Estimation and Applications of Size-biased Distributions in Forestry

Jeffrey H. Gove[1]
USDA Forest Service, Northeastern Research Station
P.O. Box 640, Durham, NH USA 03824

Abstract

Size-biased distributions arise naturally in several contexts in forestry and ecology. Simple power relationships (e.g., basal area and diameter at breast height) between variables are one such area of interest arising from a modeling perspective. Another, probability proportional to size sampling (PPS), is found in the most widely used methods for sampling standing or dead and downed material in the forest. Often, it is desirable or necessary to estimate a parametric probability density model based on size-biased data. Traditional equal probability methods may not be appropriate, or may be less efficient in such circumstances and estimation is better conducted utilizing size-biased theory. This paper surveys some of the possible uses of size-biased distribution theory in forestry and related fields.

Introduction

Size-biased distributions are a special case of the more general form known as weighted distributions. First introduced by Fisher (1934) to model ascertainment bias, weighted distributions were later formalized in a unifying theory by Rao (1965). Such distributions arise naturally in practice when observations from a sample are recorded with unequal probability, such as from probability proportional to size (PPS) designs. Briefly, if the random variable X has distribution $f(x;\boldsymbol{\theta})$, with unknown parameters $\boldsymbol{\theta}$, then the corresponding weighted distribution is of the form

$$f^w(x;\boldsymbol{\theta}) = \frac{w(x)f(x;\boldsymbol{\theta})}{\mathrm{E}\,[w(x)]}$$

where $w(x)$ is a nonnegative weight function such that $\mathrm{E}\,[w(x)]$ exists.

A special case of interest arises when the weight function is of the form $w(x) = x^\alpha$. Such distributions are known as size-biased distributions of order α and are written as (Patil and Ord, 1976; Patil, 1981; Mahfoud and Patil, 1982)

$$f^*_\alpha(x;\boldsymbol{\theta}) = \frac{x^\alpha f(x;\boldsymbol{\theta})}{\mu'_\alpha} \tag{1}$$

where $\mu'_\alpha = \int x^\alpha f(x;\boldsymbol{\theta})\,dx$ is the αth raw moment of $f(x;\boldsymbol{\theta})$. Denote X the original, or equal probability, random variable, and $X^*_\alpha \sim f^*_\alpha(x;\boldsymbol{\theta})$ the size-biased random variable. The most common cases of size-biased

[1]email: *jhgove@christa.unh.edu* or *jgove@fs.fed.us*

distributions occur when $\alpha = 1$ or 2; in the context of sampling, these special cases may be termed length- and area-biased, respectively.

Weighted distributions have numerous applications in forestry and ecology. Warren (1975) was the first to apply them in connection with sampling wood cells. Van Deusen (1986) arrived at size-biased distribution theory independently and applied it to fitting distributions of diameter at breast height (DBH) data arising from horizontal point sampling (HPS) (Grosenbaugh, 1958) inventories. Subsequently, Lappi and Bailey (1987) used weighted distributions to analyze HPS diameter increment data. More recently, weighted distributions were used by Magnussen et al. (1999) to recover the distribution of canopy heights from airborne laser scanner measurements. In ecology, Dennis and Patil (1984) use stochastic differential equations to arrive at a weighted gamma distribution as the stationary probability density function (PDF) for a stochastic population model with predation effects. In fisheries, Taillie et al. (1995) modeled populations of fish stocks using weighted distributions. In these last two examples, weighted distributions were not directly tied to the underlying sample selection method, but were simply convenient models for the observed data. Recognizing the fact that weighted distributions may be applied as convenient PDF models, Gove and Patil (1998) developed a compatible theory, unifying the DBH-frequency and basal area-DBH distributions based on the quadratic relationship between diameter and basal area. Lastly, Gove (2000) extended the work of Van Deusen (1986) by providing simulation experiments and guidelines for fitting size-biased distributions to HPS data.

This purpose of this paper is to review some of the more recent results on size-biased distributions pertaining to parameter estimation in forestry with special emphasis on the Weibull family. In addition, some new results and avenues for possible future research will be presented. Finally, a new computer program with graphical user interface (GUI) developed by the author for fitting size-biased Weibull distributions will be briefly discussed.

Size-biased Weibull distributions

Weibull distributions have found wide-spread use in forestry for modeling since they were first introduced by Bailey and Dell (1973). The two-and three-parameter Weibull PDFs are given as

$$f(x;\boldsymbol{\theta}) = \left(\frac{\gamma}{\beta}\right)\left(\frac{x}{\beta}\right)^{\gamma-1} e^{-(x/\beta)^{\gamma}} \qquad x > 0$$

$$f(x;\boldsymbol{\theta}) = \left(\frac{\gamma}{\beta}\right)\left(\frac{x-\xi}{\beta}\right)^{\gamma-1} e^{-((x-\xi)/\beta)^{\gamma}} \qquad x > \xi$$

with $\boldsymbol{\theta} = (\gamma, \beta)'$ and $\boldsymbol{\theta} = (\gamma, \beta, \xi)'$, respectively. The unknown parameters $\gamma > 0$, $\beta > 0$ and $\xi > 0$ are the shape, scale and location parameters to be estimated for a given sample of data.

