

# Describing Maritime Pine Diameter Distributions with Johnson's $S_B$ Distribution Using a New All-Parameter Recovery Approach

Teresa Fidalgo Fonseca, Carlos Pacheco Marques, and Bernard R. Parresol

**Abstract:** Forest growth and yield models that describe the heterogeneity of a stand by implicitly defining size classes are valuable tools for sustainable forest management. In stand-level growth and yield models, often stand tables representing numbers of trees by diameter class are projected through the use of probability density functions (PDFs). Theoretical knowledge shows and empirical studies corroborate that the four-parameter Johnson's  $S_B$  PDF provides greater generality in fitting diameter distributions than many of the commonly applied PDFs in forestry, such as the beta, gamma, and Weibull PDFs. Distinct parameter estimation methods are available for the Johnson's  $S_B$  distribution. However, few studies have been conducted that estimate the Johnson's  $S_B$  PDF using a parameter recovery approach. Also, for those studies that have used a parameter recovery approach, often one (i.e., the location) or two (i.e., the location and the range) of the four parameters are assumed to be known. In this article, we present a parameter recovery approach to recover all four parameters of the Johnson's  $S_B$  PDF for diameter distributions from stand variables. The location, range, and shape parameters are recovered from the median and the first three noncentral moments of the diameter distribution. The first two of these noncentral moments correspond to the average and quadratic mean diameter. The third moment is interpreted as the product of the mean diameter for the diameter distribution based on basal area, rather than stems per ha, and the squared quadratic mean diameter. The proposed methodology is demonstrated for maritime pine (*Pinus pinaster* Ait.) stands in Portugal. FOR. SCI. 55(4):367–373.

**Keywords:** stand horizontal structure, moments, *Pinus pinaster*

MARITIME PINE (*PINUS PINASTER* AIT.) is the principal softwood species in Portugal, with a coverage of 23% of the continental forested area. According to the SCC\_PP (Composition Classification System for *Pinus Pinaster*, by Fonseca et al. 2005), which classifies stand horizontal structure based on heterogeneity of tree diameters, 61% of maritime pine pure stands are homogeneous in composition with diameter distributions not exceeding 25 cm in range and 5.5 cm in SD. The remaining 39% have wider ranges of diameters, often with an irregularly shaped diameter distribution (e.g., may be multimodal). Detailed information about the variability of tree diameters in natural or altered stands is a key indicator for sustainable forest management and is needed to assess carbon sequestration of the tree components in forest ecosystems.

Attempts to numerically analyze stem diameter distributions date from the early 1900s (Loetsch et al. 1973). Since then, many different procedures to represent the variety of observed diameter distributions have been advocated and tested. For Portuguese maritime pine managed stands, good results were obtained using three probability density functions (PDFs): the three-parameter Weibull PDF, suggested for use in forestry by Bailey and Dell (1973); a distribution-free PDF fitted using the percentile-based method, developed by Borders et al. (1987); and the four-parameter

Johnson's  $S_B$  PDF (Johnson 1949). Details concerning application of each of these distributions to Portuguese forests can be found in Páscoa (1987), Marques and Bento (1989), and Fonseca (2004), respectively. Many studies support the use of the Johnson's  $S_B$  distribution to accurately describe diameter distributions (e.g., Hafley and Schreuder 1977, Von Gadow 1983, Mowrer 1986, Tham 1988, Kamziah 1998, Scolforo and Thierschi 1998, Lopes 2001, Kiviste et al. 2003, Parresol 2003, Fonseca 2004), height distributions (e.g., Hafley and Schreuder 1977), and volume ratio distributions (e.g., Newberry and Burk 1985) for a variety of tree species.

Johnson's  $S_B$  PDF (hereafter called the  $S_B$  PDF or  $S_B$  distribution) is part of a system of distributions proposed by Johnson (1949) generated by methods of translation on a standard normal variate that permits representation over the whole possible region of the plane ( $\beta_1, \beta_2$ ), where  $\beta_1$  is the square of the standardized measure of skewness and  $\beta_2$  is the standardized measure of kurtosis. The system consists of three distribution families:  $S_U$ ,  $S_L$ , and  $S_B$ , defined, respectively, for unbounded variates, variates bounded at one end, and bounded from above and below. The PDF for a variable  $X$  that follows an  $S_B$  PDF can be expressed as,

$$f(x) = \frac{\delta\lambda}{\sqrt{2\pi}(x - \xi)(\xi + \lambda - x)}$$

Teresa Fidalgo Fonseca, Forestry Department, University of Trás-os-Montes e Alto Douro, Apartado 1013, 5001-801 Vila Real, Portugal—Fax: 00351 259 350 821; tfonseca@utad.pt. Carlos Pacheco Marques, Forestry Department, University of Trás-os-Montes e Alto Douro, Apartado 1013, 5001-801 Vila Real, Portugal—cpmarque@utad.pt. Bernard R. Parresol, US Forest Service, Southern Research Station, Asheville, NC 28804—bparresol@fs.fed.us.

