

## Fatigue Statistical Distributions Useful for Modeling Diameter and Mortality of Trees

Distribuciones estadísticas de fatiga útiles para modelar diámetro y  
mortalidad de árboles

VÍCTOR LEIVA<sup>1,a</sup>, M. GUADALUPE PONCE<sup>2,b</sup>, CAROLINA MARCHANT<sup>1,c</sup>,  
OSCAR BUSTOS<sup>3,d</sup>

<sup>1</sup>DEPARTAMENTO DE ESTADÍSTICA, UNIVERSIDAD DE VALPARAÍSO, VALPARAÍSO, CHILE

<sup>2</sup>INSTITUTO DE MATEMÁTICAS Y FÍSICA, UNIVERSIDAD DE TALCA, TALCA, CHILE

<sup>3</sup>DEPARTAMENTO DE PRODUCCIÓN FORESTAL, UNIVERSIDAD DE TALCA, TALCA, CHILE

---

### Abstract

Mortality processes and the distribution of the diameter at breast height (DBH) of trees are two important problems in forestry. Trees die due to several factors caused by stress according to a phenomenon similar to material fatigue. Specifically, the force (rate) of mortality of trees quickly increases at a first stage and then reaches a maximum. In that moment, this rate slowly decreases until stabilizing at a constant value in the long term establishing a second stage of such a rate. Birnbaum-Saunders (BS) distributions are models that have received considerable attention currently due to their interesting properties. BS models have their genesis from a problem of material fatigue and present a failure or hazard rate (equivalent to the force of mortality) that has the same behavior as that of the DBH of trees. Then, BS distributions have arguments that transform them into models that can be useful in forestry. In this paper, we present a methodology based on BS distributions associated with this forest thematic. To complete this study, we perform an application of five real DBH data sets (some of them unpublished) that provides statistical evidence in favor of the BS methodology in relation to the forestry standard methodology. This application provides valuable financial information that can be used for making decisions in forestry.

**Key words:** data analysis, force of mortality, forestry, hazard rate.

### Resumen

<sup>a</sup>Professor. E-mail: victor.leiva@uv.cl

<sup>b</sup>Assistant profesor. E-mail: gponce@utalca.cl

<sup>c</sup>Assistant profesor. E-mail: carolina.marchant@uv.cl

<sup>d</sup>Assistant profesor. E-mail: obustos@utalca.cl

Los procesos de mortalidad y la distribuci3n del diámetro a la altura del pecho (DAP) de árboles son dos problemas importantes en el área forestal. Los árboles mueren debido a diversos factores causados por estrés mediante un fenómeno similar a la fatiga de materiales. Específicamente, la fuerza (tasa) de mortalidad de árboles crece rápidamente en una primera fase y luego alcanza un máximo, momento en el que comienza una segunda fase en donde esta tasa decrece lentamente estabilizándose en una constante en el largo plazo. Distribuciones Birnbaum-Saunders (BS) son modelos que han recibido una atención considerable en la actualidad debido a sus interesantes propiedades. Modelos BS nacen de un problema de fatiga de materiales y poseen una tasa de fallas (equivalente a la fuerza de mortalidad) que se comporta de la misma forma que ésa del DAP de árboles. Entonces, distribuciones BS poseen argumentos que las transforman en modelos que puede ser útiles en las ciencias forestales. En este trabajo, presentamos una metodología basada en la distribuci3n BS asociada con esta temática forestal. Para finalizar, realizamos una aplicaci3n con cinco conjuntos de datos reales (algunos de ellos no publicados) de DAP que proporciona una evidencia estadística en favor de la metodología BS en relación a la metodología estándar usada en ciencias forestales. Esta aplicaci3n entrega informaci3n que puede ser valiosa para tomar decisiones forestales.