These PDFs can be easily converted to their size-biased counterparts using (1); *viz.*,

$$
\begin{aligned}
f_\alpha^*(x;\boldsymbol{\theta}) &= \frac{x^\alpha}{\mu'_\alpha}\left(\frac{\gamma}{\beta}\right)\left(\frac{x}{\beta}\right)^{\gamma-1} e^{-(x/\beta)^\gamma} && x > 0 \\
f_\alpha^*(x;\boldsymbol{\theta}) &= \frac{x^\alpha}{\mu'_\alpha}\left(\frac{\gamma}{\beta}\right)\left(\frac{x-\xi}{\beta}\right)^{\gamma-1} e^{-((x-\xi)/\beta)^\gamma} && x > \xi
\end{aligned}
$$

for the two- and three-parameter versions, respectively, with the same restrictions on the parameters as for the equal probability PDFs. Gove and Patil (1998) have also shown that the size-biased two-parameter Weibull can be transformed, through change-of-variables techniques, to the standard gamma distribution. Such a transformation may be advantageous for simulations studies. For example, Gove (2000) used the standard gamma to draw probability-weighted samples to simulate the HPS tally distribution.

Because of their popularity in modeling the traditional DBH-frequency distribution, both the two- and three-parameter size-biased Weibull PDFs are appropriate as candidate probability models in all of the applications presented in this paper.

Size-biased Weibulls: moment estimation

Size-biased two-parameter Weibull moment estimators

The development of moment estimators for the size-biased two-parameter Weibull distribution is given in Gove (2003a). There, a modified moment estimation scheme along the lines of Cohen (1965), using the coefficient of variation, is presented. Let $\tilde{\gamma}$ and $\tilde{\beta}$ represent the moment estimates for the shape and scale parameters, respectively; then the moment equations are

$$
CV = \Gamma_\alpha \Gamma_{\alpha+1}^{-1} \sqrt{\frac{\Gamma_{\alpha+2}}{\Gamma_\alpha} - \frac{\Gamma_{\alpha+1}^2}{\Gamma_\alpha^2}} \tag{2}
$$

$$
\tilde{\beta} = \frac{\bar{x}\,\tilde{\Gamma}_\alpha}{\tilde{\Gamma}_{\alpha+1}} \tag{3}
$$

where $\bar{x}$ and CV are the sample mean and coefficient of variation, respectively, with $\Gamma_\alpha = \Gamma(\alpha/\gamma + 1)$, $\tilde{\Gamma}_\alpha = \Gamma(\alpha/\tilde{\gamma} + 1)$ and $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x}\,dx$, $k > 0$, the gamma function. The first equation (2) is solved iteratively for the shape parameter, then the scale parameter can be found directly by substitution into (3).

Size-biased three-parameter Weibull moment estimators

Unfortunately, the moment equations for the size-biased three-parameter Weibull are not easily couched in a modified scheme like that for the two-parameter where the coefficient of variation can be used. Thus, the moment equations for the first three raw moments are used; these moments can be built-up from the moments of the equal probability three-parameter Weibull (Gove, 2003a). Let $\mu^{*\prime}_{\alpha,\zeta} = \int d^{\zeta} f^{*}_{\alpha}(x;\boldsymbol{\theta})\,dx$ denote the ζth raw moment of the size-biased three-parameter Weibull distribution of order α. Then, it is straightforward to show that $\mu^{*\prime}_{\alpha,\zeta} = \frac{\mu'_{\alpha+\zeta}}{\mu'_{\alpha}}$. Now, since $\alpha = 1$ or 2 for our most common forestry applications, and $\zeta = 1, \ldots, 3$ for the first three raw moments, it is easy to see from the numerator of $\mu^{*\prime}_{\alpha,\zeta}$ that the first five raw moments of the three-parameter Weibull distribution are required for the estimating equations. The moments for the three-parameter Weibull are of the form

$$\mu'_{\alpha} = \beta^{\alpha}\Gamma_{\alpha} + \binom{\alpha}{1}\beta^{\alpha-1}\Gamma_{\alpha-1}\xi + \binom{\alpha}{2}\beta^{\alpha-2}\Gamma_{\alpha-2}\xi^{2} + \cdots + \xi^{\alpha} \tag{4}$$

where the coefficients $\binom{\alpha}{i}, i = 1, \ldots, \alpha$ follow Pascal's triangle. Thus, for example, the second raw moment from a length-biased three-parameter Weibull, $\mu^{*\prime}_{1,2}$ is

$$\frac{\mu'_3}{\mu'_1} = \frac{\beta^3\Gamma_3 + 3\beta^2\Gamma_2\xi + 3\beta\Gamma_1\xi^2 + \xi^3}{\beta\Gamma_1 + \xi}$$

