BIOMETRICS 50, 254–269
March 1994

Bayesian Estimation for the Three-Parameter Weibull Distribution with Tree Diameter Data

Edwin J. Green,[1] Francis A. Roesch, Jr.,[2] Adrian F. M. Smith,[3] and William E. Strawderman[4]

[1]Department of Natural Resources, Cook College–Rutgers University, P.O. Box 231, New Brunswick, New Jersey 08903-0231, U.S.A.
[2]U.S.D.A. Forest Service Southern Forest Experiment Station, 701 Loyola Avenue, New Orleans, Louisiana 70113, U.S.A.
[3]Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ, United Kingdom
[4]Department of Statistics, Rutgers University, New Brunswick, New Jersey 08903-0231, U.S.A.

SUMMARY

The three-parameter Weibull density is commonly used to model the distribution of tree diameters in forest stands. We demonstrate, through likelihood profiles, that maximum likelihood estimation is often inappropriate for data from young trees due to negative estimates of the location parameter. We suggest a Bayesian model and fit it, using the Gibbs sampler, to three data sets. The latter model is easy to implement and guarantees a positive estimate for the location parameter. We illustrate some novel forms of model diagnostics, demonstrating that the Bayesian model is appropriate for two of the data sets, while it is dubious for the third. A sampling–resampling method shows that the lack of fit of the model for the latter data set is due to the likelihood, and not the prior specification.

1. Introduction

The Weibull density is commonly used in many applications. For instance, it is a standard distribution in reliability analysis (e.g., see Crowder et al., 1991). It is also widely used in forestry. In the latter context, it has primarily been used to model the distribution of tree diameters (Bailey and Dell, 1973; Ek, Issos, and Bailey, 1975; Dell et al., U.S.D.A. Forest Service Research Paper SO-147, 1979; Burk and Newberry, 1984; Burk and Burkhart, FWS-1-84, Division of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, 1984; Matney, Ledbetter, and Sullivan, 1987; Lenhart, 1988; Bailey, Burgan, and Jokela, 1989; Zarnoch et al., U.S.D.A. Forest Service Research Paper SO-264, 1991, among others). However, Krug, Nordheim, and Giese (1984) cite its use in modeling tree growth, survivorship, and height distributions. Smith and Naylor (1987) noted a difficulty in using maximum likelihood estimation for the three-parameter Weibull with reliability data, namely negative estimates for the location parameter. Negative estimation for this parameter also occurs when modeling tree diameter distribution data (e.g., see Burk and Newberry, 1984). In this paper we will examine maximum likelihood estimation for the three-parameter Weibull for diameter distribution data. We will demonstrate, by means of profile likelihoods, that maximum likelihood estimation is often inappropriate for diameter data. We will suggest an alternative Bayesian method for parameter estimation, using Gibbs sampling and vague priors. The latter is easy to implement and is guaranteed to produce nonnegative estimates for the location parameter. The plan of the paper is as follows: in the next section, we discuss the data; in Section 3 we examine the log-likelihood profiles; in Section 4 we present an alternative Bayesian estimation method; and in Section 5 we present our concluding remarks.

Key words: Bayesian inference; Forestry; Gibbs sampler; Markov chain Monte Carlo methods; Weibull density.

2. Data

Our motivation arises from interest in data from 507 temporary plots in unthinned slash pine plantations in the southern United States. Plot sizes varied from .1 to .25 acre; plantation ages ranged from 8 to 47 years. The diameters of all live trees on each plot were measured to the nearest .1 inch. More detailed information on the data can be obtained from Zarnoch et al. (U.S.D.A. Forest Service Research Paper SO-264, 1991). From these 507 plots, we selected 10 plots for detailed statistical examination. Plots were selected randomly, subject to the following criteria: only .25-acre plots were included, and only one plot per plantation could be included (i.e., if a plantation were measured at two or more ages, only one measurement could be used).

It is well known that the diameters of all trees on a plot do *not* constitute a random sample of tree diameters from a plantation due to spatial correlation among the trees on a plot. However, it is common to assume that the diameters do represent a random sample and proceed accordingly. Perhaps the most compelling reason for this assumption is that spatial information on trees is rarely collected, and thus the correlation structure among the sample trees cannot be modeled. Reynolds, Burk, and Huang (1988) have investigated the effects of correlated diameters on fitted diameter distributions. They repeatedly generated samples from populations with known correlations among the individual diameters. They then fitted distributions to the samples, assuming the sample observations to be conditionally independent. Goodness-of-fit tests yielded significance levels in accord with nominal levels; hence they concluded that inter-tree correlation was not likely to be a severe problem. We will follow other investigators and assume random samples.

3. Maximum Likelihood Estimation

We will use the following parameterization for the three-parameter Weibull:

$$f(x|\mu, \tau, \gamma) = (\gamma/\tau)((x-\mu)/\tau)^{\gamma-1} \exp\{-((x-\mu)/\tau)^{\gamma}\}, \quad x \geqslant \mu, \quad \gamma \geqslant 0, \quad \tau \geqslant 0.$$

In this parameterization μ is the location parameter, while τ and γ are the scale and shape parameters, respectively. As previously noted, it is possible for maximum likelihood estimates of μ to be negative. Smith and Naylor (1987) noted this problem for survival data and recommended a Bayesian solution. While we appreciated that this problem did occur for tree diameter data, we were surprised to discover that for the 10 plots we selected for examination, five yielded negative maximum likelihood estimates (MLEs) for μ. Evidently, this is not an uncommon problem. We selected three of the five samples for further investigation. Descriptive statistics for these samples are presented in Table 1.

In order to investigate why maximum likelihood yielded negative estimates for μ, we examined profiles of the log-likelihood function. Given the observation $X = (x_1, x_2, \ldots, x_n)$, where n is the sample size, the log-likelihood is

$$L(\mu, \tau, \gamma) = \sum_{i=1}^{n} \log f(x_i|\mu, \tau, \gamma).$$

In Figure 1 we present profile log-likelihoods for μ for the three samples. The profile log-likelihood for μ is

$$L^{\mu}(\mu, \tau, \gamma) = \sup_{\tau,\gamma} \{L(\mu, \tau, \gamma)\}.$$

The profiles clearly indicate that the MLE for μ is negative. All three samples had histograms that were noticeably skewed to the right (Figure 2). Furthermore, these samples, along with the other two that yielded negative MLEs for μ, where the five youngest plantations of the 10 originally selected. This is sensible; younger plantations tend to have smaller trees, hence the minimum

Table 1

Descriptive sample statistics for tree diameters in three samples (diameters measured to the nearest .1 inch)

Sample	Sample size	Mean diameter	Sample variance	Minimum diameter	Maximum diameter	Plantation age (yr)
1	125	7.7	3.2	2.8	11.9	17
2	127	5.5	1.3	1.4	8.0	11
3	200	5.4	1.0	1.9	7.2	13

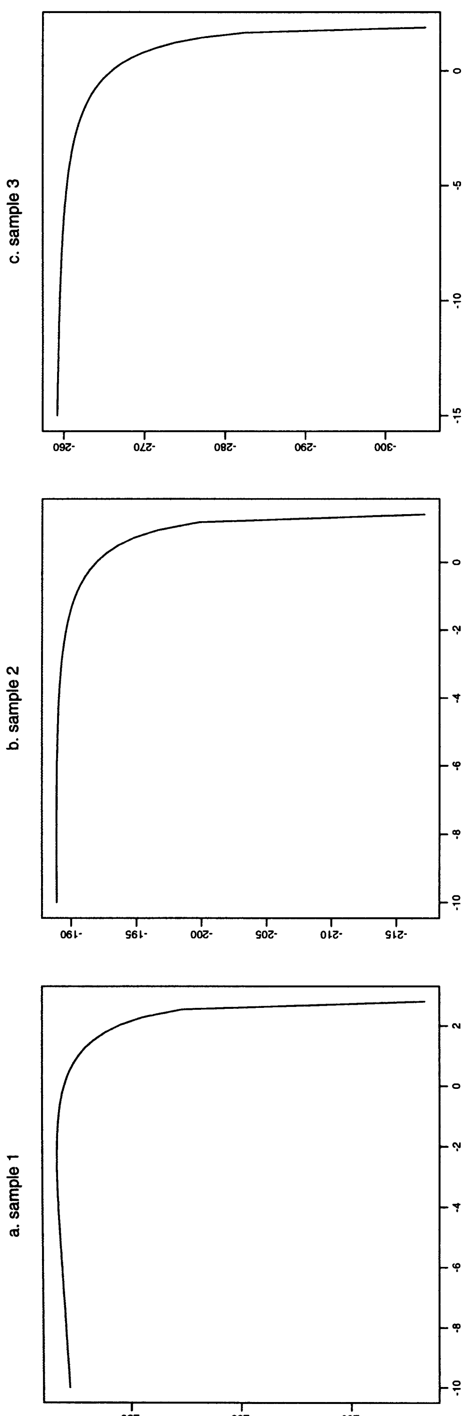


Figure 1. Profile log-likelihoods for μ for three samples.