**Palabras clave:** análisis de datos, fuerza de mortalidad, silvicultura, tasa de riesgo.

## 1. Introduction

The determination of the statistical distribution of the diameter at breast height (DBH) of trees, and its relationship to the age, composition, density and geographical location where a forest is localized are valuable information for different purposes (Bailey & Dell 1973, Santelices & Riquelme 2007). Specifically, the distribution of the DBH is frequently used to determine the volume of wood from a stand allowing us to make decisions about: (i) productivity (quantity); (ii) diversity of products (quality); (iii) tree ages (mortality); and (iv) harvest policy and trees pruning (regeneration). Then, to know the DBH distribution may help to plan biological and financial management aspects of a forest in a more efficient way (Rennolls, Geary & Rollinson 1985). For example, trees with a large diameter are used for wood production, while trees with a small diameter are used for cellulose production. Thus, the four mentioned concepts (quality, quantity, mortality and regeneration) propose a challenge to postulate models that allow us to describe the forest behavior based on the DBH distribution.

Several statistical distributions have been used in the forestry area mainly to model the DBH. These distributions (in chronological order) are the models:

- (i) Exponential (Meyer 1952, Schmelz & Lindsey 1965);
- (ii) Gamma (Nelson 1964);
- (iii) Log-normal (Bliss & Reinker 1964);

- (iv) Beta (Clutter & Bennett 1965, McGee & Della-Bianca 1967, Lenhart & Clutter 1971, Li, Zhang & Davis 2002, Wang & Rennolls 2005);
- (v) Weibull (Bailey & Dell 1973, Little 1983, Rennolls et al. 1985, Zutter, Oderwald, Murphy & Farrar 1986, Borders, Souter, Bailey & Ware 1987, McEwen & Parresol 1991, Maltamo, Puumalainen & Päivinen 1995, Pece, de Benítez & de Galíndez 2000, García-Güemes, Cañadas & Montero 2002, Wang & Rennolls 2005, Palahí, Pukkala & Trasobares 2006, Podlaski 2006);
- (vi) Johnson SB (Hafley & Schreuder 1977, Schreuder & Hafley 1977);
- (vii) Log-logistic (Wang & Rennolls 2005);
- (viii) Burr XII (Wang & Rennolls 2005) and
- (ix) Birnbaum-Saunders (BS) (Podlaski 2008).

The most used distribution is the Weibull model and the most recent is the BS model. In spite of the wide use of different statistical distributions to describe the DBH, the model selection has been based in empirical arguments supported by goodness-of-fit methods and not by theoretical arguments that justify its use. In order to propose DBH distributions with better arguments, mortality models based on cumulative stress can be considered (Podlaski 2008).

A statistical distribution useful for describing non-negative data that has recently received considerable attention is the BS model. This two-parameter distribution is unimodal and positively skewed. For more details about the BS distribution, see Birnbaum & Saunders (1969) and Johnson, Kotz & Balakrishnan (1995, pp. 651-663). The interest for the BS distribution is due to its theoretical arguments based on the physics of materials, its properties and its relation to the normal distribution. Some extensions and generalization of the BS distributions are attributed to Díaz-García & Leiva (2005); Vilca & Leiva (2006); Guiraud, Leiva & Fierro (2009). In particular, the BS-Student-*t* distribution has been widely studied (Azevedo, Leiva, Athayde & Balakrishnan 2012). Although BS distributions have their origin in engineering, these have been applied in several other fields, such as environmental sciences and forestry (Leiva, Barros, Paula & Sanhueza 2008, Podlaski 2008, Leiva, Sanhueza & Angulo 2009, Leiva, Vilca, Balakrishnan & Sanhueza 2010, Leiva, Athayde, Azevedo & Marchant 2011, Vilca, Santana, Leiva & Balakrishnan 2011, Ferreira, Gomes & Leiva 2012, Marchant, Leiva, Cavieres & Sanhueza 2013). Podlaski (2008) employed the BS model to describe DBH data for silver fir (*Abies alba* Mill.) and European beech (*Fagus sylvatica* L.) from a national park in Poland, using theoretical arguments. In addition, based on goodness-of-fit methods, he discovered that the BS distribution was the model that best described these data, displacing the Weibull distribution.

The aims of the present work are: (i) to introduce a methodology based on BS distributions (one of them being novel) for describing DBH data that can be useful for making decisions in forestry and (ii) to carry out practical applications of real DBH data sets (some of them unpublished) that illustrate this methodology. The article is structured as follows: In the second section, we explain the methods

employed in this study, including a theoretical justification for the use of the BS distribution to model DBH data. In the third section, we establish an application with five real data sets of DBH using a methodology based on BS distributions. This methodology furnishes statistical evidence in its favor, in relation to the standard methodology used in forestry. This application provides valuable financial information that can be used for making decisions in forestry. Finally, we sketch some discussions and conclusions.

