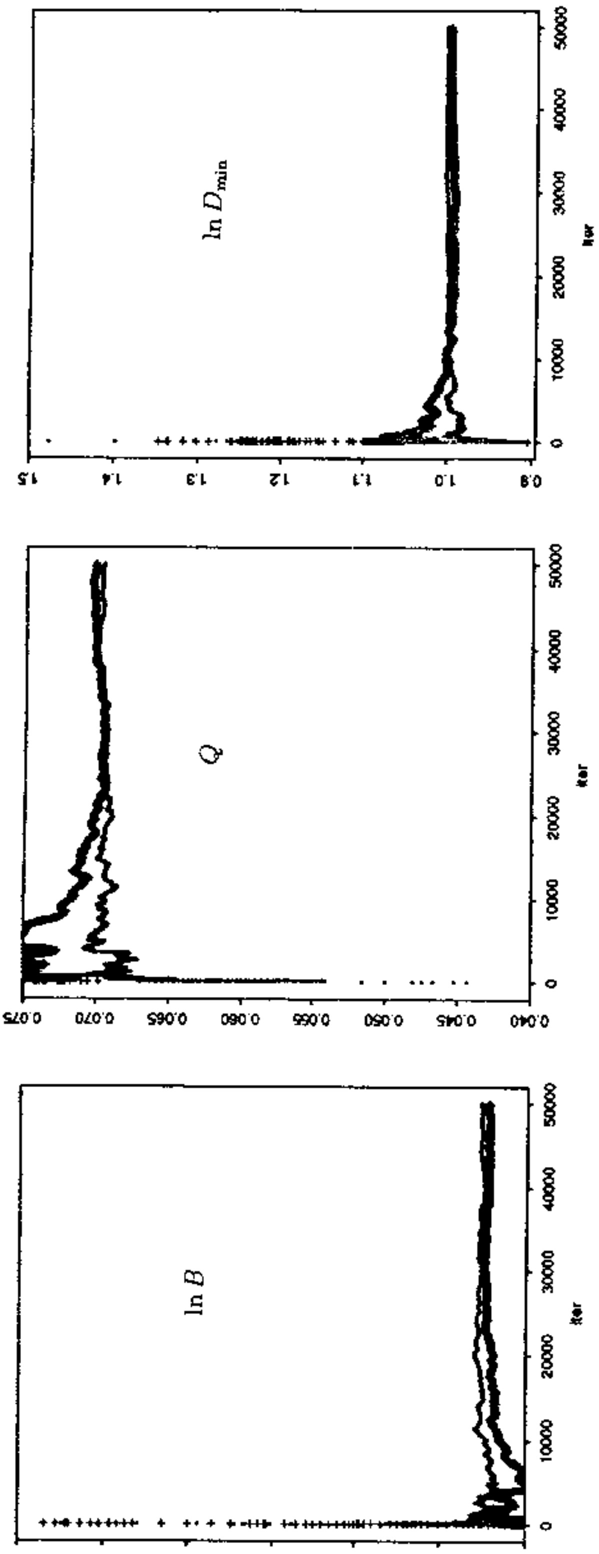
4.5 *Implementation of the Gibbs sampler*

Sampling from each full conditional distribution was accomplished with the ratio-of-uniforms method (Wakefield *et al.*, 1991). We simulated $\ln\sigma_i^2$ and logit(Q), because ratio-of-uniforms works best for variables defined over the whole real line.

For each observation in the sample, N diameters were generated as mentioned above. Heights and then volumes were stochastically generated for each diameter using the ordinary least-squares (OLS) estimates for θ_i, $i = 1, 2$, and the predictive variances $\{\mathrm{MSE}[1 + \mathbf{z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}']\}$, with $\mathbf{Z}$ the appropriate design matrix, and $\boldsymbol{z}$ the vector of independent variables for which a prediction is sought. Here, we introduce the mild assumption that the error terms for equations (6) and (7) are normally distributed. Collecting the N volumes yielded an observation on the total volume.

We ran the Gibbs sampler twice, each time with unique, randomly chosen starting values. Each chain was run for 50 000 iterations. Plotting the values generated for $\ln B$, $\ln D_{0.93}$, $\ln D_{\min}$, Q and β_{52} against the iteration number revealed no indications of convergence difficulties. Next, the ergodic averages for these five variables were computed. The results are displayed in Fig. 1. The figure suggests that the sampler seems to converge after about 20 000 iterations.