Acknowledgments: Part of this study was developed within the project POAgro 372 "Experimentation and Demonstration of Silvicultural and Sustainable Management Techniques in Maritime Pine Stands" funded by the Portuguese Institute of Research (INIAP). We extend special thanks to an anonymous associate editor for valuable suggestions that improved the organization and content presentation of the article.

$$\exp\left(-\frac{1}{2}\left[\gamma + \delta \ln\left(\frac{x - \xi}{\xi + \lambda - x}\right)\right]^2\right),$$

$$\xi < x < \xi + \lambda, \quad (1)$$

where  $\lambda, \delta > 0$ ,  $-\infty < \xi < \infty$ ,  $-\infty < \gamma < \infty$ . The parameter  $\lambda$  gives the range,  $\xi$  is the location parameter (lower bound),  $\delta$  and  $\gamma$  are shape parameters, and  $\gamma = 0$  indicates symmetry.

The  $S_B$  PDF possesses two properties that make it well suited to represent biological variables. First, because of the lower bound,  $\xi$ , and upper bound,  $\xi + \lambda$ , the  $S_B$  PDF can represent variables that have physical or natural constraints on their range. Second, the two parameters controlling the shape ( $\delta$  and  $\gamma$ ) allow a considerable amount of flexibility to fit a broad spectrum of distribution.

Parameters of the  $S_B$  PDF can be estimated by the percentile method, maximum likelihood, moments, and linear or nonlinear regression methods. Most approaches require a priori specification of one of the bounds of the distribution (an evaluation and comparison of the existing methods can be found in Zhou and McTague 1996, Kamziah et al. 1999, Scolforo et al. 2003, and Zhang et al. 2003). The use of the  $S_B$  distribution for diameter distributions in forest growth and yield models, fitted using parameter prediction or recovery methods (e.g., Hyink and Moser 1983), have been tested, although to a lesser extent. The parameter prediction approach has the following disadvantages: this approach does not assure compatibility between stand values generated from a distribution and the estimated (or observed) stand values and functions to relate one or more parameters to stand variables usually account for only a small proportion of the variation in parameters. For example, the shape parameter  $\delta$  has been shown to only weakly correlate with age (Kamziah 1998, Scolforo et al. 2003, Fonseca 2004). Better results are achieved through a parameter recovery-based approach or a mixed parameter recovery/parameter prediction approach. With the documented error index (Reynolds et al. 1988, see definition under the Assessment Criteria section of this article) values for comparable maritime pine stands, the  $S_B$  PDF provides a better goodness of fit than the distribution-free percentile method or fitting the Weibull function (Fonseca 2004).

Rennolls and Wang (2005) presented a definition of the  $S_B$  distribution in terms of the inverse transformation from normality, pointing out that the new parameterization is more natural than the one commonly used. According to Wang and Rennolls (2005), a new distribution model, named by the authors as the “logit-logistic” distribution, expands the coverage of the original  $S_B$  distribution in the skew-kurtosis shape-space.

Because Johnson’s  $S_B$  was shown to be accurate in representing diameter distributions for many species including maritime pine, it was selected to represent diameter distributions in the new growth-and-yield model ModisPinaster (*Model with diameter distribution for Pinus Pinaster*, by Fonseca and Marques 2006) for managed maritime pine stands. The algorithm used to incorporate the  $S_B$  in ModisPinaster was based on the parameter recovery method in combination with parameter prediction proposed by Par-

resol (2003). In brief, in his approach, Parresol assumed that the minimum location parameter was prespecified (set to 0.8 of minimum diameter in ModisPinaster). The range and two-shape parameters were then recovered from the median and the first two noncentral moments of the diameter distribution (average diameter and quadratic mean diameter). However, the location parameter can be a difficult variable to predict accurately as smaller trees in a stand are greatly influenced by microsite heterogeneity and silvicultural factors. In addition, the minimum diameter for a stand will change over time and with management practices. For example, thinning stands from below results in a change in the location parameter.

The objective of this article was to present a method to recover all four parameters of the Johnson’s  $S_B$  distribution by extending Parresol’s (2003) research. With this new method, the parameters of the  $S_B$  PDF can be solved by equating parameters of the empirical diameter distribution to their analytical counterparts from the  $S_B$  PDF. The method is then used to fit a diameter distribution for maritime pine stands.

### Three-Parameter Recovery Method

The common procedure to fit the  $S_B$  PDF is to reduce the four-parameter distribution to a two (or three)-parameter distribution. One or both of the location and range parameters are predicted directly, as a function of extreme diameter values in the sample data. The remaining parameters are then often solved from percentiles or moments. Knoebel and Burkhart (1991) relied on the two-percentile method to recover  $\gamma$  and  $\delta$ . Scolforo et al. (2003) described a moment approach to solve for the shape parameters making use of the proximate solutions given by Johnson and Kitchen (1971b). Parresol (2003) developed a percentile-moment method to simultaneously solve for the range and the shape parameters.

As an alternative, using parameter recovery models, stand-average attributes are directly predicted and then used to obtain estimates of the underlying diameter distribution (Hyink and Moser 1983). Parresol (2003) described a three-parameter recovery approach, in which the location parameter was separately estimated. He estimated  $\xi$  using regression techniques to extrapolate the random variable breast height diameter ( $d$ ) to the lower bound. Using this estimated location parameter, he then solved for the range and two-shape parameters, using a transformation presented in Johnson and Kotz (1970). Although there is no closed form expression for the  $S_B$  PDF, if the random variable  $X \sim S_B(\xi, \lambda, \gamma, \delta)$ , where  $X$  is the diameter, then

$$z = \gamma + \delta \ln[(x - \xi)/(\xi + \lambda - x)] \sim N(0, 1). \quad (2)$$