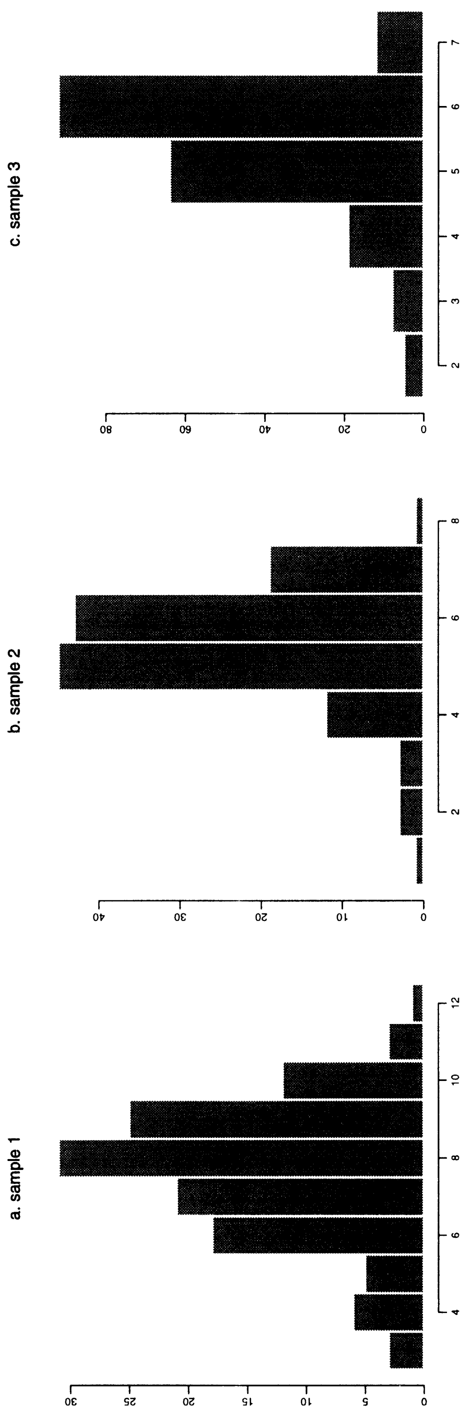


Figure 2. Histograms of diameter distributions for three samples.

diameters of such plantations are closer to zero and there is a greater chance that the MLE for the location parameter will be negative. This was also borne out by the minimum diameters: the five samples with negative MLEs for μ had the five smallest minimum diameters.

The log-likelihood profile for sample 3 is particularly intriguing. It shows the likelihood increasing with decreasing μ. The maximum is actually achieved at $\hat{\mu} = -111.203$. The profiles for τ and γ for this sample are similar, yielding the MLEs $\hat{\tau} = 117.025$ and $\hat{\gamma} = 154.468$. As reported values for the last two parameters rarely exceed, say, 9.0, this solution conflicts with common sense. For brevity, we will henceforth confine our discussion to samples 1–3.

When Bailey and Dell (1973) introduced the Weibull distribution to the forestry community, they recognized that MLEs are generally better than moment-based estimators. However, at that time the computer-intensive, iterative procedure necessary for computing MLEs for the Weibull was a serious drawback. Hence they recommended moment-based estimators. These have continued to be widely used in forestry. In some algorithms for moment-based estimates, it is still possible for the estimate of μ to be negative (e.g., Burk and Newberry, 1984), while in others (e.g., Matney et al., 1987) it is not. Many investigators simply specify μ to be a deterministic function of the minimum observed diameter [e.g., see Lenhart (1988) or Zarnoch et al. (U.S.D.A. Forest Service Research Paper SO-264, 1991)]. In tests of method of moments versus constrained MLE (where μ is forced to be positive), it often appears that method of moments performs better. However, we believe this is due to the constraint on μ. It is unusual for investigators to report measures of dispersion, such as standard errors, when estimating Weibull parameters for diameter distributions. This is especially true for method of moment techniques, in which parameters are often determined through nonlinear transformations of sample moments. In contrast, as will be seen below, with our procedure one obtains not only a point estimate, but the entire marginal posterior distribution for each parameter. This information could be valuable in simulation experiments.

4. Bayesian Model

The Bayesian model is completed by specifying a prior distribution for μ, τ, and γ. The joint posterior distribution is then obtained by application of Bayes' theorem:

$$p(\mu, \tau, \gamma|X) \propto l(\mu, \tau, \gamma|X)p(\mu, \tau, \gamma),$$

where the first term on the right-hand side is the likelihood and the second is the prior. We will take the priors for μ, τ, and γ to be independent, noting that dependencies in the posterior will arise from the data. Inferences about sets of (μ, τ, γ) should be based on the joint posterior, while if interest is focused on any particular parameter, say μ, its marginal posterior distribution is obtained by integrating the others (τ, γ) out of the joint posterior.

Although the Weibull distribution has been used in a large number of diameter distribution models, investigators do not often report the parameter estimates from individual samples (instead, predicted values are usually presented), so data-based priors are not readily available. Thus we choose vague priors (some use will be made of our expectation of the ranges of τ and γ later). As τ and γ are constrained to be positive, we adopt Jeffreys' prior for positive parameters (Jeffreys, 1961), i.e., $p(\tau) \propto (1/\tau)$ and $p(\gamma) \propto (1/\gamma)$. Regarding μ, we know for certain that it is in the interval $(0, x_{[1]}]$, where $x_{[1]}$ is the first order statistic, i.e., the minimum diameter in the sample. Beyond this we have little reason to prefer one value for μ over another. Thus we choose the prior for this parameter to be uniformly distributed on $\mathbf{R}^+$. The condition $x_{[1]} \geqslant \mu$ ensures that μ will be in the desired interval. The full Bayesian model is thus

$$p(\mu, \tau, \gamma|x) \propto (\gamma^{n-1}/\tau^{n+1})\left\{\prod_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma-1}\right\} \exp\left\{-\sum_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma}\right\},$$
$$x_{[1]} \geqslant \mu > 0, \quad \tau > 0, \quad \gamma > 0. \qquad (1)$$

The normalizing constant for (1) seems impossible to obtain analytically. Smith and Naylor (1987) performed computations using an adaptive quadrature technique, but it appears that the rather sophisticated numerical-analytic basis of their method has inhibited its routine adoption and use. We therefore opt for a stochastic simulation procedure, which although not computationally efficient when compared with finely tuned quadrature, has the advantage of simplicity of implementation. The Gibbs sampler, first introduced by Geman and Geman (1984) and popularized for Bayesian computation by Gelfand and Smith (1990) and Gelfand et al. (1990), provides a convenient method for generating samples from a posterior distribution. Technical details concerning the Gibbs sampler are provided in the above papers, and will merely be summarized here. To use the sampler, it must be possible to sample from the full conditional distribution of each parameter. Following an initial guess at the value of each parameter, the Gibbs sampler proceeds by iteratively generating a new value for each parameter conditional on the data and the most recent values of the other parameters. For instance, suppose μ_0, τ_0, and γ_0 are our initial guesses. We generate μ_1 from $p(\mu|\tau_0, \gamma_0, X)$, then τ_1 from $p(\tau|\mu_1, \gamma_0, X)$, and finally γ_1 from $p(\gamma|\mu_1, \tau_1, X)$. This constitutes one cycle of the sampler. In the next cycle, initial values are replaced by those from the first cycle, and values from the first cycle by those from the second, and so on. It can be shown [e.g., see the appendix and references in Smith and Roberts (1993)] that, under rather weak regularity conditions, the generated values for triples (μ, τ, γ) are, asymptotically, realizations from the joint posterior distribution, and running (ergodic) averages of realized values are consistent estimators of posterior expectations.