We tested for convergence in the following *ad hoc* manner. First, for each of the two runs of 50 000, we discarded the first 30 039 iterates and then formed approximate identically and independently distributed (iid) samples of size 500, by selecting every 40th iterate. Then, for each variable (i.e. $\ln B$, $\ln D_{0.93}$, $\ln D_{\min}$, Q and β_{52}), we tested the means and variances of the two samples for equality with the usual t- and F-tests. The value of each test statistic ($t^\star$ and $F^\star$) and the p-values are shown in Table 1.


ln D_{min}
Q
ln B
iter
iter
iter
0
10000
20000
30000
40000
50000
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.075
0.070
0.065
0.060
0.055
0.050
0.045
0.040
4.85
4.80
4.75
4.70
4.65
4.60
4.55

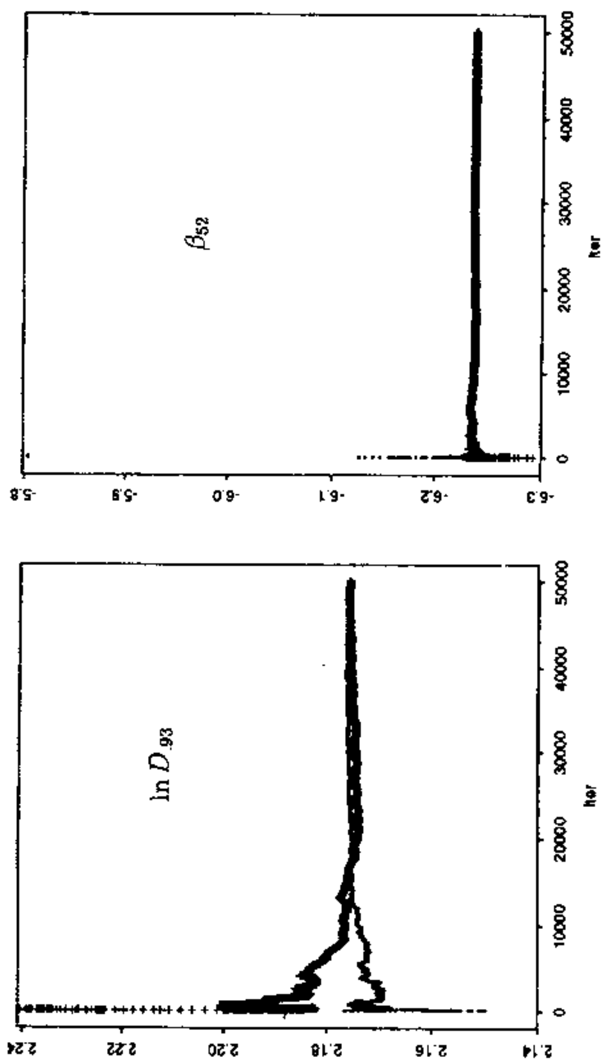


FIG. 1. Ergodic means of $\ln B$, Q, $\ln D_{\min}$, $\ln D_{0.93}$ and β_{52} vs iteration number for two runs of Gibbs sampler.

TABLE 1. Test statistics, by variable, for two samples of size 500

Variable	t^*	p	F^*	p
$\ln B$	1.44	0.15	1.12	0.18
$\ln D_{0.93}$	−1.84	0.06	1.09	0.33
$\ln D_{\min}$	1.57	0.12	1.12	0.19
Q	−1.61	0.11	1.07	0.43
β_{52}	1.34	0.18	0.81	0.02

It appears from the results in Table 1 that the sampler has converged for $\ln B$, $\ln D_{0.93}$, $\ln D_{\min}$ and Q, but there is some question about β_{52}. The means of the two samples for β_{52} agree reasonably well but the variances differ somewhat. We investigated this further by constructing Q–Q plots for the two samples of size 500 for each variable. These are displayed in Fig. 2. The Q–Q plots show that the samples for $\ln B$, $\ln D_{0.93}$, $\ln D_{\min}$ and Q agree well; furthermore, they show that the samples for β_{52} agree throughout most of the distribution. It appears that sample 2 (vertical axis) included some extreme values at the low end of the distribution, which were missed by sample 1 (horizontal axis). The inclusion of these values in sample 2 caused its estimated variance to be larger and the F^* value for the test of equality of variance to correspond to a low p-value.

It is well known that more iterations of the Gibbs sampler are required to 'nail down' the tails of the marginal posterior distributions than are required to characterize adequately the main body of the distribution (see, for example, Raftery & Lewis, 1992). Hence, after viewing the Q–Q plots, our conclusion is that we have run the sampler long enough to characterize adequately the marginal posterior distribution of β_{52}. This variable is not one of primary interest (we are mainly interested in volume and, to a lesser extent, B and N). Hence, we concluded that the sampler had, for practical purposes, converged.