Given a new variable,

$$y = (x - \xi)/\lambda, \quad (3)$$

it follows from Equation 2 that

$$z = \gamma + \delta \ln[y/(1 - y)] \sim N(0, 1). \quad (4)$$

The new random variable,  $Y$ , will follow a distribution with the same shape parameters as  $X$  (Johnson and Kotz

1970). Using the  $Y$  random variable, the  $S_B$  PDF (Equation 1) becomes

$$f(y) = \frac{\delta}{y(1-y)\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\gamma + \delta \ln\left(\frac{y}{1-y}\right)\right]^2\right), \quad 0 < y < 1. \quad (5)$$

Setting  $z$  in Equation 4 equal to zero and rearranging in terms of parameter  $\gamma$  gives

$$\gamma = \delta \ln(1/y_{0.50} - 1) \quad (6)$$

where  $y_{0.50}$  is the median of  $Y$ .

By using the statistical expectation of  $X^p$  in terms of the  $Y$  variable,

$$E(X) = \bar{d} = E(\xi + \lambda Y) = \xi + \lambda \mu'_1(Y), \quad (7)$$

$$E(X^2) = dg^2 = E(\xi + \lambda Y)^2 = \xi^2 + 2\xi\lambda\mu'_1(Y) + \lambda^2\mu'_2(Y), \quad (8)$$

where  $\mu'_r$  denotes the  $r$ th noncentral moments of  $Y$  and  $\bar{d}$  is the average tree diameter. Quadratic mean diameter ( $dg$ ) is functionally related to number of trees per unit area ( $N$ ) and basal area per unit area ( $G$ ):  $G = kNdg^2$ . Hence, Equation 8 is equivalent to

$$G = kN[\xi^2 + 2\xi\lambda\mu'_1(Y) + \lambda^2\mu'_2(Y)]. \quad (9)$$

Because  $z$  is a unit normal variate, the  $r$ th noncentral moment of  $Y$  is

$$\mu'_r(Y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} [1 + e^{-(z-\gamma)/\delta}]^{-r} e^{-z^2/2} dz. \quad (10)$$

The relationship in Equation 6 is first used to eliminate  $\gamma$  in Equations 7 and 9 by substitution on Equation 10. The resulting system of two equations and two unknown parameters is nonlinear and must be solved by numerical procedures (see Parresol 2003 for details on the technical solution). Given estimates of  $G$ ,  $N$ ,  $\bar{d}$ , median tree diameter ( $d_{0.50}$ ), and the  $\xi$  parameter and Equations 7 and 9, the system is iteratively solved for  $\lambda$  and  $\delta$ . Parameter  $\gamma$  is then determined from Equation 6.

## All-Parameter Recovery Method

Alternatively, the four parameters of the  $S_B$  PDF can be recovered using a percentile-moment based procedure. To recover all of the parameters of the  $S_B$  distribution using Parresol's approach, a fourth equation must be added to Equations 6, 7, and 9. Our proposed approach relies on the 50th percentile and the first three noncentral moments of diameter distribution to solve this problem.

Johnson and Kitchen (1971a) described the construction of tables for fitting  $S_B$  distributions, using the first four moments. In their approach, they determined  $\delta$  and  $\gamma$ , given values of the moment ratios  $\sqrt{\beta_1}$  and  $\beta_2$ . Given shape parameter values, the location and range parameter could be

subsequently obtained from the mean and the SD of  $X$ . To our knowledge, the use of the third noncentral moment of  $X$  for moment-based recovery of a PDF has never been attempted for the  $S_B$  distribution. Burk and Newberry (1984) selected the third noncentral model as their first choice for their three-parameter Weibull function recovery system. However, they warned that  $\mu'_3$  was as not readily understood for diameter distributions as  $\mu'_1(\bar{d})$  and  $\mu'_2(dg^2)$ . More recently, Gove and Patil (1998) showed that weighted distribution theory provides an interpretation of the third moment in the context of easily determined variables. Specifically, interpretation arises when diameter distributions are displayed with respect to tree basal area (basal area-sized distribution [BASD]), rather than to tree frequency. The expected value of a BASD of tree diameters ( $d_3^*$ ) can be expressed as the ratio of successive noncentral moments of the original unweighted diameter distribution, specifically,

$$E[d_3^*] = \frac{\mu'_3(d)}{\mu'_2(d)}.$$

Designating the BASD mean as  $\bar{d}_G$ , it follows that the third noncentral moment of the diameter distribution is the product of the mean of the BASD and the squared quadratic mean diameter, that is,

$$\mu'_3(d) = \bar{d}_G dg^2.$$

Using the statistical expectation of  $X^3$  for the  $S_B$  distribution follows

$$E(X^3) = E[(X)^3] = E[(\xi + \lambda Y)^3].$$

Hence

$$\bar{d}_G dg^2 = \xi^3 + 3\xi^2\lambda\mu'_1(Y) + 3\xi\lambda^2\mu'_2(Y) + \lambda^3\mu'_3(Y). \quad (11)$$

Adding Equation 11 to Parresol's system results in recovery of all parameters, including the location ( $\xi$ ) parameter. The BASD mean ( $\bar{d}_G$ ) can be calculated from the plot diameters as a weighted mean using as the weight function the individual basal area ( $g$ ) values,

$$\bar{d}_G = \frac{\sum_{i=1}^n g_i d_i}{\sum_{i=1}^n g_i}.$$

Equations 6, 7, 9, and 11 are our proposed system to recover all the parameters of the  $S_B$  distribution. The relationship in Equation 6 is used to eliminate  $\gamma$  in Equations 7, 9, and 11. We are then left with a system of three nonlinear equations with three unknown parameters. Theoretically, this system allows recovery for the  $S_B$  parameters. However, solving a system of three nonlinear equations that involve the moments of  $Y$  is not straightforward.