There are many different strategies for running the Gibbs sampler. Some investigators [e.g., Gelfand et al. (1990), Lange, Carlin, and Gelfand (1992), Wakefield et al. (1994, first example)] run multiple chains until "convergence," accepting the values from the last cycle of each chain to form samples from the joint posterior. Others [e.g., Green and Strawderman (1992), Wakefield et al. (1994, second example)] run one long chain. Subsequent to "convergence," the generated values are accepted as a sample from the posterior. If an independent, identically distributed (iid) sample is required, then suitably spaced variates may be accepted, say every tenth. The former strategy (multiple chains) seems to be less efficient than the latter, but may be preferable for exploratory pilot studies. Here, "convergence" is to be interpreted as modulo some graphical or numerical stopping rule diagnostics. For more discussion on strategies and convergence issues see the above papers and Hills and Smith (1992), Raftery and Lewis (1992), Gelman and Rubin (1992), Geyer (1992), and Smith and Roberts (1993). In the present study we used the latter approach (one long chain). Experimentation with different starting values convinced us that the chain was converging and covering the entire posterior distribution. We discuss monitoring convergence later.

4.1 *Random Variable Generation*

As mentioned above, to run the Gibbs sampler we must be able to sample from the full conditional distribution for each parameter. The full conditionals are (up to proportionality):

$$p(\tau|\mu, \gamma, X) \propto \tau^{-(n\gamma+1)} \exp\left\{-\sum_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma}\right\},$$

$$p(\gamma|\mu, \tau, X) \propto \gamma^{n-1}\left\{\prod_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma-1}\right\} \exp\left\{-\sum_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma}\right\},$$

$$p(\mu|\tau, \gamma, X) \propto \left\{\prod_{i=1}^{n} (x_i - \mu)^{\gamma-1}\right\} \exp\left\{-\sum_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma}\right\}.$$

Generation of a random variable from $p(\tau|\mu, \gamma, X)$ is easily accomplished by drawing a random variate z from a gamma distribution with parameter n and then letting $\tau = (\sum_{i=1}^{n} (x_i - \mu)^{\gamma}/z)^{1/\gamma}$. No similar transformation suggests itself for $p(\gamma|\mu, \tau, X)$ or $p(\mu|\tau, \gamma, X)$. However, it is easy to show that $p(\gamma|\mu, \tau, X)$ is log-concave [i.e., the first derivative of $\ln p(\gamma|\mu, \tau, X)$ decreases monotonically with increasing γ]. Hence we use the highly efficient adaptive rejection technique of Gilks and Wild (1992) to generate random values of γ.

The final conditional, $p(\mu|\tau, \gamma, X)$, is not always log-concave (this condition depends on the value of γ). Hence to generate values for μ we use an envelope rejection method. At each cycle of the Gibbs sampler, let

$$h(\mu) = \left\{\prod_{i=1}^{n} (x_i - \mu)^{\gamma-1}\right\} \exp\left\{-\sum_{i=1}^{n} ((x_i - \mu)/\tau)^{\gamma}\right\},$$

$$g(\mu) = \gamma(x_{[1]} - \mu)^{\gamma-1}/x_{[1]}^{\gamma},$$

$$c = \sup_{\delta}\left[\left\{\prod_{i=2}^{n} (x_{[i]} - \delta)^{\gamma-1}\right\} \exp\left\{-\sum_{i=1}^{n} ((x_i - \delta)/\tau)^{\gamma}\right\}\right],$$

where the current values are used for τ and γ, and $x_{[i]}$ is the ith order statistic. We sample a value μ^* from $g(\mu)$, and a value $r \sim$ Uniform(0, 1). If $r \leqslant h(\mu^*)/(cg(\mu^*))$, we accept μ^*; otherwise we draw new values for μ^* and r and reapply the test. The accepted values are distributed according to $h(\mu)$ [for proof, see Ripley (1987) or Dagpunar (1988)], which is identical to $p(\mu|\tau, \gamma, X)$, up to a proportionality constant.

4.2 *Convergence of Gibbs Sampler*

We determined convergence in two manners. First, as previously mentioned, we ran the sampler with several different starting values for μ, τ, and γ. Once we were convinced that the ergodic averages were converging to the same points, we ran one long "production" chain. We determined convergence of this chain by monitoring the ergodic averages for the three parameters. We found convergence for τ and γ to be relatively quick, but slow for μ. In general, approximately 10,000 iterations were required for the sample of the latter to converge. However, computation time was not excessive, and so we did not search extensively for a more efficient simulation method. A suitable reparameterization would probably speed the sampler up considerably [e.g., see Hills and Smith (1992) or Wakefield (1993)], but this would require sophisticated user input, which conflicts with our desire for a simply implemented procedure.

In any event, we ran the sampler for 50,000 iterations for each data set. This required about 8.2 hours of system time on a Sun cluster. We saved each iterate, but for computational ease, we base our posterior samples on the 500 observations obtained by discarding the first 30,000 iterations and then accepting every 40th iteration. We emphasize that we do this simply to make the posterior sample manageable, and not to generate iid posterior samples (although that is the effective result). As we ran the sampler so long after apparent convergence, ours was a very conservative procedure.

4.3 *Posterior Distributions*

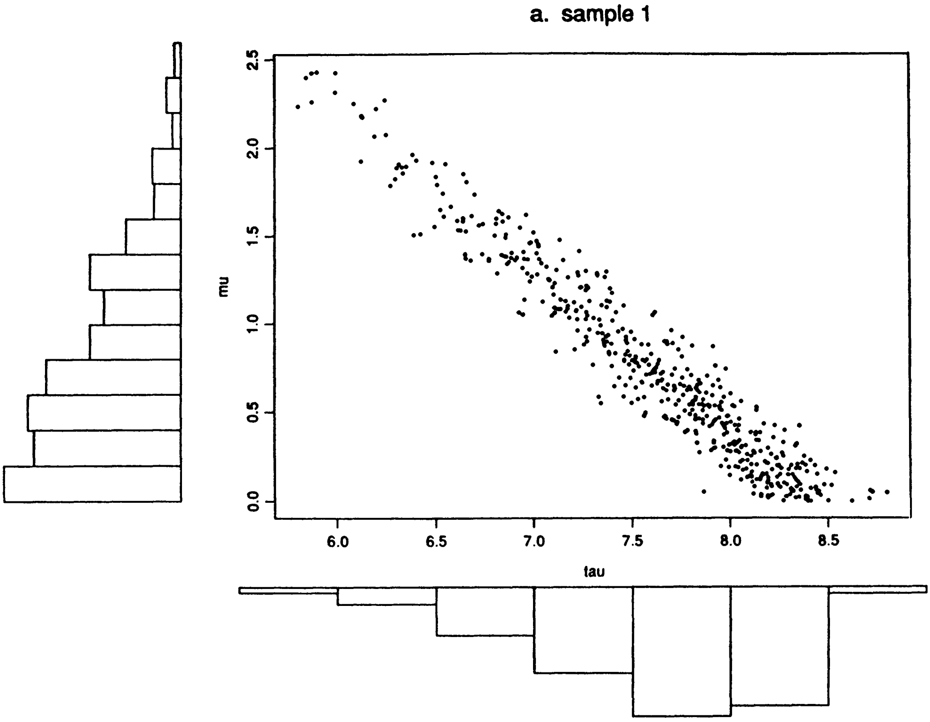
The posterior distributions for samples 1–3 are displayed in Figure 3 as scatterplots from the appropriate bivariate posteriors, with histograms of the marginal posteriors along the axes. Of course, the prior on μ (Uniform(0, ∞)) ensures that the marginal posterior for this parameter puts mass only on positive values, as desired. However, even though the mode of the marginal posterior for μ is near zero for each sample (data set), there is still considerable mass placed on larger values (up to about 2.0 for sample 1, and .8 for samples 2 and 3). Hence the usual solution in the forestry literature of setting $\mu = 0$ whenever the estimate is negative (see, e.g., Lenhart, 1988) can be highly misleading. The reason for the slow convergence of μ is apparent in Figures 3a, 3d, and 3g. This parameter is highly correlated with τ. Convergence is attained most rapidly when the bivariate posterior distributions are spherical (Wakefield, 1993). We performed a number of transformations on the data in Figure 3g in an attempt to produce a spherical distribution. None of the obvious transformations succeeded, and we consider this to be an area for future work.