Our method for assessing convergence is admittedly *ad hoc*. For more rigorous tests, we refer the reader to Gelman and Rubin (1992), Geyer (1992), Geweke (1992), Raftery and Lewis (1992), Roberts (1992) and Zellner and Min (1994), among others.

5 Results

For computational ease from here on, we will use one of the samples of size 500. Once samples are obtained from the marginal posterior distributions for the desired variables, either the means or modes of these samples are sensible point estimates. Here, we choose the means, primarily because of computational ease. The means of the samples agree remarkably well with the point estimates available from the published model. In Table 2, we present results for a plantation with $A = 20$, $T = 600$ and $S = 60$.

Perhaps more valuable than the point estimates, however, are the marginal posterior densities available from the Bayes model. In Fig. 3, we display the marginal posteriors for B, N, H_D and TV. These histograms indicate substantial variability about the mean or mode for B, N and TV. This information is unavailable in the original model. Given the data sumarized in Fig. 2, it is trivial

TABLE 2. Predicted values for growth and yield model variables from published model and Bayes model

Variable	Published	Bayes
N	387.6	393.9
H_D	51.2	51.8
B	97.3	98.4
D_{min}	2.7	2.9
$D_{0.93}$	8.8	8.8
TV	2534	2604
a	1.3	1.4
b	5.8	5.7
c	4.0	4.0

TABLE 3. Approximate 95% credible regions for N, B, H_D, D_{min}, $D_{0.93}$ and TV

Variable	0.05 percentile	0.95 percentile
N	282.4	496.0
H_D	51.4	52.1
B	77.2	120.3
D_{min}	1.3	4.8
$D_{0.93}$	8.0	9.8
TV	2005	3204

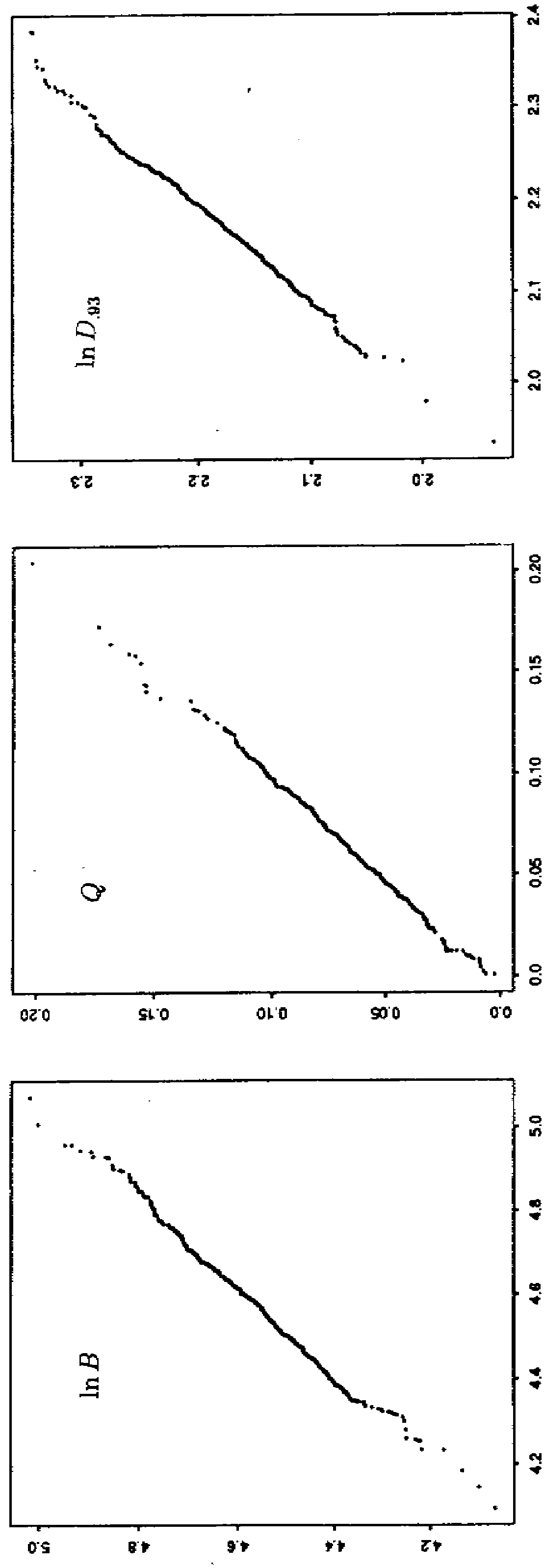
to develop Bayesian credible regions (the Bayesian analog to confidence intervals). In Table 3, we display approximate 95% credible regions. These were obtained by simply recording the fifth and 95th percentile of each marginal posterior sample.