To solve the four-parameter  $S_B$  recovery-model, we developed a technical procedure using the three-dimensional Levenberg-Marquardt algorithm. The partial derivatives needed to implement the technical procedure are given in the Appendix. Details of pseudomoments evaluation are given in Parresol (2003). The procedure was programmed in

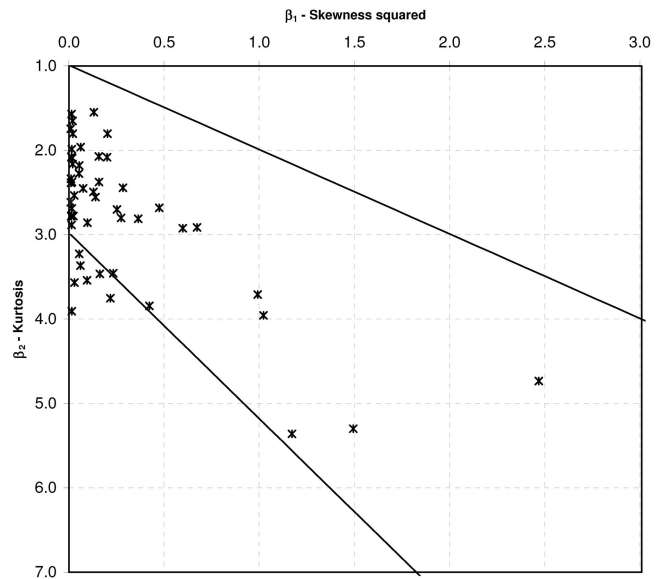
SAS (SAS Institute, Cary, NC) language, using optimization theory (details about the  $S_B$  parameter recovery programs are presented in Parresol et al. 2009). In this article, no further details on technical fitting procedures are presented.

## Parameter Recovery for Maritime Pine Stands

### Data Description

This study used information from the database on maritime pine (Data\_Pinaster) created and maintained over the last two decades at the Forestry Department of the University of Trás-os-Montes e Alto Douro. A total of 50 diameter distributions from sampled stands in Tâmega's Valley in North Portugal (latitude range: 41°15'–41°52'N; longitude range: 7°20'–8°00'W) were randomly selected for evaluation purposes. In each stand, circular 0.05-ha plots had been established. Available tree characteristics were diameter outside bark at breast height ( $d$ , cm) of all living trees exceeding 0 mm  $d$  at 1.30 m height; total height ( $h$ , m) and height to live crown ( $h_c$ , m) for a subset of trees; and mean height of the 100 largest trees per ha for stand dominant height ( $h_d$ , m) and site index (SI, m) evaluation. Diameters were measured to the nearest mm and heights to the nearest dm. Stand age ( $t$ , years) was evaluated in the dominant trees. Characterization of stand variables is shown in Table 1. Values of SI, at the index age of 35 years, were estimated using Marques' (1991) model. According to the SCC\_PP system (Fonseca et al. 2005), 29 sampled stands are homogeneous in diameter variability, whereas the remaining 21 stands are classified as heterogeneous, with 18 stands having irregularly shaped diameter distributions.

The estimated skewness,  $\sqrt{b_1}$ , values range from  $-1.57$  to 1.22 and estimated kurtosis,  $b_2$ , values vary from 1.55 to 5.36. Figure 1 shows these estimates in the  $(\beta_1, \beta_2)$  space, together with two reference lines. Certain combinations of  $\beta_1$  and  $\beta_2$  are mathematically impossible and occur in the region above the line  $\beta_2 - \beta_1 - 1 = 0$ . From Figure 1, 43 empirical distributions are in the  $S_B$  region (three are seen to be close to the  $S_L$  top limit) and the remaining seven are in the  $S_U$  region, all being included in the logit-logistic shape domain (see Wang and Rennolls 2005 for details on the



**Figure 1.** Representation of the 50 observations in the  $(\beta_1, \beta_2)$  space of skewness squared and kurtosis. Each mark represents a single distribution that has a position determined by the values of estimated skewness and kurtosis. The area above the top line ( $\beta_2 - \beta_1 - 1 = 0$ ) defines the impossible region. The lower line corresponds to the  $S_L$  distribution and separates the  $S_B$  region (area between the impossible region and the  $S_L$  line) from the  $S_U$  region (area below the  $S_L$  line).

logit-logistic domain limits). The graph suggests that 20% of study cases might be more adequately represented by another distribution rather than Johnson's  $S_B$ .

### Assessment Criteria

Model adequacy was evaluated by analysis of the goodness of fit, using 5-cm wide diameter classes, which is the usual range adopted for softwood species in Portugal. Goodness of fit was evaluated both by the Kolmogorov-Smirnov test and by the error index (Reynolds et al. 1988). The error index (EI) of each diameter distribution was calculated as the sum of the absolute deviations of the class basal areas,

$$e = \sum_{j=1}^k |\hat{G}_j - G_j|,$$

where  $\hat{G}_j$  and  $G_j$  are the estimated and the observed basal area of diameter class  $j$ .