## 2. Methods

### 2.1. A Fatigue Model

The BS distribution is based on a physical argument that produces fatigue in the materials (Birnbaum & Saunders 1969). This argument is the Miner or cumulative damage law (Miner 1945). Birnbaum & Saunders (1968) provided a probabilistic interpretation of this law. The BS or fatigue life distribution was obtained from a model that shows failures to occur due to the development and growing of a dominant crack provoked by stress. This distribution describes the total time elapsed until a type of cumulative damage inducted by stress exceeds a threshold of resistance of the material producing its failure or rupture. Birnbaum & Saunders (1969) demonstrated that the failure rate (hazard rate or force of mortality) associated with their model has two phases. During the first phase, this rate quickly increases until a maximum point (change or critical point) and then a second phase starts when the failure rate begins to slowly decrease until it is stabilized at a constant greater than zero. Fatigue processes have failure rates which usually present in this way. In addition, these processes can be divided into three stages:

- (A1) The beginning of an imperceptible fissure;
- (A2) The growth and propagation of the fissure, which provokes a crack in the material specimen due to cyclic stress and tension; and
- (A3) The rupture or failure of the material specimen due to fatigue.

The stage (A3) occupies a negligible lifetime. Therefore, (A2) contains most of the time of the fatigue life. For this reason, statistical models for fatigue processes are primarily concerned with describing the random variation of lifetimes associated with (A2) through two-parameter life distributions. These parameters allow those specimens subject to fatigue to be characterized and at the same time predicting their behavior under different force, stress and tension patterns.

Having explained the physical framework of the genesis of the BS distribution, it is now necessary to make the statistical assumptions. Birnbaum & Saunders (1969) used the knowledge of certain type of materials failure due to fatigue to develop their model. The fatigue process that they used was based on the following:

- (B1) A material specimen is subjected to cyclic loads or repetitive shocks, which produce a crack or wear in this specimen;
- (B2) The failure occurs when the size of the crack in the material specimen exceeds a certain level of resistance (threshold), denoted by  $\omega$ ;
- (B3) The sequence of loads imposed in the material is the same from one cycle to another;
- (B4) The crack extension due to a load  $l_i$  ( $X_i$  say) during the  $j$ th cycle is a random variable (r.v.) governed by all the loads  $l_j$ , for  $j < i$ , and by the actual crack extension that precedes it;
- (B5) The total size of the crack due to the  $j$ th cycle ( $Y_i$  say) is an r.v. that follows a statistical distribution of mean  $\mu$  and variance  $\sigma^2$ ; and
- (B6) The sizes of cracks in different cycles are independent.

Notice that the total crack size due to the  $(j + 1)$ th cycle of load is  $Y_{j+1} = X_{jm+1} + \dots + X_{jm+m}$ , for  $j, m = 0, 1, 2, \dots$ . Thus, the accumulated crack size at the end of the  $n$ th stress cycle is  $S_n = \sum_{j=1}^n Y_j$ . Then, based on it, (B1)-(B6) and the central limit theorem, we have  $Z_n = [S_n - n\mu]/\sqrt{n\sigma^2} \sim N(0, 1)$ , as  $n$  approaches to  $\infty$ , i.e.,  $Z_n$  follows approximately a standard normal distribution. Now, let  $N$  be the number of stress cycles until the specimen fails. The cumulative distribution function (c.d.f.) of  $N$ , based on the total probability theorem, is  $P(N \leq n) = P(N \leq n, S_n > \omega) + P(N \leq n, S_n \leq \omega) = P(S_n > \omega) + P(N \leq n, S_n \leq \omega)$ . Notice that  $P(N \leq n, S_n \leq \omega) > 0$ , because  $S_n$  follows approximately a normal distribution, but this probability is negligible, so that  $P(N \leq n) \approx P(S_n > \omega)$ , and hence