It should be clear that the moment equations for the length- and area-biased versions differ. For comparison, the second raw moment from an area-biased three-parameter Weibull is given as $\mu^{*\prime}_{2,2}$, and is therefore more complicated; *viz.*,

$$\frac{\mu'_4}{\mu'_2} = \frac{\beta^4\Gamma_4 + 4\beta^3\Gamma_3\xi + 6\beta^2\Gamma_2\xi^2 + 4\beta\Gamma_1\xi^3 + \xi^4}{\beta^2\Gamma_2 + 2\beta\Gamma_1\xi + \xi^2}$$

The first three moment equations are set equal to the first three sample moments and solved simultaneously for the estimates $\tilde{\gamma}, \tilde{\beta}, \tilde{\xi}$. Further details are given in Gove (2002) and Gove (2003a).

Size-biased Weibulls: maximum likelihood estimation

The maximum likelihood estimators (MLEs) for size-biased Weibulls can be found by building-up from the equal probability likelihood, just as in the case of the three-parameter moment estimators in the previous section. The equal probability three-parameter Weibull log-likelihood is

$$\ln\mathcal{L} = n\ln\left(\frac{\gamma}{\beta^{\gamma}}\right) + (\gamma-1)\sum_{i=1}^{n}\ln(x_i-\xi) - \frac{1}{\beta^{\gamma}}\sum_{i=1}^{n}(x_i-\xi)^{\gamma}$$

and the two-parameter log-likelihood follows directly by setting $\xi = 0$.

The size-biased form was first given by Van Deusen (1986), where he noted that it was composed of the equal probability log-likelihood plus a constant and a correction term. He also noted that the purpose of the correction was to account for the fact that the observations are drawn with unequal probability. The general form of the size-biased log-likelihood is given as

$$\ln \mathcal{L}^* = \ln \mathcal{L} + \alpha \sum_{i=1}^{n} \ln x_i - n \ln \mu'_\alpha$$

where the second term is constant, depending only on the data, and thus may be dropped if desired.

In addition, the gradient vector and Hessian matrix of first and second-order partial derivatives are also of the same form (Gove, 2003a). For example, the gradient equations for the size-biased three-parameter Weibull follow the form

$$\begin{aligned}
\frac{\partial \ln \mathcal{L}^*}{\partial \gamma} &= \frac{\partial \ln \mathcal{L}}{\partial \gamma} - n\rho_\gamma(\alpha) \\
\frac{\partial \ln \mathcal{L}^*}{\partial \beta} &= \frac{\partial \ln \mathcal{L}}{\partial \beta} - n\rho_\beta(\alpha) \\
\frac{\partial \ln \mathcal{L}^*}{\partial \xi} &= \frac{\partial \ln \mathcal{L}}{\partial \xi} - n\rho_\xi(\alpha)
\end{aligned}$$

Notice that the correction term ($n\rho_{\boldsymbol{\theta}_i}(\alpha)$) depends on the size-biased order α. Thus, there are unique corrections associated with length- and area-biased log-likelihoods. The Hessian matrix follows the same pattern, being composed of the equal probability and correction components. Detailed equations for the three-parameter gradient and Hessian are presented in Gove (2003a). In the two-parameter size-biased Weibull, the equations are much simpler due to the simpler nature of the raw moment μ'_α in that distribution. The gradient equations for the two-parameter case are given in Gove (2000).

The basal area-size distribution

As mentioned earlier, the basal area-size distribution (BASD) is the size-biased distribution of order $\alpha = 2$ of the traditional DBH-frequency distribution (Gove and Patil, 1998). The relation can be easily shown algebraically and arises, not from sampling theory, but purely from the quadratic relationship between DBH and basal area. If the random variable X is tree diameter, then $X \sim f(x; \boldsymbol{\theta})$ is the DBH-frequency distribution. From it, we normally calculate the number of trees in the ith diameter class (N_i), once the parameters $\boldsymbol{\theta}$ have been estimated from sample data; *viz.*

$$N_i = N \int_{x_{l_i}}^{x_{u_i}} f(x; \boldsymbol{\theta})\, dx \tag{5}$$

where x_{l_i} and x_{u_i} are the lower and upper diameter class limits, respectively and N is the total number of trees per hectare.