## Results and Discussion

Initial stand conditions (stem diameter distribution median, mean and quadratic mean values, and BASD mean value) were input into the SAS program for parameter recovery. Convergent solutions were obtained for all stands. Location and range parameter estimates were found to be all greater than zero and, in general, conformed to two logical constraints:  $\check{\xi}$  being less than the observed minimum diameter ( $d_{\min}$ ) and  $\check{\xi} + \check{\lambda}$  being greater than the maximum value of the observed diameter ( $d_{\max}$ ).

Using the solutions of  $\check{\xi}$ ,  $\check{\lambda}$ ,  $\check{\delta}$ , and  $\check{\gamma}$ , the 50 estimated

**Table 1.** Stand characteristics of the test material

Variable	Minimum	Average	Maximum	SD
$t$ (years)	16	36	55	10.6
$h_d$ (m)	7.4	14.7	24.3	4.0
SI (m)	10.2	15.2	21.3	2.6
$N$ (trees ha <sup>-1</sup> )	260	1092	2940	767
$G$ (m <sup>2</sup> ha <sup>-1</sup> )	8.9	27.9	66.8	11.9
$d_{0.50}$ (cm)	7.2	19.8	41.0	7.1
$\bar{d}$ (cm)	7.4	20.0	40.4	7.0
$\bar{d}_g$ (cm)	7.8	20.8	40.9	7.0
$\bar{d}_G$ (cm)	9.1	22.8	42.3	7.3
$d_{\min}$ (cm)	1.9	10.1	23.4	6.2
$d_{\max}$ (cm)	14.5	31.3	51.2	8.3

$t$ , stand age;  $h_d$ , dominant height; SI, site index at 35 years reference age;  $G$ , basal area;  $N$ , number of trees per hectare;  $d_{0.50}$ , median diameter;  $\bar{d}$ , average diameter;  $\bar{d}_g$ , quadratic mean diameter;  $\bar{d}_G$ , mean basal area-sized distribution;  $d_{\min}$ , minimum diameter;  $d_{\max}$ , maximum diameter.

distributions were generated and compared with the observed distributions. The Kolmogorov-Smirnov goodness-of-fit tests for the 50 distributions resulted in rejection of the null hypothesis at the  $\alpha = 0.05$  level for two diameter distributions with  $(b_1, b_2)$  values that fell in the  $S_B$  region. The median value of the error index for the 50 distributions was  $5.80 \text{ m}^2 \text{ ha}^{-1}$ . For comparison purposes, the error index analysis was carried out also for the three-parameter recovery model proposed by Parresol (2003). With Parresol's approach, a 100% convergence rate was also achieved. Figure 2 gives the frequency polygons for the EI values for both the four-parameter and three-parameter recovery approaches. The averages (SD) for the EI values using the three-parameter versus the four-parameter recovery approach were  $8.01 (6.16)$  and  $7.70 (5.58) \text{ m}^2 \text{ ha}^{-1}$ , respectively. The averages and the shapes of the frequency polygons show that differences among simulated and observed distributions, weighted by basal area, are generally alike between the two parameter recovery approaches, indicating that the four-parameter recovery approach compares with the earlier one, while providing an estimate of the location parameter simultaneously. Analyzing the differences in EI values for the complete data set through the Wilcoxon (Wackerly et al. 1996) signed-rank test corroborates the resemblance ( $-36.0$  with a  $P$  value of  $0.732$ ) of fit between the estimated distributions using both approaches. For a few stands both approaches gave poor fits, suggesting that the empirical distributions are not of the assumed form or that the parameters have not been related to stand variables in a suitable way (Reynolds et al. 1988).

For this study, we used observed values of the stand variables as input, but predicted or projected stand variables can also be used. In addition, it should be pointed out that the supplementary variable needed for the complete parameter recovery procedure can be easily predicted from stand variables. For the case study data set, we obtained correlations greater than  $0.75$  between  $\bar{d}_G$  and the stand variables  $t$ ,  $N$ ,  $h_d$ , and  $dg$ . Similar correlation values were obtained for  $\bar{d}$  and  $d_{0.50}$  and  $t$ ,  $h_d$ , and  $dg$ . Accordingly, the four-parameter recovery approach is an interesting alternative to former

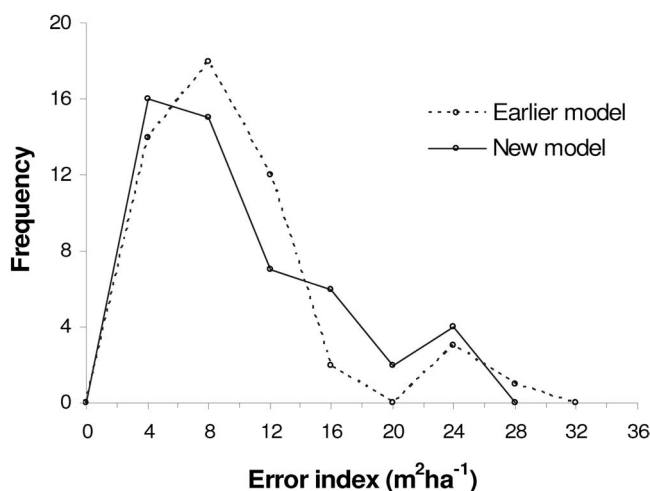


Figure 2. Frequency distributions for the error index for the four-parameter versus the three-parameter recovery approaches for the 50 distributions of this case study.

models in a recovery-approach context for growth and yield studies as it does not require prespecification of location and/or range parameter values. As pointed out by a reviewer, in the case of symmetrical distributions,  $d_{0.50}$  equals  $\bar{d}$ , and the four-equation recovery method reduces to a three-equation recovery method.