6 Conclusions

It is difficult to evaluate the variability implied in our model (Fig. 3). To do so, we would need large numbers of plots on the same site quality land, planted at the same density and currently at the same age. Such data are rare at best. However, we feel that the most important contribution of our model is not necessarily getting the variability exactly correct, but, instead, it is reminding practitioners that the point estimates available from growth and yield models are subject to substantial error.

Notes

1. After $\mathbf{Y}$ has been observed, it is known without error. Hence, the posterior expectation of $\mathbf{Y}$ is $\mathbf{Y}$.
2. We ignore the error variance σ_5^2 in the distribution for $\ln H_D$, because we will always assume that it is the mean $\ln H_D$ that is desired. Hence, only the variation in the slope of the regression line is important (the intercept having been removed via constraint).


$\ln D_{.93}$
Q
$\ln B$

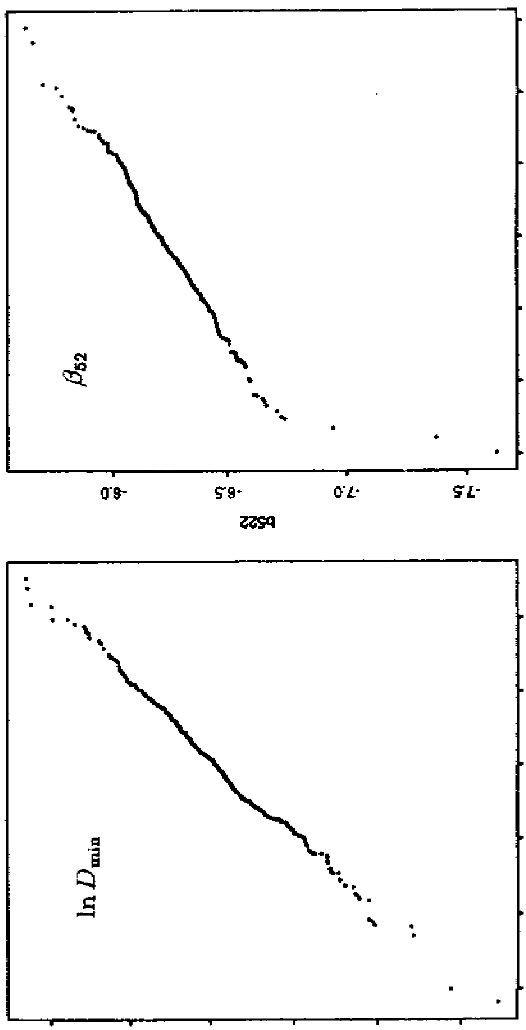


FIG. 2. Q–Q plots for samples of size 500 from each run of Gibbs sampler for $\ln B$, Q, $\ln D_{\min}$, $\ln D_{0.93}$ and β_{52}.