It is worth emphasizing one desirable characteristic of the marginal posterior distributions obtained with the Gibbs sampler: once the marginal posterior for a parameter has been obtained, the marginal posterior distribution of any function of that parameter is available through a direct transformation of the values in the posterior sample (Gelfand and Smith, 1990).

4.4 *Model Diagnostics*

In order to see whether the problems encountered with maximum likelihood were due to a lack of fit of the Weibull density, we performed several diagnostics. First, for each triple (μ, τ, γ) generated from the joint posterior, we generated a predictive sample of X with size equal to the original size. To make these samples readily comparable with the observed data histogram, the generated values were grouped into .1-inch classes. If the Bayesian model (likelihood and prior) is appropriate, then the observed histogram should not be "in the tail" of the predictive histograms. Working with the cumulative distribution functions (cdf's), those of the predictive samples are shown as unconnected dots in Figures 4a–c, while the empirical distribution functions from the data sets are shown as solid lines. Evidently, the data are compatible with the model for samples 1 and 2, while there is an indication of incompatibility for sample 3.

We then used a version of the Bayesian significance testing procedure proposed in work as yet unpublished by Gelman et al. to quantify the above graphical impressions. We used the one-sample Kolmogorov–Smirnoff (KS) test, for which the test statistic is D, the maximum absolute difference between the observed empirical distribution function (edf) and the hypothesized cdf. The Bayesian test is performed as follows. First, draw a triple (μ, τ, γ) from the joint posterior distribution $p(\mu, \tau, \gamma|X)$ and calculate D, given the observed X and (μ, τ, γ). Next, generate a predictive sample X^*, given (μ, τ, γ), and calculate D_{pred}, the test statistic arising from X^* and (μ, τ, γ).



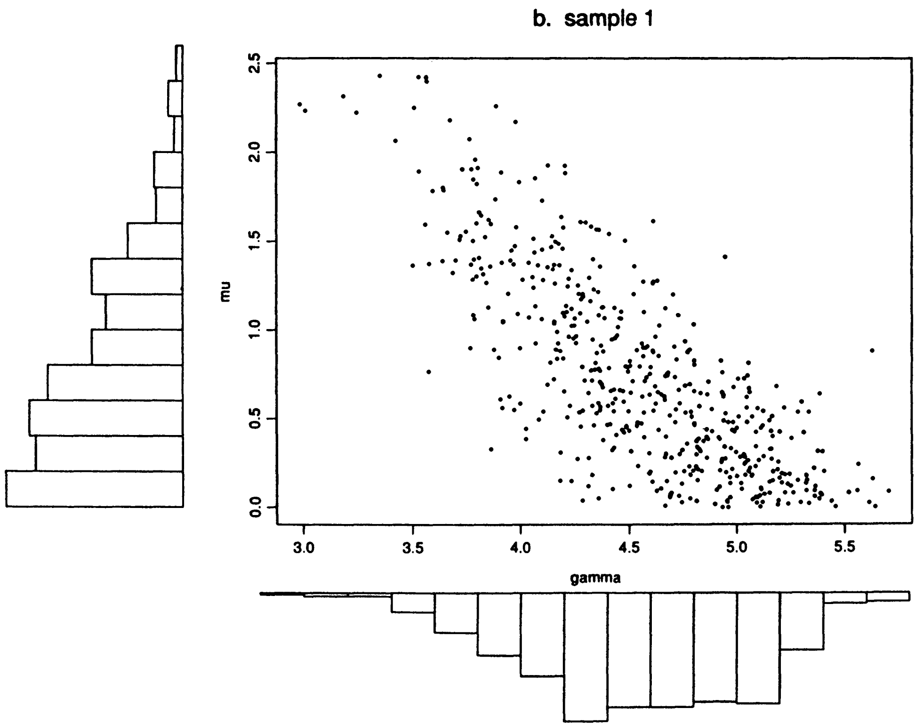


Figure 3. Scatterplots of bivariate posterior distributions, with histograms of marginal posteriors on the axes, for three samples.

Repeat the procedure a large number of times; the estimated P-value is the number of times D_{pred} exceeds D (see unpublished report of Gelman et al. for more details).

We performed the Bayesian version of the KS test using the same samples of size 500 from the joint posterior distributions that were used in Figures 3 and 4. The resulting values of D and D_{pred}

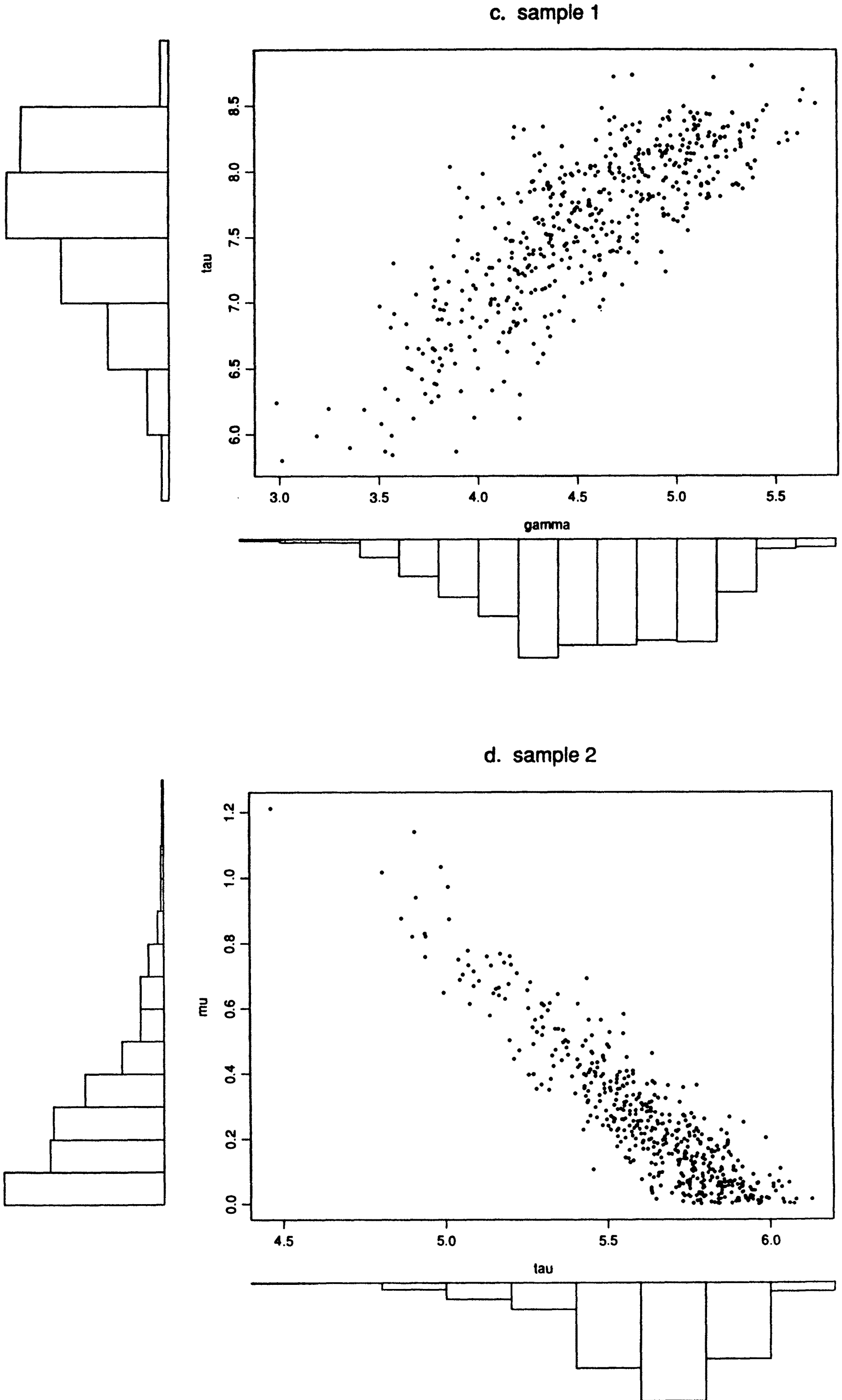


Figure 3. Continued.

are displayed in Figure 5. The points above the 45° line are those for which $D_{\text{pred}} > D$. For samples 1 and 2, we obtained P-values of .412 and .490, respectively, while for sample 3 we obtained .084. These quantitative results confirm our graphical conclusions from Figures 4a–c.