## Concluding Remarks

In this article, an all-parameter recovery method to obtain the four parameters of the Johnson's  $S_B$  PDF was presented. The approach uses stand-level variables that are easily understood. Moments such as mean diameter and quadratic mean diameter are well understood in terms of their relationship to stand characteristics, such as density, age, and site index. The fourth variable required in the new  $S_B$  model not only has a meaningful interpretation (as shown by Gove and Patil 1998), but also it can be easily calculated from tree diameters or predicted without difficulty from stand inventory variables.

Moments of  $S_B$  distributions are difficult to obtain analytically and hence parameter estimation by the method of moments is not easy. We applied the three-dimensional Levenberg-Marquardt routine to the resulting system of equations to solve for the parameters. Designing a system of logical and objective constraints using optimization theory led to a 100% convergence rate for the 50 maritime pine stands tested. From a technical perspective, the  $S_B$  parameter recovery model worked extremely well. We want to point out that a simpler system could be obtained using four percentiles. Although the percentile method is easier, it has three main disadvantages: it does not provide a unique solution because parameter estimates depend on the percentile points chosen; it is not clear what percentiles should be chosen; and percentiles, in general, have no ready biological interpretation as do moments.

Our case study results show that the all-parameter recovery model provides good approximations of the observed maritime pine diameter distributions. These results indicate that the all-parameter approach is a useful alternative for recovering parameters of the  $S_B$  distribution for diameter distributions of other species.

## Literature Cited

- BAILEY, R.L., AND T. R. DELL. 1973. Quantifying diameter distributions with the Weibull-function. *For. Sci.* 19:97-104.
- BORDERS, B.E., R.A. SOUTER, R.L. BAILEY, AND K.D. WARE. 1987. Percentile-based distributions characterize forest stand tables. *For. Sci.* 33:570-576.
- BURK, T.E., AND J.D. NEWBERRY. 1984. A simple algorithm for moment-based recovery of Weibull distribution parameters. *For. Sci.* 30:392-332.
- FONSECA, T.F. 2004. *Modelação do crescimento, mortalidade e distribuição diamétrica, do pinhal bravo no Vale do Tâmega*. Ph.D. dissertation, Univ. of Trás-os-Montes e Alto Douro, Vila Real, Portugal. 248 p.
- FONSECA, T.F., J.A. FREIRE, J.S. BENTO, AND C.P. MARQUES. 2005. SCC\_PP: Um sistema para classificação da constituição dos povoamentos de *Pinus pinaster*. P. 1-7 in *Actas das comunicações do V Congresso Florestal Nacional*, Sociedade Portuguesa de Ciências Florestais. Silva R., and F. Páscoa (eds.).