$$P(N \leq n) \approx P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} > \frac{\omega - n\mu}{\sigma\sqrt{n}}\right) = \Phi\left(\frac{\sqrt{\omega\mu}}{\sigma} \left[\sqrt{\frac{n}{\omega/\mu}} - \sqrt{\frac{\omega/\mu}{n}}\right]\right) \quad (1)$$

where  $\Phi(\cdot)$  is the normal standard c.d.f. However, we must suppose the probability that  $Y_j$  given in (B5) takes negative values is zero. Birnbaum & Saunders (1969) used (1) to define their distribution, considering the discrete r.v.  $N$  as a continuous r.v.  $T$ , i.e., the number of stress cycles until to fail  $N$  is replaced by the total time until to fail  $T$  and the  $n$ th cycle by the time  $t$ . Thus, considering the reparameterization  $\alpha = \sigma/\sqrt{\omega\mu}$  and  $\beta = \omega/\mu$ , and that (1) is exact instead of approximated, we obtain the c.d.f. of the BS distribution for the fatigue life with shape ( $\alpha$ ) and scale ( $\beta$ ) parameters given by

$$F_T(t) = \Phi\left(\frac{1}{\alpha} \left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right]\right), \quad t > 0, \alpha > 0, \beta > 0 \quad (2)$$

To suppose (1) is exact, it means to suppose  $Y_j$  follows exactly a  $N(\mu, \sigma^2)$  distribution in (B5).

## 2.2. Birnbaum-Saunders Distributions

If an r.v.  $T$  has a c.d.f. as in (2), then it follows a BS distribution with shape ( $\alpha > 0$ ) and scale ( $\beta > 0$ ) parameters, which is denoted by  $T \sim \text{BS}(\alpha, \beta)$ . Here, the parameter  $\beta$  is also the median. Hence, BS ( $T$  say) and normal standard ( $Z$  say) r.v.'s are related by

$$T = \beta \left[ \frac{\alpha Z}{2} + \sqrt{\left\{ \frac{\alpha Z}{2} \right\}^2 + 1} \right]^2 \sim \text{BS}(\alpha, \beta) \text{ and } Z = \frac{1}{\alpha} \left[ \sqrt{\frac{T}{\beta}} - \sqrt{\frac{\beta}{T}} \right] \sim \text{N}(0, 1) \quad (3)$$

In addition,  $W = Z^2$  follows a  $\chi^2$  distribution with one degree of freedom (d.f.), denoted by  $W \sim \chi^2(1)$ . The probability density function (p.d.f.) of  $T$  is

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2} \left[ \frac{t}{\beta} + \frac{\beta}{t} - 2 \right]\right) \frac{1}{2\alpha\beta} \left[ \left\{ \frac{t}{\beta} \right\}^{-1/2} + \left\{ \frac{t}{\beta} \right\}^{-3/2} \right], \quad t > 0 \quad (4)$$

The  $q$ th quantile of  $T$  is  $t_q = \beta[\alpha z_q/2 + \sqrt{\{\alpha z_q/2\}^2 + 1}]^2$ , for  $0 < q < 1$ , where  $t_q = F_T^{-1}(q)$ , with  $F_T^{-1}(\cdot)$  being the inverse c.d.f. of  $T$ , and  $z_q$  the  $\text{N}(0, 1)$   $q$ th quantile. The mean, variance and coefficient of variation (CV) of  $T$  are

$$E[T] = \frac{\beta}{2} [2 + \alpha^2], \quad V[T] = \frac{\beta^2 \alpha^2}{4} [4 + 5\alpha^2] \text{ and } \text{CV}[T] = \frac{\alpha\sqrt{4+5\alpha^2}}{2+\alpha^2} \quad (5)$$

Although the BS distribution can be useful to model the DBH, there are several reasons to consider that the DBH distribution could start from a value greater than zero. In such a situation, a shifted version of the BS (ShBS) distribution, with shape ( $\alpha > 0$ ), scale ( $\beta > 0$ ) and shift ( $\gamma \in \mathbb{R}$ ) parameters, is needed, which is denoted by  $T \sim \text{ShBS}(\alpha, \beta, \gamma)$ . Leiva et al. (2011) characterized this distribution assuming that if  $T = \beta[\alpha Z/2 + \sqrt{\{\alpha Z/2\}^2 + 1}]^2 \sim \text{ShBS}(\alpha, \beta, \gamma)$ , then,  $Z = [1/\alpha][\sqrt{\{T - \gamma\}/\beta} - \sqrt{\beta/\{T - \gamma\}}] \sim \text{N}(0, 1)$  and so again  $W = Z^2 \sim \chi^2(1)$ . Therefore, in this case, the p.d.f. and c.d.f. of  $T$  are