The BASD comes about by redistributing the probability mass in terms of basal area, rather than tree frequency. The random variable in both cases is still DBH. The BASD can then be used to calculate the basal area (B_i) in the ith DBH class as

$$B_i = B \int_{x_{l_i}}^{x_{u_i}} f_2^*(x; \boldsymbol{\theta})\, dx$$

where B is the stand basal area per hectare. Thus, $X_2^* \sim f_2^*(x; \boldsymbol{\theta})$.

Gove and Patil (1998) presented several examples of stands fitted with a parameter recovery model, all with the same basal area and number of trees, but spanning a wide range of the two-parameter Weibull parameter space. As an example, the stand in their Figure 1d has been re-fitted with a three-parameter Weibull model and is presented in Figure 1. This figure shows the empirical histogram for the DBH-frequency distribution along with the Weibull curve fitted by ML. Also shown is the corresponding BASD curve, which shares the same estimated parameter vector $\hat{\boldsymbol{\theta}}$ from ML.

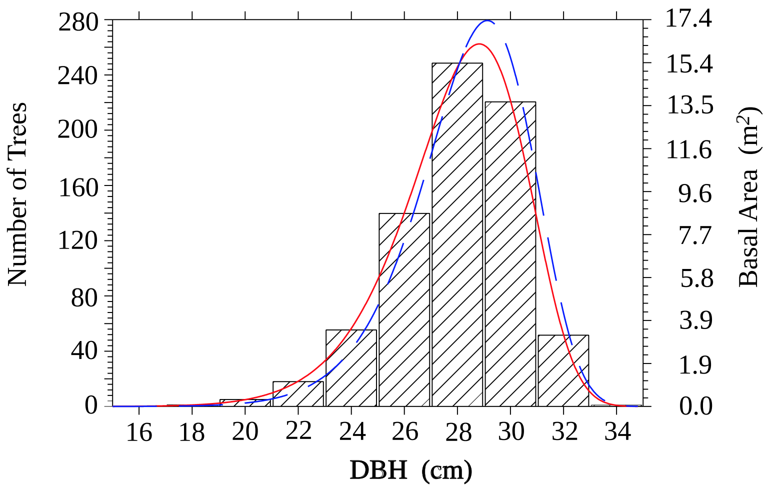

Figure 1: Example diameter distribution (shaded) from Gove and Patil (1998) with $B = 45.9$ and $N = 741$ showing the estimated DBH-frequency distribution (solid) and associated BASD (dashed) for a three-parameter Weibull with MLEs: $\hat{\gamma} = 11.23$, $\hat{\beta} = 23.44$, $\hat{\xi} = 5.57$.

Estimation of Weibull parameters under size-biased sampling

Arguably, the two most useful forms of size-biased distributions arising in forestry are the length- and area-biased models. Length-biased data arises from line intersect samples (LIS) (Kaiser, 1983), horizontal and vertical line samples (HLS, VLS) (Grosenbaugh, 1958), and from transect relascope sampling (TRS) (Ståhl, 1998). Area-biased data arises naturally from HPS, and vertical point sampling (VPS) (Grosenbaugh, 1958), and from point relascope sampling (PRS) (Gove et al., 1999) for down coarse woody debris. In this section these links are explored in more detail with special emphasis on the distribution of HPS tally tree diameters.

Because of the intrinsic link between basal area and HPS, it is not surprising that the distribution of tally diameters from a HPS turns out once again to be the size-biased distribution of order $\alpha = 2$ (Van Deusen, 1986; Gove, 2000). Thus, if the underlying population diameter distribution for a given stand is $f(x;\boldsymbol{\theta})$, then the corresponding HPS tally distribution is given by $f_2^*(x;\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a shared parameter set. Having sampled from $f_2^*(x;\boldsymbol{\theta})$ using a prism or suitable angle gauge with HPS, we next must estimate $\boldsymbol{\theta}$, usually by ML. In the following sections some strategies for estimation are discussed with regard to this problem.

Fitting single horizontal point samples

Van Deusen (1986) first discussed fitting Weibull distributions to diameter data arising from single horizontal point samples. The most common reason for doing this being the subsequent fitting of parameter prediction models (Hyink and Moser, 1983). Later, Gove (2000) used simulations to address in more detail the possible problems with parameter estimation, using two-parameter Weibulls for illustration. The main results of the latter study are discussed in this section.

Briefly, it is possible to estimate $\boldsymbol{\theta}$ either by fitting a Weibull to the estimated stand table (number of trees per hectare by DBH) diameters from a single HPS, or by fitting the area-biased Weibull directly to the tally diameters. However, in theory, $\boldsymbol{\theta}$ is supposed to be a shared parameter set between $f(x;\boldsymbol{\theta})$ and $f_2^*(x;\boldsymbol{\theta})$. A problem arises because one can fit both distributions to their respective data for any given HPS and, in so doing, two different estimates of $\boldsymbol{\theta}$ normally result in the process. Then the question becomes, which estimate is the best? This question does not arise when fitting distributions to diameters sampled on fixed area plots, because in either instance we are estimating $f(x;\boldsymbol{\theta})$ (Gove, 2000).