- Viseu, Portugal.
- FONSECA, T., AND C.P. MARQUES. 2006. *ModisPinaster Manual*. POAgro 372. Internal report. Univ. of Trás-os-Montes e Alto Douro, Vila Real, Portugal.
- GOVE, J.H., AND G.P. PATIL. 1998. Modeling the basal area-size distribution of forest stands: A compatible approach. *For. Sci.* 44:285–297.
- HAFLEY, W.L., AND M.T. SCHREUDER. 1977. Statistical distributions for fitting diameter and height data in even-aged stands. *Can. J. For. Res.* 7:481–487.
- HYINK, D.M., AND J.W. MOSER, JR. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. *For. Sci.* 29:85–95.
- JOHNSON, N.L. 1949. Systems of frequency curves generated by methods of translation. *Biometrika* 36:149–176.
- JOHNSON, N.L., AND J.O. KITCHEN. 1971a. Some notes on tables to facilitate fitting  $S_B$  curves. *Biometrika* 58:223–226.
- JOHNSON, N.L., AND J.O. KITCHEN. 1971b. Tables to facilitate fitting  $S_B$  curves II: Both terminals known. *Biometrika* 58: 657–668.
- JOHNSON, N.L., AND S. KOTZ. 1970. *Continuous univariate distributions*. Vol. 1. John Wiley & Sons, New York, NY. 300 p.
- KAMZIAH, A.K. 1998. *Development of diameter distribution yield prediction models for simulation of Acacia mangium plantations*. M.Sc. thesis, Univ. of Putra Malaysia, Malaysia. 180 p.
- KAMZIAH, A.K., M.I. AHMAD, AND J. LAPONGAN. 1999. Nonlinear regression approach to estimating Johnson  $S_B$  parameters for diameter data. *Can. J. For. Res.* 29:310–314.
- KIVISTE, A., A. NILSON, M. HORDO, AND M. MERENÄKK. 2003. Diameter distribution models and height-diameter equations for Estonian forest. P. 169–179 in *Modelling Forest Systems*. Amaro, A., D. Reed, and P. Soares (eds.). CABI Publishing, Wallingford, UK.
- KNOEBEL, B.R., AND H.E. BURKHART. 1991. A bivariate distribution approach to modeling forest diameter distribution at two points in time. *Biometrics* 47:241–253.
- LOETSCH, F., F. ZÖHRER, AND K.E. HALLER. 1973. *Forest Inventory*. Vol. II. BLV Verlagsgesellschaft, München, Germany. 469 p.
- LOPES, S. 2001. *Modelação matemática da distribuição de diâmetros em povoamentos de pinheiro bravo*. M.Sc. thesis, Faculdade de Ciências e Tecnologia da Univ. de Coimbra, Coimbra, Portugal. 102 p.
- MARQUES, C. P. 1991. Evaluating site quality of even-aged maritime pine stands in northern Portugal using direct and indirect methods. *For. Ecol. Manag.* 41:193–204.
- MARQUES, C.P., AND J.S. BENTO. 1989. Application of a system of adjacent percentiles to characterize diameter distributions. P. 153–158 in *Artificial intelligence and growth models for forest management decisions, Proc. of a IUFRO S4.01 symposium*, Burkhart, H.E., H.M. Rauscher, and K. Johann (eds.). School of Forestry and Wildlife Resources, Blacksburg, VA.
- MOWRER, H.T. 1986. *ASPENORM: A normal distribution growth and yield model for aspen in the Central Rocky Mountains*. US For. Ser. Res. Paper RM-264. 12 p.
- NEWBERRY, J.D., AND T.E. BURK. 1985.  $S_B$  distribution-based models for individual tree merchantable volume-total volume ratios. *For. Sci.* 31:389–398.
- PARRESOL, B.R. 2003. *Recovering parameters of Johnson's  $S_B$  distribution*. US For. Ser. Res. Paper SRS-31. 9 p.
- PARRESOL, B.R., T.F. FONSECA, AND C.P. MARQUES. 2009. Numerical details and SAS programs for parameter recovery of the  $S_B$  distribution. US For. Serv. Gen. Tech. Rep. SRS-xx. Southern Res. Stn., Asheville, NC. In press.
- PÁSCOA, F. 1987. *Estrutura, crescimento e produção em povoamentos de pinheiro bravo. Um modelo de simulação*. Ph.D. dissertation, Instituto Superior de Agronomia/Univ. Técnica de Lisboa, Lisbon, Portugal. 241 p.
- RENNOLLS, K., AND M. WANG. 2005. A new parametrization of Johnson's  $S_B$  distribution with application to fitting forest tree diameter data. *Can. J. For. Res.* 35:575–579.
- REYNOLDS, M.R., JR., T.E. BURK, AND W. HUANG. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. *For. Sci.* 34:373–399.
- SCOLFORO, J.R.S., F.C.V. TABAI, R.L.G. MACEDO, F.W. ACERBI JR., AND A.L. ASSIS. 2003.  $S_B$  distribution's accuracy to represent the diameter distribution of *Pinus taeda*, through five fitting methods. *For. Ecol. Manag.* 157:489–496.
- SCOLFORO, J.R.S., AND A. THIERSCHI. 1998. Estimativas e testes da distribuição de frequência diamétrica para *Eucalyptus camaldulensis*, através da distribuição  $S_B$  de Johnson, por diferentes métodos de ajuste. *Sci. Forest.* 54:93–106.
- THAM, Å. 1988. Estimate and test frequency distributions with the Johnson  $S_B$  function from stand parameters in young mixed stands after different thinning treatments. P. 255–262 in *Forest growth modeling and prediction: Proc. IUFRO Conference*. Vol. 1. Ek, A.R., S.R. Shiffley, and T.E. Burk (eds.). US For. Serv. Gen. Tech. Rep. NC-120. North Central Forest Experiment Station, Minneapolis, MN.
- VON GADOW, K. 1983. Fitting distributions in *Pinus patula* stands. *Suid-Afrik. Bosbouydskrif.* 126:20–29.
- WACKERLY, D.D., W. MENDENHAL, AND R.L. SCHEAFFER. 1996. *Mathematical statistics with applications*. Duxbury Press, North Scituate, MA. 789 p.
- WANG, M., AND K. RENNOLLS. 2005. Tree diameter distribution modelling: Introducing the logit-logistic. *Can. J. For. Res.* 35:1305–1313.
- ZHANG, L., K.C. PACKARD, AND C. LIU. 2003. A comparison estimation methods for fitting Weibull and Johnson's  $S_B$  distributions to mixed spruce-fir stands in northeastern North America. *Can. J. For. Res.* 33:1340–1347.
- ZHOU, B., AND J.P. MCTAGUE. 1996. Comparison and evaluation of five methods of estimation of Johnson system parameters. *Can. J. For. Res.* 26:928–935.