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2} \left[ \frac{t-\gamma}{\beta} + \frac{\beta}{t-\gamma} - 2 \right]\right) \frac{1}{2\alpha\beta} \left[ \left\{ \frac{t-\gamma}{\beta} \right\}^{-1/2} + \left\{ \frac{t-\gamma}{\beta} \right\}^{-3/2} \right] \quad (6)$$

and  $F_T(t) = \Phi([1/\alpha][\sqrt{\{t - \gamma\}/\beta} - \sqrt{\beta/\{t - \gamma\}}])$ , for  $t > \gamma$ , respectively. In addition, the  $q$ th quantile of  $T$  is similar to that from the non-shifted case plus the value  $\gamma$  at the end of such an expression. The mean, variance and CV of  $T$  are now

$$E[T] = \frac{\beta}{2} \left[ 2 + \alpha^2 + \frac{2\gamma}{\beta} \right], \quad V[T] = \frac{\beta^2 \alpha^2}{4} [4 + 5\alpha^2] \text{ and } \text{CV}[T] = \frac{\alpha\beta\sqrt{4+5\alpha^2}}{\beta[2+\alpha^2]+2\gamma} \quad (7)$$

## 2.3. Birnbaum-Saunders- $t$ -Student Distributions

If an r.v.  $T$  follows a BS- $t$  distribution with shape ( $\alpha > 0$ ,  $\nu > 0$ ) and scale ( $\beta > 0$ ) parameters, then the notation  $T \sim \text{BS-}t(\alpha, \beta; \nu)$  is used. Thus, if  $T = \beta[\alpha Z/2 + \sqrt{\{\alpha Z/2\}^2 + 1}]^2 \sim \text{BS-}t(\alpha, \beta; \nu)$ , then  $Z = [1/\alpha][\sqrt{T/\beta} - \sqrt{\beta/T}] \sim t(\nu)$ ,

with  $\nu$  d.f., and  $W = Z^2 \sim \mathcal{F}(1, \nu)$ . Therefore, in this case, the p.d.f. and c.d.f. of  $T$  are

$$f_T(t; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left[ 1 + \left\{ \frac{t}{\beta} + \frac{\beta}{t} - 2 \right\} / \{2\alpha^2\nu\} \right]^{-\frac{[\nu+1]}{2}} \frac{1}{2\alpha\beta} \left[ \left\{ \frac{t}{\beta} \right\}^{-1/2} + \left\{ \frac{t}{\beta} \right\}^{-3/2} \right]$$

$$F_T(t; \nu) = \Phi_t(t) = \frac{1}{2} \left[ 1 + I_{\frac{[1/\alpha^2][t/\beta + \beta/t - 2]}{[1/\alpha^2][t/\beta + \beta/t - 2] + \nu}} \left( \frac{1}{2}, \frac{\nu}{2} \right) \right], \quad t > 0 \tag{8}$$

respectively, where  $I_x(a, b) = [\int_0^x t^{a-1}\{1-t\}^{b-1} dt] / \int_0^1 t^{a-1}\{1-t\}^{b-1} dt$  is the incomplete beta function ratio. The  $q$ th quantile of  $T$  is

$$t_q = \beta[\alpha z_q/2 + \sqrt{\{\alpha z_q/2\}^2 + 1}]^2,$$

where  $z_q$  is the  $q$ th quantile of the  $t(\nu)$  distribution. The mean, variance and CV of  $T$  are now

$$E[T] = \frac{\beta}{2} [2 + A\alpha^2], \quad V[T] = \frac{\beta^2\alpha^2}{4} [4A + 5B\alpha^2] \quad \text{and} \quad CV[T] = \frac{\alpha\sqrt{4A+5B\alpha^2}}{2+A\alpha^2} \tag{9}$$

where  $A = \nu/[\nu - 2]$ , for  $\nu > 2$ , and  $B = \nu^2[\nu - 1]/[\{\nu - 6\}\{\nu - 2\}^2]$ , for  $\nu > 6$ .