The simulations presented were extensive and will not be discussed in detail here. However, they were designed to assess the effects of both expected sample size (in terms of number of trees tallied) per point, and the shape of the population distribution $f(x;\boldsymbol{\theta})$ on estimation. The key findings were as follows. First, as the sample size per point increases, both parameter estimates tend to converge to the population values. However, the rate at which they do so depends in large part on the shape of the underlying population of diameters. In the case of fairly symmetric population distributions, both parameter estimates converged at the same rate and had very similar root-mean squared errors (RMSEs). However, as the population diameter distribution tended more towards a reverse J-shape, associated with typical uneven-aged stands, the parameter estimates from the size-biased distribution fit both converged more quickly and had lower RMSE, often by more than a half.

The reasons for the results are two-fold. First, because the size-biased form is theoretically linked to the underlying sampling mechanism, its shape more nearly parallels that of the population distribution of HPS diameters and is therefore estimated more efficiently. This is particularly true, as illustrated in Figure 4a of Gove (2000), when the population diameter distribution is reverse J-shaped. As the population distribution of diameters become more symmetric, the shapes of $f(x;\boldsymbol{\theta})$ and $f_2^*(x;\boldsymbol{\theta})$ tend to be more alike and estimation is therefore essentially equivalent for either density. Second, in the reverse J-shaped population, sampling with probability proportional to basal area is akin to sampling for rare events in terms of frequency. The vast majority of probability density for the associated tally distribution is confined to diameters of essentially merchantable size. Therefore, it is very difficult to realize a large enough sample of smaller diameter trees on any one point, to actually shift the estimated stand table from unimodel to reverse J-shaped. For example, the result of $m = 1000$ simulations from a reverse J-shaped distribution with population shape parameter $\gamma = 1.0$, resulted in an estimated stand table shape parameter of $\hat{\gamma} = 1.54$ with $N^* = 40$ trees per point sampled. In contrast, the estimate for the size-biased shape parameter from the tally data for the same simulations was 1.07, with RMSE equal to $\frac{1}{4}$ that of the stand table estimate for the shape parameter.

The most important conclusion that should be kept in mind from this study, is that concerning the overall purpose of the inventory. Horizontal point sample inventories are a rich reservoir of data for estimating forest characteristics. However, the normal recommendation of choosing an angle that selects 5 to 12 trees per point on average for estimating stand data (Avery and Burkhart, 1994, p. 218), generally will not suffice for parameter estimation of assumed diameter distributions. Therefore, the goals of parameter estimation and inventory may conflict and it is possible that, depending on the shape of the population diameter distribution, alternative inventory protocols may be required.

Fitting with multiple points

Fitting diameter distributions to a single HPS for use with parameter prediction model construction is undoubtedly a rather infrequent use of such data. It is probably more likely the case that one would be interested in fitting diameter distributions to sample data arising from more than one HPS point, say, for example, to a stand diameter distribution taken over n sample points. In this case, the question posed in the previous study are still valid. However, the support for parameter estimation naturally increases with the increased sample size and one would expect that the ML estimates would continue to converge in both the stand and tally estimates to the respective population values. The problem can be viewed from two different perspectives based on the degree of homogeneity of the target population diameter distribution.

Homogeneous stands

In this case, one would envision that the diameter distribution from one point to the next in a HPS inventory is relatively homogeneous within the population of interest. Thus, for parameter estimation purposes, the stand table can be computed directly from the sample of n points to estimate $f(x;\boldsymbol{\theta})$. Similarly, the tally from all n points can simply be pooled to estimate $f_2^*(x;\boldsymbol{\theta})$. Furthermore, let $\hat{\boldsymbol{\theta}} = (\hat{\gamma}, \hat{\beta})$ and $\hat{\boldsymbol{\theta}}^* = (\hat{\gamma}^*, \hat{\beta}^*)$ be the MLEs for $f(x;\boldsymbol{\theta})$ and $f_2^*(x;\boldsymbol{\theta})$, respectively.

Two sets of simulations were conducted to extend the previous study to the multiple point case. The two populations chosen were those that showed the poorest convergence in the single HPS estimates: the reverse J-shaped and mild positively skewed populations. The expected number of trees sampled per point was fixed at $N^* = 10$, and the sample sizes ranged from $n = 5$ to 40 points for the simulations.

The results of the simulations are presented in Table 1. These results clearly show that in both cases, as the sample size increases, the parameter estimates converge to the population values more rapidly for the tally distribution. Not only is the overall bias less, but the RMSE is also significantly reduced. This is particularly true for the reverse J-shaped population, but also still holds rather convincingly for the mildly skewed population.