## Appendix

The three-dimensional Levenberg-Marquardt algorithm can be written in vector notation as

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\mathbf{J}'(\mathbf{x}_i)\mathbf{J}(\mathbf{x}_i) + c_i\mathbf{I}]^{-1}\mathbf{J}'(\mathbf{x}_i)\mathbf{f}(\mathbf{x}_i),$$

where  $c_i \geq 0$  is a scaling factor and the Jacobian at each iteration point  $\mathbf{x}_i$  is

$$\mathbf{J}_i = \left( \frac{\partial f_j}{\partial x^{(k)}} \right)_{\mathbf{x}=\mathbf{x}_i}$$

Let us denote the relationships in Equations 7, 9, and 11 as the  $f_j$ :

$$f_1 = \xi + \lambda\mu'_1 \quad (\text{A1})$$

$$f_2 = [\xi^2 + 2\xi\lambda\mu'_1 + \lambda^2\mu'^2_2]kN \quad (\text{A2})$$

$$f_3 = \xi^3 + 3\xi^2\lambda\mu'_1 + 3\xi\lambda^2\mu'^2_2 + \lambda^3\mu'^3_3 \quad (\text{A3})$$

The Jacobian matrix of this system is

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial \xi} & \frac{\partial f_1}{\partial \lambda} & \frac{\partial f_1}{\partial \delta} \\ \frac{\partial f_2}{\partial \xi} & \frac{\partial f_2}{\partial \lambda} & \frac{\partial f_2}{\partial \delta} \\ \frac{\partial f_3}{\partial \xi} & \frac{\partial f_3}{\partial \lambda} & \frac{\partial f_3}{\partial \delta} \end{bmatrix} \quad (A4)$$

The partial derivatives of Equations A1–A3 with respect to the  $x^{(k)}$  parameters ( $\xi$ ,  $\lambda$ , and  $\delta$ ) depend on the partial derivatives of the moments of  $Y(\mu'_r)$ , see Equation 10). The partial derivatives of the moments of  $Y$  can be expressed as

$$\frac{\partial \mu'_r}{\partial \xi} = \frac{r\lambda}{(\lambda - d_{0.50} + \xi)(d_{0.50} - \xi)} (\mu'_{r+1} - \mu'_r), \quad (A5)$$

$$\frac{\partial \mu'_r}{\partial \lambda} = \frac{r}{\lambda - d_{0.50} + \xi} (\mu'_{r+1} - \mu'_r), \quad (A6)$$

$$\frac{\partial f_1}{\partial \xi} = 2 \left[ \xi + \lambda \mu'_1 + \frac{\xi \lambda^2 (\mu'_2 - \mu'_1) + \lambda^3 (\mu'_3 - \mu'_2)}{(\lambda - d_{0.50} + \xi)(d_{0.50} - \xi)} \right] kN \quad (A11)$$

$$\frac{\partial f_2}{\partial \lambda} = 2 \left[ \xi \mu'_1 + \lambda \mu'_2 + \frac{\xi \lambda (\mu'_2 - \mu'_1) + \lambda^2 (\mu'_3 - \mu'_2)}{\lambda - d_{0.50} + \xi} \right] kN \quad (A12)$$

$$\frac{\partial f_2}{\partial \delta} = \frac{2\lambda}{\delta^2} [\xi(\tilde{\mu}'_2 - \tilde{\mu}'_1) + \lambda(\tilde{\mu}'_3 - \tilde{\mu}'_2)] kN \quad (A13)$$

$$\frac{\partial f_3}{\partial \xi} = 3 \left[ \frac{\xi^2 \lambda^2 (\mu'_2 - \mu'_1) + 2\xi \lambda^3 (\mu'_3 - \mu'_2) + \lambda^4 (\mu'_4 - \mu'_3)}{(\lambda - d_{0.50} + \xi)(d_{0.50} - \xi)} + \xi^2 + 2\xi \lambda \mu'_1 + \lambda^2 \mu'_2 \right] \quad (A14)$$

$$\frac{\partial f_3}{\partial \lambda} = 3 \left[ \frac{\xi^2 \lambda (\mu'_2 - \mu'_1) + 2\xi \lambda^2 (\mu'_3 - \mu'_2) + \lambda^3 (\mu'_4 - \mu'_3)}{\lambda - d_{0.50} + \xi} + \xi^2 \mu'_1 + 2\xi \lambda \mu'_2 + \lambda^2 \mu'_3 \right] \quad (A15)$$

$$\frac{\partial f_3}{\partial \delta} = \frac{3\lambda}{\delta^2} [\xi^2(\tilde{\mu}'_2 - \tilde{\mu}'_1) + 2\xi \lambda(\tilde{\mu}'_3 - \tilde{\mu}'_2) + \lambda^2(\tilde{\mu}'_4 - \tilde{\mu}'_3)] \quad (A16)$$

Substituting Equations A8 through A16 into Equation A4 gives the three-dimensional Levenberg-Marquardt formula to solve the system of equations.

$$\frac{\partial \mu'_r}{\partial \delta} = \frac{r}{\delta^2} (\tilde{\mu}'_{r+1} - \tilde{\mu}'_r), \quad (A7)$$

where  $\tilde{\mu}'_r$  in Equation A7 denotes the pseudomoment of  $Y$  (Parresol 2003, p. 4).

Using the results of Equations A5–A7, the partial derivatives of  $f_1$ ,  $f_2$ , and  $f_3$  with respect to the  $\xi$ ,  $\lambda$ , and  $\delta$  parameters are

$$\frac{\partial f_1}{\partial \xi} = 1 + \frac{\lambda^2 (\mu'_2 - \mu'_1)}{(\lambda - d_{0.50} + \xi)(d_{0.50} - \xi)} \quad (A8)$$

$$\frac{\partial f_1}{\partial \lambda} = \mu'_1 + \frac{\lambda (\mu'_2 - \mu'_1)}{\lambda - d_{0.50} + \xi} \quad (A9)$$

$$\frac{\partial f_1}{\partial \delta} = \frac{\lambda}{\delta^2} (\tilde{\mu}'_2 - \tilde{\mu}'_1) \quad (A10)$$