Such as in the case of the BS distribution, we can define a new shifted version of the BS- $t$  (ShBS- $t$ ) distribution, with shape ( $\alpha > 0, \nu > 0$ ), scale ( $\beta > 0$ ) and shift ( $\gamma \in \mathbb{R}$ ) parameters, which is denoted by  $T \sim \text{ShBS-}t(\alpha, \beta, \gamma; \nu)$ . Thus, if  $T = \beta[\alpha Z/2 + \sqrt{\{\alpha Z/2\}^2 + 1}]^2 \sim \text{ShBS-}t(\alpha, \beta, \gamma; \nu)$ , then  $Z = [1/\alpha][\sqrt{\{T - \gamma\}}/\beta - \sqrt{\beta/\{T - \gamma\}}] \sim t(\nu)$  and so again  $W = Z^2 \sim \mathcal{F}(1, \nu)$ . Therefore, in this case, the p.d.f. and c.d.f. of  $T$  are

$$f_T(t; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left[ 1 + \left\{ \frac{t-\gamma}{\beta} + \frac{\beta}{t-\gamma} - 2 \right\} / \{2\alpha^2\nu\} \right]^{-\frac{[\nu+1]}{2}} \frac{1}{2\alpha\beta} \left[ \left\{ \frac{t-\gamma}{\beta} \right\}^{-1/2} + \left\{ \frac{t-\gamma}{\beta} \right\}^{-3/2} \right]$$

$$F_T(t; \nu) = \Phi_t(t - \gamma) = \frac{1}{2} \left[ 1 + I_{\frac{[1/\alpha^2][\{t-\gamma\}/\beta + \beta/\{t-\gamma\} - 2]}{[1/\alpha^2][\{t-\gamma\}/\beta + \beta/\{t-\gamma\} - 2] + \nu}} \left( \frac{1}{2}, \frac{\nu}{2} \right) \right], \quad t > \gamma \tag{10}$$

respectively. The  $q$ th quantile of  $T$  is obtained in an analogous way as in the ShBS case. The mean, variance and CV of  $T$  respectively are now

$$E[T] = \frac{\beta}{2} \left[ 2 + A\alpha^2 + \frac{2\gamma}{\beta} \right], \quad V[T] = \frac{\beta^2\alpha^2}{4} [4A + 5B\alpha^2] \quad \text{and} \quad CV[T] = \frac{\alpha\beta\sqrt{4A+5B\alpha^2}}{\beta[2+A\alpha^2]+2\gamma} \tag{11}$$

where  $A$  and  $B$  are as given in (9).

### 2.4. Force of Mortality

Hazard can be defined as the probability that a dangerous event that could develop into an emergency or disaster. Origin of this event can be provoked by an environmental agent that could have an adverse effect. Then, hazard is a chance

and not a real fact. This means that hazard should be evaluated as the frequency or intensity of an r.v., e.g., the DBH. A useful function in hazard analysis is the hazard rate (h.r.) or force of mortality defined as  $h_T(t) = f_T(t)/[1 - F_T(t)]$ , where  $f_T(\cdot)$  and  $F_T(\cdot)$  are the p.d.f. and c.d.f. of the r.v.  $T$ , respectively (Johnson et al. 1995). The h.r. can be interpreted as the velocity or propensity that a specific event occurs, expressed per unit of the r.v. (in general, time, but in the case of DBH is a unit of length). A characteristic of the h.r. is that it allows us to identify statistical distributions. For example, distributions with shapes similar for their p.d.f.'s could have h.r.'s which are totally different (such as is the case with the BS and Weibull distributions). As mentioned in Subsection 2.1, the BS distribution has a non-monotone h.r., because it is first increasing, until a critical point in its phase I and then it is decreasing until its stabilization at a positive constant greater than zero in its phase II. Specifically, for the BS case, if  $t$  approaches to  $\infty$ , then the h.r.  $h_T(t)$  converges to the constant  $1/[2\alpha^2\beta] > 0$ , for  $t > 0$ . Figure 1(a) shows the behavior of the BS p.d.f. for some values of the shape parameter ( $\alpha$ ). Notice that, as  $\alpha$  decreases, the shape of the BS p.d.f. is approximately symmetrical. Graphical plots for different values of the parameter  $\beta$  were not considered, because this parameter only modifies the scale. Figure 1(b) displays the behavior of the BS h.r. for some values of  $\alpha$ . Notice that, as  $\alpha$  decreases, the shape of the h.r. is approximately increasing. For a recent study of the BS- $t$  h.r., the interested reader is referred to (Azevedo et al. 2012).