Table 1: Simulation results for $m = 250$ replications of $N^* = 10$ trees per point on n multiple HPS points drawn from a two-parameter Weibull population of tree diameters with $\boldsymbol{\theta} = (\gamma, \beta)$.

		Average				%RMSE			
$\boldsymbol{\theta}$	n	$\hat{\gamma}^*$	$\hat{\beta}^*$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}^*$	$\hat{\beta}^*$	$\hat{\gamma}$	$\hat{\beta}$
(1,8)	5	1.05	8.50	1.48	11.16	19.1	34.2	75.0	71.7
	10	1.02	8.20	1.39	10.61	12.7	24.1	60.2	58.3
	20	1.02	8.23	1.32	10.20	8.9	16.6	51.1	49.1
	30	1.0	7.99	1.25	9.55	6.1	12.5	41.3	39.7
	40	1.0	7.95	1.22	9.21	5.6	10.9	38.7	36.9
(2,15)	5	2.08	15.18	2.32	15.57	16.4	11.4	32.6	16.0
	10	2.04	15.04	2.18	15.25	11.8	8.6	23.2	12.0
	20	2.02	15.07	2.15	15.32	7.3	5.7	19.4	10.5
	30	2.0	14.97	2.08	15.06	6.1	4.6	15.7	8.9
	40	2.01	15.04	2.13	15.24	5.2	3.9	16.8	10.4

Because the reverse J-shaped case is closely linked to uneven-aged management, which seems to be gaining in popularity in the U.S., it is of interest to look at this problematic case a little more closely. The results of the simulations are presented graphically in Figure 2. In both cases, the population line is shown as solid, and the average density (dashed) lines generally approach it as the number of points increases. It is quite apparent from these graphs that the estimated densities for the stand table data are never quite able to estimate the true reverse J-shape. On the other hand, it takes in the neighborhood of 25–30 HPS points to arrive at the correct estimates when using the tally data and $f_2^*(x;\boldsymbol{\theta})$ in such stands.

These simulations mirror the trends in the single HPS case exactly, but because of the increase sample size, show that convergence is better in the multiple point scenario. The conclusions to be drawn then also parallel the single point case: When sampling from stand conditions that approach that of the classic reverse J-shaped distribution, or show some degree of positive skewness, parameter estimation should be undertaken using the size-biased likelihood approach.

Heterogeneous stands

Consider a stand (or larger area) where the diameter distribution varies, possibly considerably, throughout, but where it is still desired to estimate $f(x;\boldsymbol{\theta})$. In such cases it may or may not be feasible to stratify.