The above tests cast some doubt on the model for sample 3. But the nature of the Bayesian significance test is such that it tests the Bayesian model in its entirety, i.e., likelihood *and* prior.

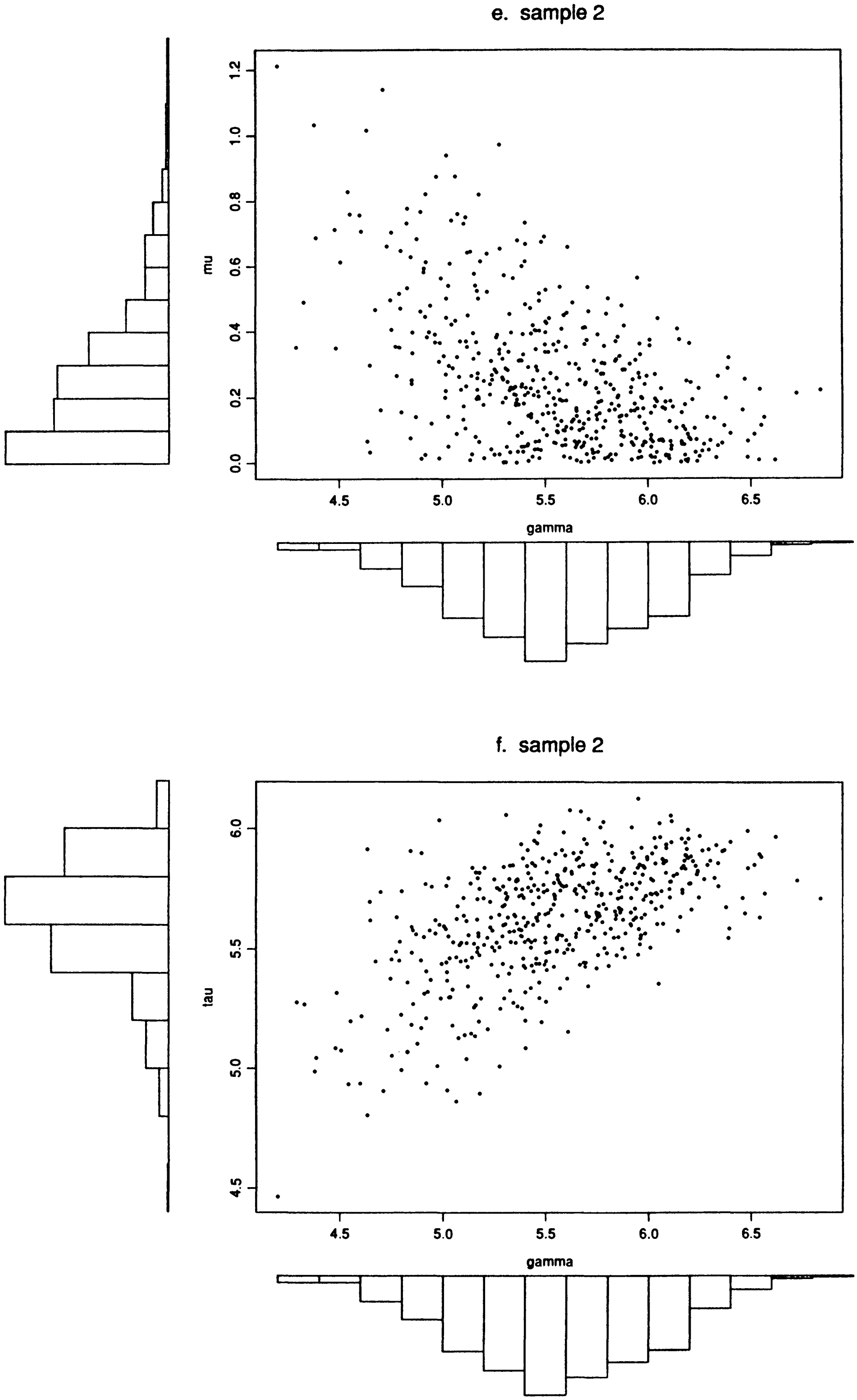
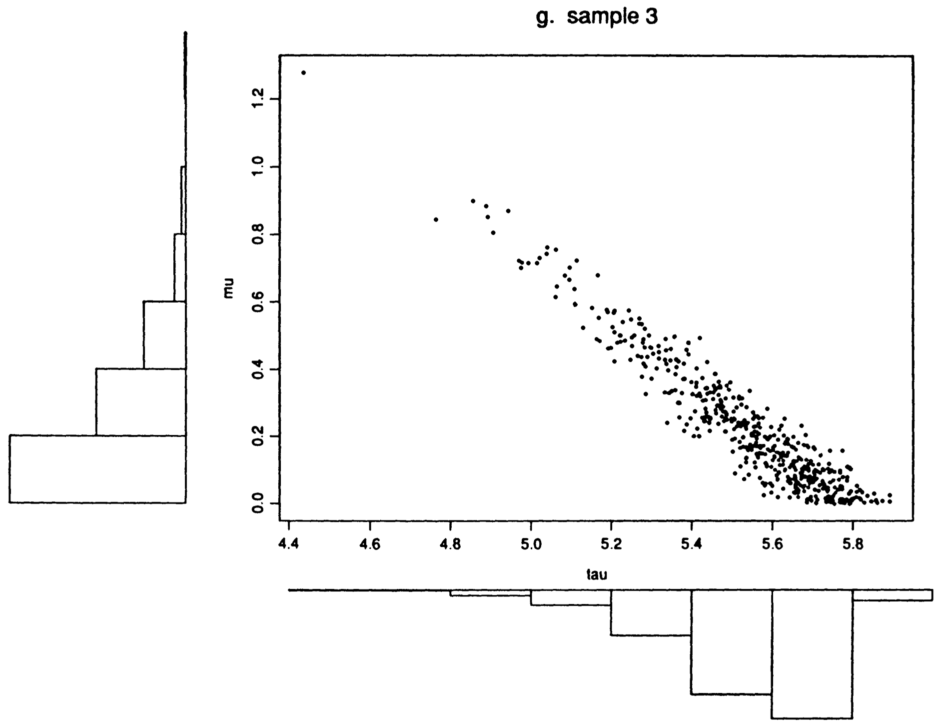


Figure 3. Continued.

Given the large sample size ($n = 200$), we were confident that the problem was with the likelihood and not the weak prior we specified. However, for the sake of completeness, we performed an analysis on the robustness of the posterior to the prior. The sampling–resampling method presented by Smith and Gelfand (1992) has made such investigations easy. As an illustration, we retained our original prior for μ but assigned Uniform(0, 20) distributions for both τ and γ since, as mentioned



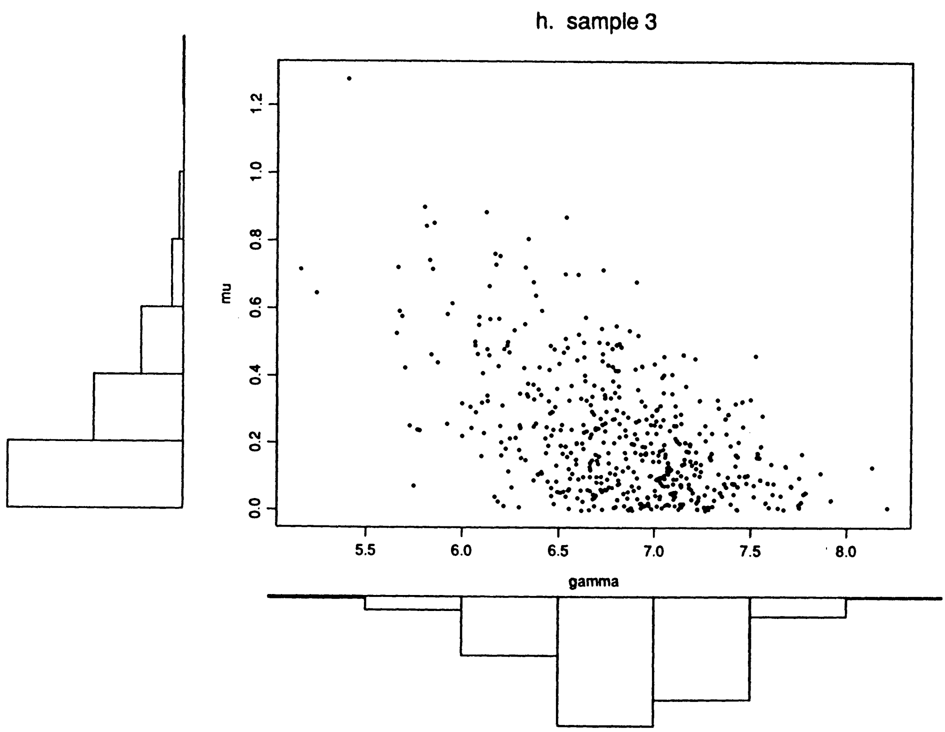


Figure 3. Continued.