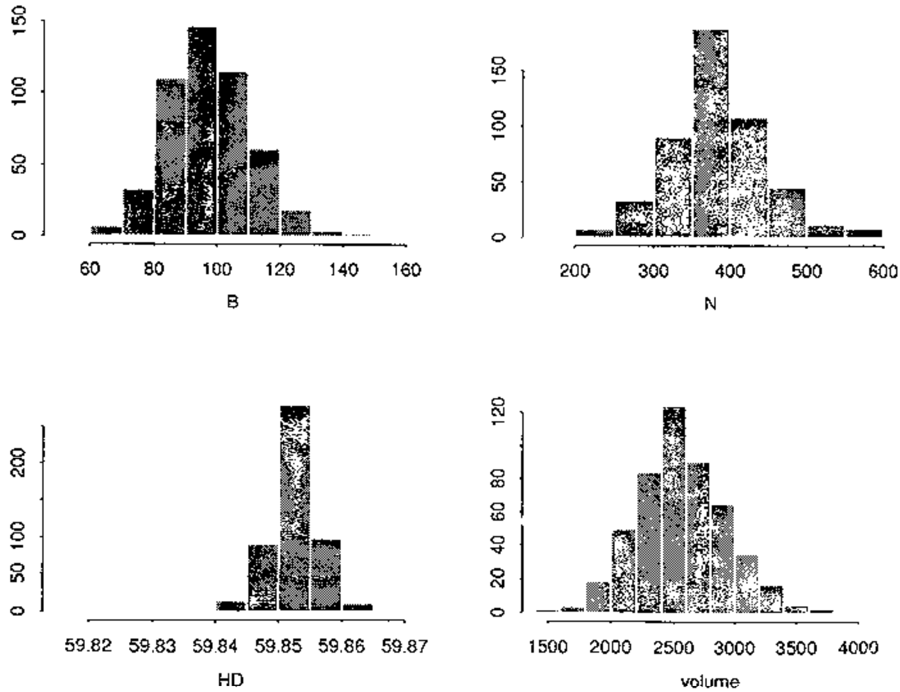


FIG. 3. Histograms from marginal posterior samples of size 500 for B, N, H_D and TV.

REFERENCES

GELFAND, A. E. & SMITH, A. F. M. (1990) Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, 85, pp. 398–409.

GELFAND, A. E., HILLS, S. E., RACINE-POON, A. & SMITH, A. F. M. (1990) Illustration of Bayesian inference in normal data models using Gibbs sampling, *Journal of the American Statistical Association*, 85, pp. 972–985.

GELMAN, A. & RUBIN, D. (1992) Inference form iterative simulation using multiple sequences (with discussion), *Statistical Science*, 7, pp. 457–472.

GEMAN, S. & GEMAN, D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, pp. 721–741.

GEWEKE, J. (1992) Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *In:* BERGER, BERNARDO, DAWID, LINDLEY & SMITH (Eds), *Bayesian Statistics 4, Proceedings of the 4th Valencia International Meeting*, pp. 169–194 (Oxford, Oxford University Press).

GEYER, C. J. (1992) Practical Markov chain Monte Carlo (with discussion), *Statistical Science*, 7, pp. 473–482.

GREEN, E. J. (1983) Evaluating the predictive abilities of regressions with PRESS, *Forest Science*, 29, pp. 712–714.

GREEN, E. J. & STRAWDERMAN, W. E. (1992) A comparison of hierarchical Bayes and empirical Bayes methods with a forestry application, *Forest Science*, 38, pp. 350–366.

GREEN, E. J., ROESCH, JR, F. A., SMITH, A. F. M. & STRAWDERMAN, W. E. (1994) Bayesian estimation of the three-parameter Weibull distribution with tree diameter data, *Biometrics*, 50, pp. 254–269.

JAYNES, E. T. (1982) On the rationale of maximum-entropy methods, *Proceedings of the IEEE*, 70, pp. 939–952.

RAFTERY, A. E. & LEWIS, S. M. (1992) How many iterations in the Gibbs sampler? *In:* BERGER, BERNARDO, DAWID, LINDLEY & SMITH (Eds), *Bayesian Statistics 4, Proceedings of the 4th Valencia International Meeting*, pp. 763–774 (Oxford, Oxford University Press).

ROBERTS, G. O. (1992) Convergence diagnostics of the Gibbs sampler. *In:* BERGER, BERNARDO, DAWID, LINDLEY & SMITH (Eds), *Bayesian Statistics 4, Proceedings of the 4th Valencia International Meeting*, pp. 775–782 (Oxford, Oxford University Press).

SMITH, A. F. M. & ROBERTS, G. O. (1993) Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo, *Journal of the Royal Statistical Society, Series B*, 55, pp. 3–23.

WAKEFIELD, J. C., GELFAND, A. E. & SMITH, A. F. M. (1991) Efficient generation of random variates via the ratio-of-uniforms method, *Statistics and Computing*, 1, pp. 129–133.

WAKEFIELD, J. C., SMITH, A. F. M., RACINE-POON, A. & GELFAND, A. E. (1994) Bayesian analysis of linear and nonlinear models using the Gibbs samples, *Applied Statistics*, 43, pp. 201–221.

ZARNOCH, S. J., FEDUCCIA, D. P., BALDWIN, V. C. & DELL, T. R. (1991) Growth and yield predictions for thinned and unthinned slash pine plantations on cutover sites in the West Gulf region, *USDA Forest Service Research Paper S0-264*.

ZELLNER, A. (1994) Bayesian method of moments/instrumental variable (BMOM/IV) analysis of mean and regression models, unpublished report, HGB Alexander Research Foundation, Graduate School of Business, University of Chicago.

ZELLNER, A. & MIN, C. (1994) Gibbs sampler convergence criteria (GSC^2), unpublished report, HGB Alexander Research Foundation, Graduate School of Business, University of Chicago.