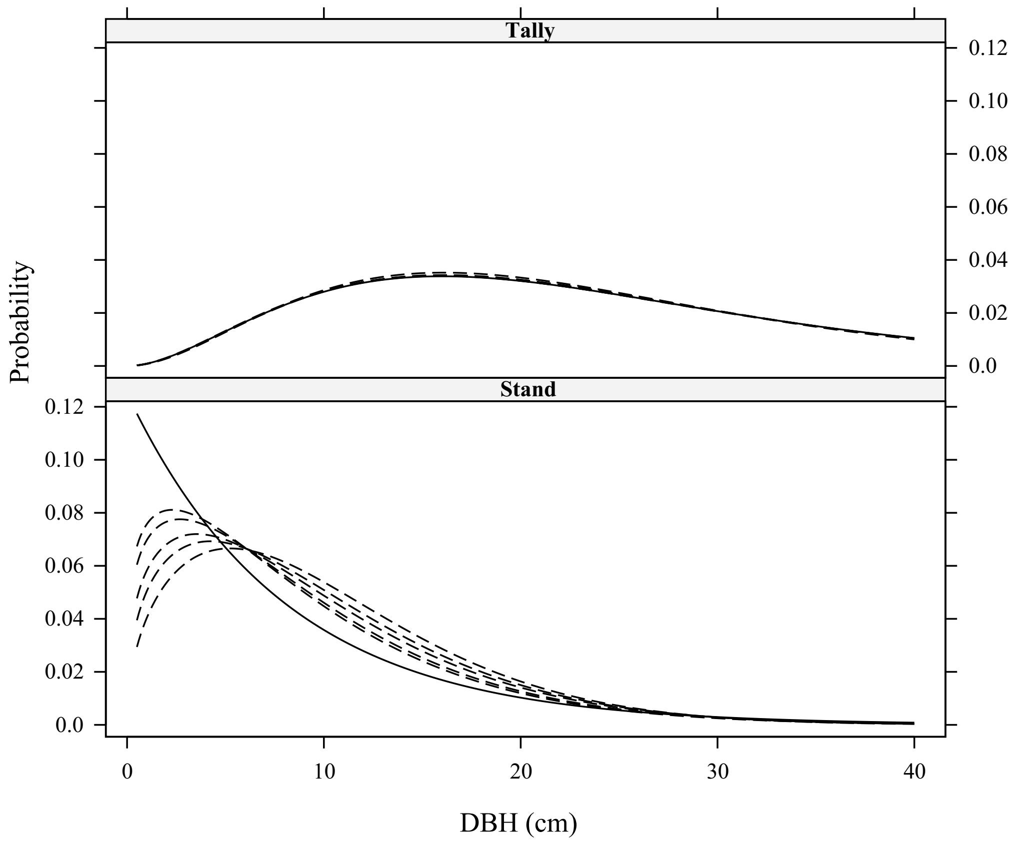

Figure 2: Simulated average distribution results for the homogeneous reverse J-shaped population with the tally densities (top) and stand table densities (bottom); the mean densities (dashed) converge to the population curve (solid) with increasing sample size (see Table 1 for details).

Stating that the diameter distribution varies, is another way of saying that $\boldsymbol{\theta}$ is not constant throughout. For example, consider a HPS with n points in which $\boldsymbol{\theta}$ varies from point to point according to some stochastic process. Then $\boldsymbol{\theta}$ may be considered a random variable and may exhibit a spatial covariance structure between points. Such a scenario might possibly be modeled using continuous mixtures.

For illustration, assume that the conditional distribution of tallied diameters given $\boldsymbol{\theta}$ is two-parameter Weibull; *viz.*, $X_2^*|\boldsymbol{\theta} \sim f_2^*(x|\boldsymbol{\theta})$. It would then make sense to use a bivariate distribution to model the variation in $\boldsymbol{\theta}$ over the stand. One candidate probability model for the joint distribution of $\boldsymbol{\Theta} \sim f(\boldsymbol{\theta}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is bivariate normal with mean and covariance matrix $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively. Other bivariate distributions could also be considered. With the bivariate normal, particular care must be taken to ensure that, for all practical purposes, $\boldsymbol{\theta} > 0$. Thus, extreme variability between HPS points coupled with small scale or shape parameters might argue against its use. However, for the sake of illustration it is a useful model.

With this modeling scheme, the bivariate normal is the mixing distribution and the marginal stand tally distribution for X_2^* under HPS would be given by

$$f_2^*(x; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \iint f_2^*(x|\boldsymbol{\theta}) f(\boldsymbol{\theta}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\, d\gamma\, d\beta \tag{6}$$

Undoubtedly, the marginal distribution given in (6) does not exist in closed form and the integration would require numerical methods. However, it does pose an interesting interpretation for the final density once γ and β have been integrated out. The only two remaining parameters are $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Thus, the following method might be used to fit such a distribution based on the techniques discussed earlier in this paper

1. Fit a size-biased two-parameter Weibull PDF to each of n individual HPS points in the stand using the methods in previous sections.
2. Calculate the sample mean vector $\hat{\boldsymbol{\mu}}$ and respective sample covariance matrix $\boldsymbol{S}$ from the parameter estimates on the n individual HPS sample points, as estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively.

The mixture density $f_2^*(x; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ may now be estimated by $f_2^*(x; \hat{\boldsymbol{\mu}}, \boldsymbol{S})$. However, it must be kept in mind that the above in no way has proven that $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{S}$ have any of the desirable properties of say, MLEs, for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. It is simply a possible model for a heterogeneous stand parameter estimation scenario.

Discussion

The discussion on estimation and applications of size-biased distributions to this point demonstrate that they have both a solid theoretical underpinning and practical use in forestry. Well-known relationships between basal area and horizontal point sampling, for example, are preserved under this theory. It should not be surprising then, that other results will also hold for size-biased distributions. For example, Gove (2003b) has shown that the relationship between the quadratic mean stand diameter and the harmonic mean basal area from a HPS holds for area-biased distributions; the result is shown to apply to the BASD too.

In fact, size-biased distributions can also lend new insight into previously unknown relationships. For example, Gove and Patil (1998) showed that the third raw moment of the DBH-frequency distribution has an intuitive and consistent interpretation through the BASD—a result that had been missed prior to the application of this theory. Similarly, it can be shown analytically (Gove, 2003b) that $f(x; \boldsymbol{\theta})$ and $f_2^*(x; \boldsymbol{\theta})$ will *always* cross at the quadratic mean stand diameter ($\bar{D}_q$). To illustrate, refer back to Figure 1, for this stand $\bar{D}_q = 28.08$ cm, and this is exactly where the two PDFs cross.