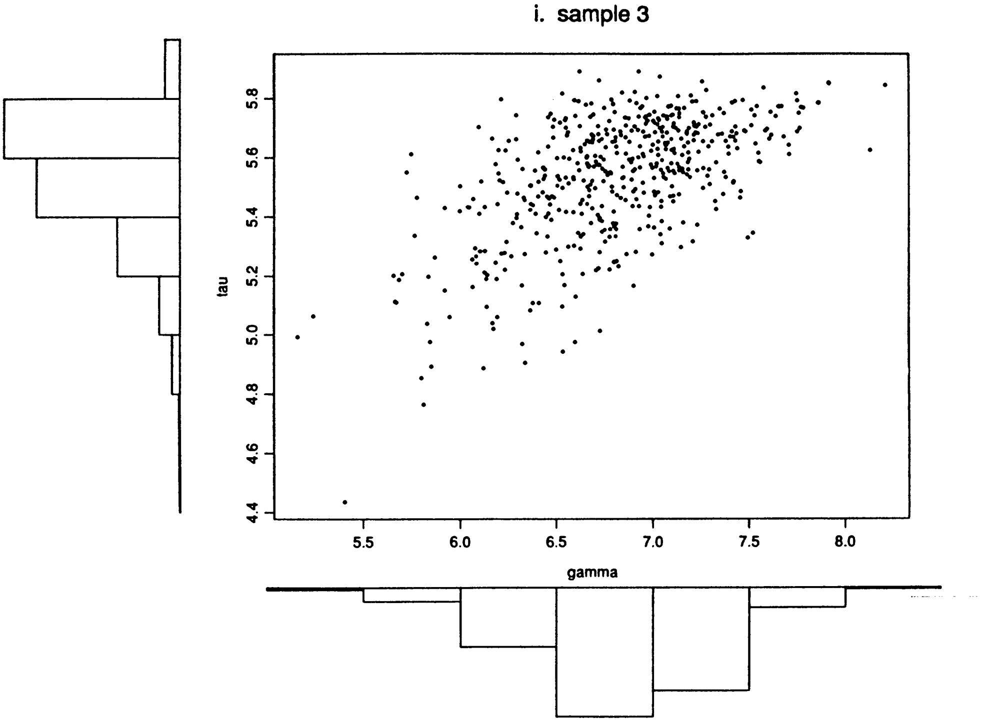


Figure 3. Continued.

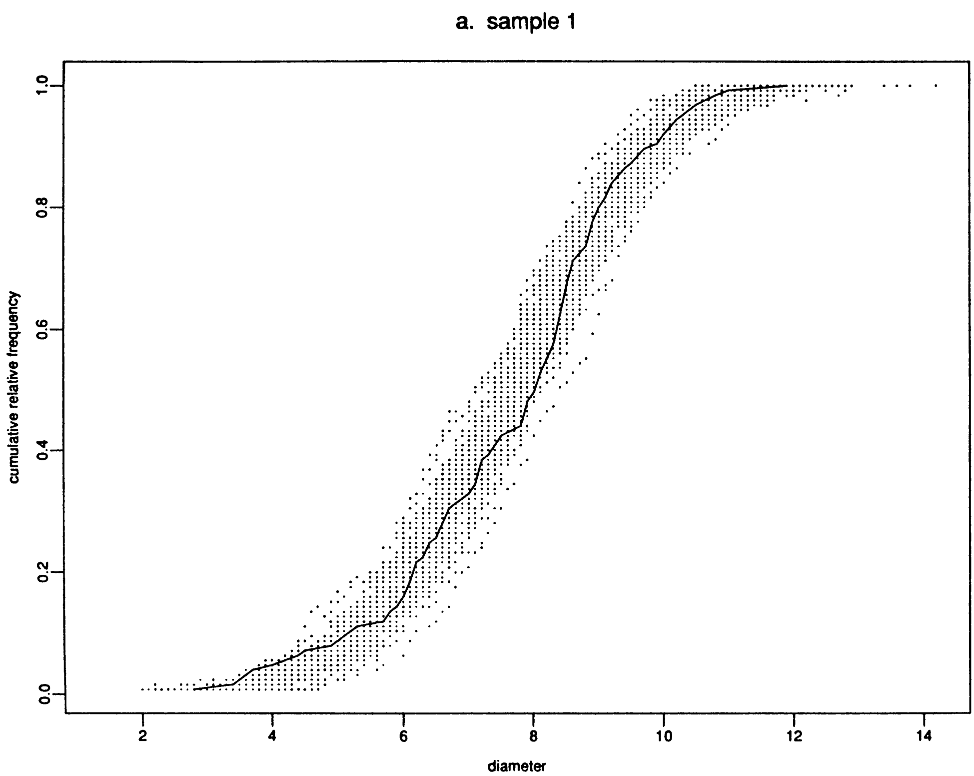
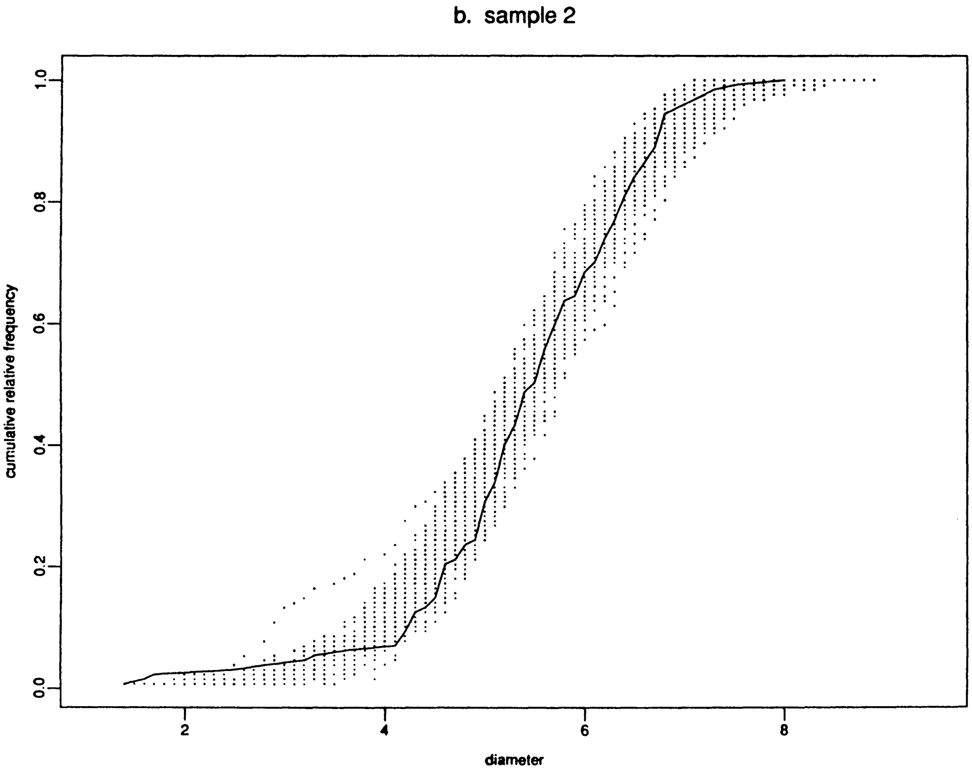


Figure 4. Predictive cdf's (unconnected dots) and edf (solid line) for three samples.