A new computer program (*Balance*) (Gove, 2002) has been developed to facilitate the use of size-biased distributions in forestry. *Balance* was written in FORTRAN-90, and is fully integrated with a graphical user interface and runs under Microsoft Windows© operating systems. Currently, *Balance* allows the user to fit two- and three-parameter equal probability Weibull distributions. In addition, both length- and area-biased versions of these PDFs can also be fitted. *Balance* computes the moment estimates and then uses these as starting values for ML. Results are presented in three windows; *viz.*, a listing of the input data in a grid window, a summary report window with fit statistics, and a graphics window with various graphical displays. The latter may be exported in encapsulated PostScript format, an example of which is shown in Figure 1.

Notice in this figure that, even though the equal probability density was estimated for the DBH-frequency distribution, *Balance* also shows the related BASD.

Clearly, size-biased distributions provide a useful paradigm for sampling and modeling in forestry research. The availability of computer programs such as *Balance* to make fitting such distributions easier, should serve to increase their application.

References

Avery, T. E. and H. E. Burkhart (1994). *Forest measurements* (Fourth ed.). N.Y.: McGraw-Hill.

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

Cohen, A. C. (1965). Maximum likelihood estimation in the weibull distribution based on complete and on censored samples. *Technometrics 7*, 579–588.

Dennis, B. and G. Patil (1984). The gamma distribution and weighted multimodal gamma distributions as models of population abundance. *Mathematical Biosciences 68*, 187–212.

Fisher, R. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics 6*, 13–25.

Gove, J. H. (2000). Some observations on fitting assumed diameter distributions to horizontal point sampling data. *Canadian Journal of Forest Research 30*, 521–533.

Gove, J. H. (2002). *Balance*: A system for fitting diameter distribution models. General Technical Report NE-xx, USDA Forest Service. (In review).

Gove, J. H. (2003a). Moment and maximum likelihood estimators for Weibull distributions under length- and area-biased sampling. *Ecological and Environmental Statistics*. (Submitted).

Gove, J. H. (2003b). A note on the relationship between the quadratic mean stand diameter and harmonic mean basal area under size-biased distribution theory. *Canadian Journal of Forest Research*. (Submitted).

Gove, J. H. and G. P. Patil (1998). Modeling the basal area-size distribution of forest stands: A compatible approach. *Forest Science 44*(2), 285–297.

Gove, J. H., A. Ringvall, G. Ståhl, and M. J. Ducey (1999). Point relascope sampling of downed coarse woody debris. *Canadian Journal of Forest Research 29*, 1718–1726.

Grosenbaugh, L. R. (1958). Point sampling and line sampling: probability theory, geometric implications, synthesis. Occasional Paper 160, USDA Forest Service, Southern Forest Experiment Station.

Hyink, D. M. and J. W. Moser, Jr (1983). A generalized framework for projecting forest yield and stand structure using diameter distributions. *Forest Science 29*, 85–95.

Kaiser, L. (1983). Unbiased estimation in line-intercept sampling. *Biometrics 39*, 965–976.

Lappi, J. and R. L. Bailey (1987). Estimation of diameter increment function or other tree relations using angle-count samples. *Forest Science 33*, 725–739.

Magnussen, S., P. Eggermont, and V. N. LaRiccia (1999). Recovering tree heights from airborne laser scanner data. *Forest Science 45*(3), 407–422.

Mahfoud, M. and G. P. Patil (1982). On weighted distributions. In G. Kallianpur, P. Krishnaiah, and J. Ghosh (Eds.), *Statistics and probability: Essays in honor of C.R. Rao*, New York, pp. 479–492. North-Holland.

Patil, G. P. (1981). Studies in statistical ecology involving weighted distributions. In J. K. Ghosh and J. Roy (Eds.), *Applications and new directions*, Calcutta, India, pp. 478–503. Proceedings of the Indian Statistical Institute Golden Jubilee: Statistical Publishing Society.

Patil, G. P. and J. K. Ord (1976). On size-biased sampling and related form-invariant weighted distributions. *Sankhȳa, Series B 38*(1), 48–61.

Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment. In G. P. Patil (Ed.), *Classical and contagious discrete distributions*, Calcutta, pp. 320–332. Statistical Publishing Society.

Ståhl, G. (1998). Transect relascope sampling—a method for the quantification of coarse woody debris. *Forest Science 44*(1), 58–63.

Taillie, C., G. P. Patil, and R. Hennemuth (1995). Modeling and analysis of recruitment distributions. *Ecological and Environmental Statistics 2*(4), 315–329.

Van Deusen, P. C. (1986). Fitting assumed distributions to horizontal point sample diameters. *Forest Science 32*, 146–148.

Warren, W. (1975). Statistical distributions in forestry and forest products research. In G. P. Patil, S. Kotz, and J. K. Ord (Eds.), *Statistical distributions in scientific work*, Volume 2, Dordrecht, The Netherlands, pp. 369–384. D. Reidel.