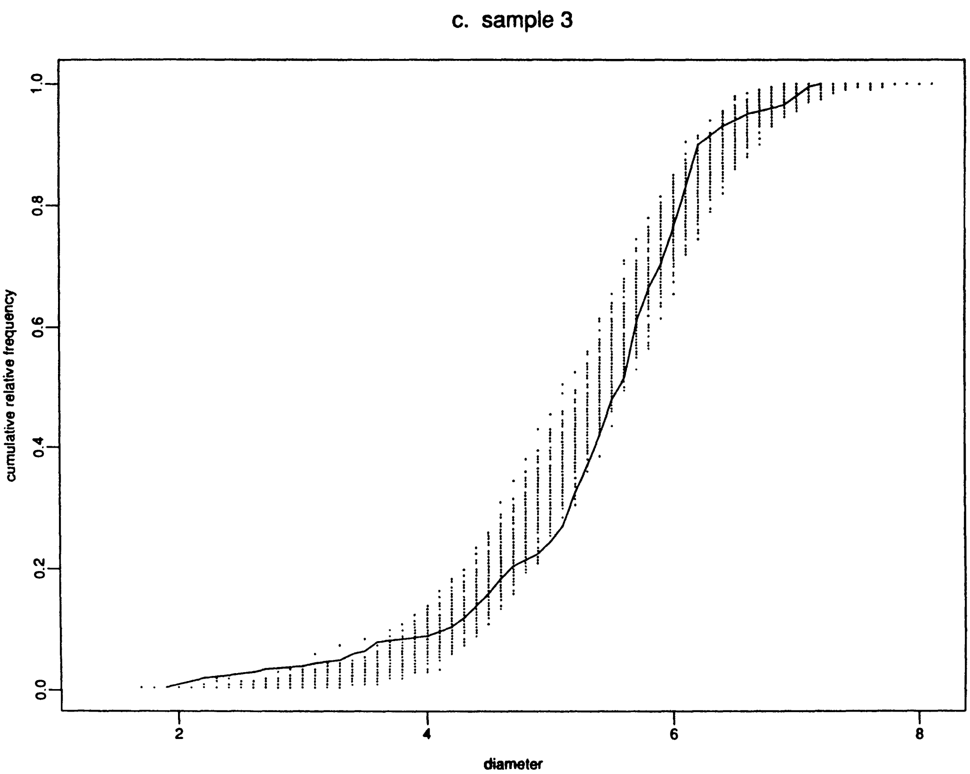


Figure 4. Continued.

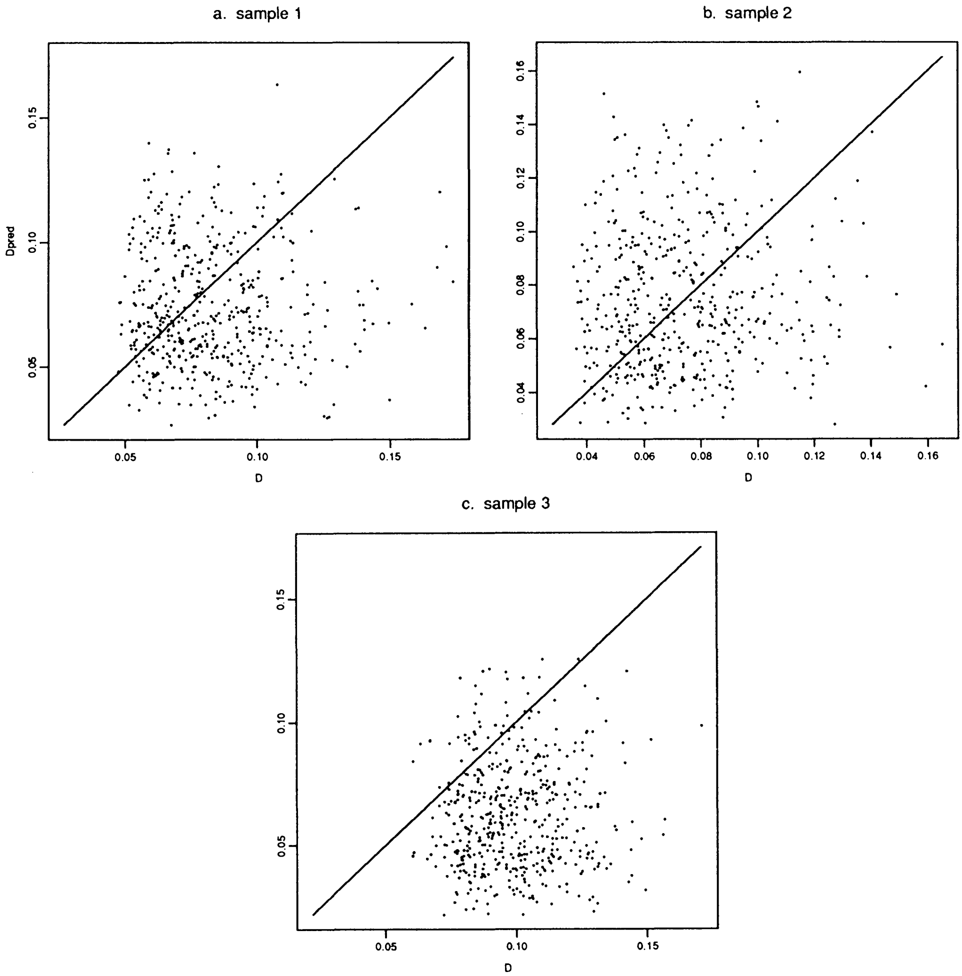


Figure 5. Scatterplots of D_{pred} versus D for three samples.

earlier, reported values for τ and γ are generally below about 9.0. We employed Smith and Gelfand's weighted bootstrap sampling–resampling method to generate a joint posterior sample according to this prior specification. We generated a posterior sample of size 500, and then performed the diagnostics described above on this sample. The results are shown in Figure 6. Figure 6a is nearly identical to Figure 4c, and the significance test gave a P-value of .080 for the resampled posterior. Hence we conclude that the prior is not the reason for the apparent lack of fit for sample 3, and that it is instead the likelihood—i.e., the Weibull density—that is questionable.

5. Conclusions

Our purpose here has been to demonstrate that easily implemented Bayesian procedures are available for more subtle inferential and diagnostic approaches to the three-parameter Weibull distribution than are typically found in the literature. We have shown that maximum likelihood estimation can be problematic when diameter distributions are modeled with the Weibull density, especially with data from young plantations or distributions with long left tails. In contrast, the Bayesian model is easy to fit and yields results in line with our prior expectations. In particular, widely used ad hoc rules, such as "*set* $\hat{\mu} = 0$ *whenever* $\hat{\mu} < 0$," are unnecessary with the Bayesian model.

If the Weibull density is to be used in a forestry yield model, then predictions would probably be made using the modal values of the joint posterior sample of the parameters. On the other hand, in the Bayesian approach, once having obtained samples from the full joint posterior, these can be straightforwardly used to great advantage for simulating any predictive distribution of interest.

Finally, we have demonstrated the use of some diagnostic methods for Bayesian models. Plots of the predictive cdf's versus the edf, and the Bayesian significance test indicated a lack of fit for one

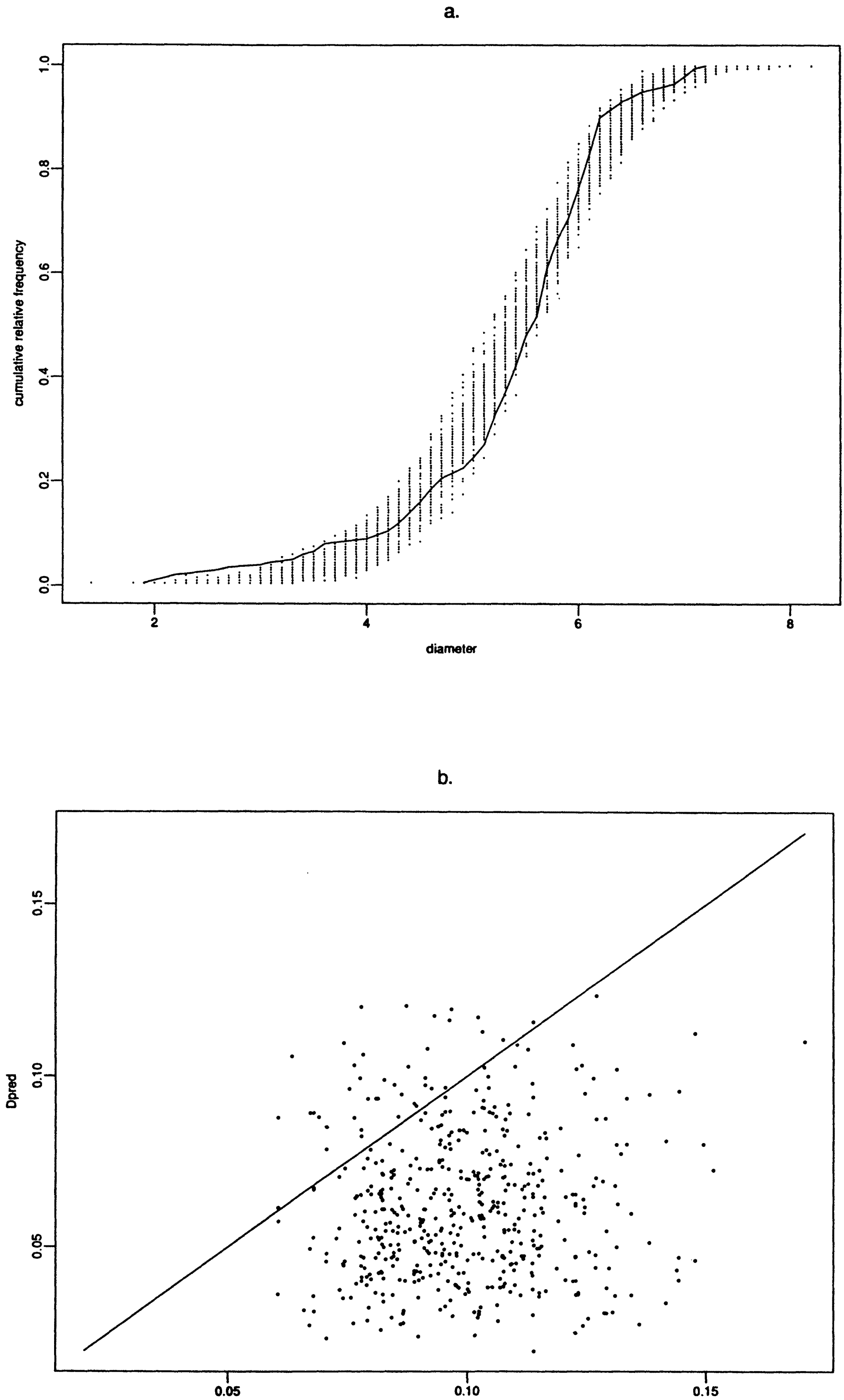


Figure 6. **a.** Predictive cdf (unconnected dots) and edf (solid line) for sample 3 with resampled posterior; **b.** Scatterplot of D_{pred} versus D for sample 3 with resampled posterior.

of the data sets. Further testing of the robustness of the prior via the sampling–resampling approach convinced us that it was indeed the likelihood and not the prior that was the source of the lack of fit.

It is difficult to say how important the lack of fit detected for sample 3 is. If we were building a yield model, with data from many plots, and this data set were the only offending one, then we would probably not be too concerned. On the other hand, if the plantation from which sample 3 arose were of special importance, or if many of the samples showed a similar lack of fit, then we would probably investigate other likelihoods. As usual, the practical significance of the statistical significance depends on the context.

Acknowledgements

This is New Jersey Agricultural Experiment Station Publication No. D-17503-2-94, supported by U.S. McIntire–Stennis Act funds and the U.S.D.A. Forest Service Southern Forest Experiment Station. The work reported herein was completed while the first author was an Academic Visitor at Imperial College, London, We are grateful to Dr S. K. Upadhyay for helpful discussions and permission to use his Gibbs sampler computer program.

Résumé

La loi de Weibull à trois paramètres est habituellement utilisée pour modéliser la distribution des diamètre des arbres dans les forêts. Nous montrons, à travers les profils de vraisemblance, que l'estimation du maximum de vraisemblance est souvent non appropriée pour les données provenant de jeunes arbres, car elle conduit à une estimation négative du paramètre de valeur centrale. Nous suggérons un modèle Bayesien que nous calculons, en utilisant l'échantillonnage de Gibbs, pour trois ensemble de données. Ce dernier modèle est facile à mettre en oeuvre et garantit une estimation positive du paramètre de valeur centrale. Nous détaillons les validations de quelques nouvelles configurations du modèle, montrant que le modèle Bayesien est approprié pour deux des exemples, alors qu'il est douteux pour le troisième. Une méthode d'échantillonnage–reéchantillonnage montre que le défaut de calcul du modèle pour le dernier ensemble de données est due à la vraisemblance, et non aux spécifications à priori.

References

Bailey, R. L., Burgan, T. M., and Jokela, E. J. (1989). Fertilized midrotation-aged slash pine plantations—stand structure and yield prediction models. *Southern Journal of Applied Forestry* **13,** 76–80.

Bailey, R. L. and Dell, T. R. (1973). Quantifying diameter distributions with the Weibull function. *Forest Science* **19,** 97–104.

Burk, T. E. and Newberry, J. D. (1984). A simple algorithm for moment-based recovery of Weibull distribution parameters. *Forest Science* **30,** 329–332.

Crowder, M. J., Kimber, A. C., Smith, R. L., and Sweeting, T. J. (1991). *Statistical Analysis of Reliability Data*. London: Chapman and Hall.

Dagpunar, J. (1988). *Principles of Random Variate Generation*. London: Oxford University Press.

Ek, A. R., Issos, J. N., and Bailey, R. L. (1975). Solving for Weibull diameter distribution parameters to obtain specified mean diameters. *Forest Science* **21,** 290–292.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* **85,** 972–985.

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

Gelman, A. and Rubin, D. (1992). Inferences from iterative simulation using multiple sequences (with Discussion). *Statistical Science* **7,** 457–511.

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

Geyer, C. (1992). Practical Markov chain Monte Carlo (with Discussion). *Statistical Science* **7,** 473–511.

Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics* **41,** 337–348.

Green, E. J. and Strawderman, W. E. (1992). A comparison of hierarchical Bayes and empirical Bayes methods. *Forest Science* **38,** 350–366.

Hills, S. E. and Smith, A. F. M. (1992). Parameterization issues in Bayesian inference. In *Bayesian Statistics* **4,** J. M. Bernardo et al. (eds). London: Oxford University Press.

Jeffreys, H. (1961). *Theory of Probability*, 3rd edition. London: Oxford University Press.

Krug, A. G., Nordheim, E. V., and Giese, R. L. (1984). Determining initial values for parameters of a Weibull model: A case study. *Forest Science* **30,** 573–581.

Lange, N., Carlin, B. P., and Gelfand, A. E. (1992). Hierarchical Bayes models for the progression of HIV infection using longitudinal CD4 T-cell numbers (with Discussion). *Journal of the American Statistical Association* **87,** 615–632.

Lenhart, J. D. (1988). Diameter-distribution yield-prediction system for unthinned loblolly and slash pine plantations on non-old-fields in east Texas. *Southern Journal of Applied Forestry* **12,** 239–242.

Matney, T. G., Ledbetter, J. R., and Sullivan, A. D. (1987). Diameter distribution yield systems for unthinned cutover site-prepared slash pine plantations in southern Mississippi. *Southern Journal of Applied Forestry* **11,** 32–36.

Raftery, A. E. and Lewis, S. (1992). How many iterations in the Gibbs sampler? In *Bayesian Statistics* **4,** J. M. Bernardo et al. (eds). London: Oxford University Press.

Reynolds, M. R., Jr., Burk, T. E., and Huang, W. C. (1988). Goodness-of-fit tests and model selection procedures for diameter distribution models. *Forest Science* **34,** 373–399.

Ripley, B. (1987). *Stochastic Simulation*. New York: Wiley.

Smith, A. F. M. and Gelfand, A. E. (1992). Bayesian statistics without tears: A sampling–resampling approach. *The American Statistician* **46,** 84–88.

Smith, A. F. M. and 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,** 3–23.

Smith, R. L. and Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimation for the three-parameter Weibull distribution. *Applied Statistics* **36,** 358–369.

Wakefield, J. C. (1993). Discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society, Series B* **55,** 56–57.

Wakefield, J. C., Smith, A. F. M., Racine-Poon, A., and Gelfand, A. E. (1994). Bayesian analysis of linear and nonlinear population models using the Gibbs sampler. *Applied Statistics* **43,** 201–221.

Received January 1993; *revised May* 1993; *accepted June* 1993.