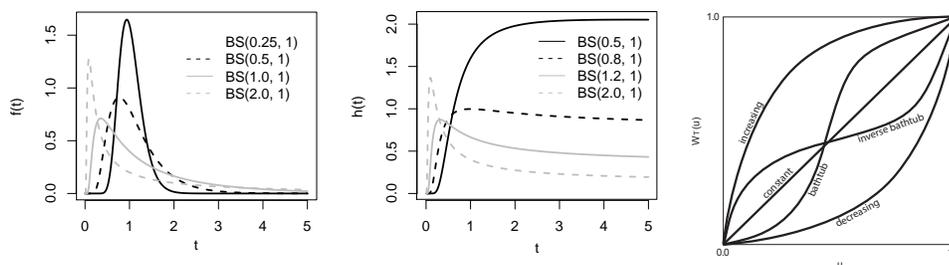


FIGURE 1: BS p.d.f. (left), BS h.r. (center) and theoretical TTT plots (right) for the indicated values.

When continuous data are analyzed (for example, DBH data) and we want to propose a distribution for modeling such data, one usually constructs a histogram. This graphical plot is an empirical approximation of the p.d.f. However, it is always convenient to look also for the h.r. of the data. The problem is that to approximate empirically the h.r. is not an easy task. A tool that is being used for this purpose is the total time on test (TTT) plot, which allows us to have an idea about the shape of the h.r. of an r.v. and, as consequence, about the distribution that the data follows. The TTT function of the r.v.  $T$  is given by  $H_T^{-1}(u) = \int_0^{F_T^{-1}(u)} [1 - F_T(y)] dy$  and its scaled version by  $W_T(u) = H_T^{-1}(u)/H_T^{-1}(1)$ , for  $0 \leq u \leq 1$ , where once again  $F_T^{-1}(\cdot)$  is the inverse c.d.f. of  $T$ . Now,  $W_T(\cdot)$  can be approximated allowing us to construct the empirical scaled TTT curve by plotting the points  $[k/n, W_n(k/n)]$ , where

$W_n(k/n) = [\sum_{i=1}^k T_{(i)} + [n - k] T_{(k)}] / \sum_{i=1}^n T_{(i)}$ , for  $k = 1, \dots, n$ , with  $T_{(i)}$  being the  $i$ th order statistic, for  $i = 1, \dots, n$ . Specifically, if the TTT plot is concave (convex), then a model with increasing (decreasing) h.r. is appropriate. Now, if the TTT plot is first concave (convex) and then convex (concave), an inverse bathtub (IBT) shaped (bathtub –BT–) h.r. must be considered. If the TTT plot is a straight line, then the exponential distribution must be used. For example, the normal distribution is in the increasing h.r. class, while the gamma and Weibull distributions admit increasing, constant and decreasing h.r.'s. However, the BS and log-normal distributions have non-monotone h.r.'s, because these are initially increasing until their change points and then decreasing (IBT shaped h.r.) to zero, in the log-normal case, or to a constant greater than zero, in the BS case. This last case must be highlighted because biological entities (such as humans, insects and trees) have h.r.'s of this type (Gavrilov & Gavrilova 2001). In Figure 1(c), we see several theoretical shapes of the TTT plot, which correspond to a particular type of h.r. (Aarset 1987).

## 2.5. Model Estimation and Checking

Parameters of the BS, ShBS, BS- $t$  and ShBS- $t$  distributions can be estimated by the maximum likelihood (ML) method adapted by a non-failing algorithm (Leiva et al. 2011). To obtain the estimates of the parameters of these distributions, their corresponding likelihood functions must be constructed using (4), (6), (8) and (10), respectively. When these parameters have been estimated, we must check goodness-of-fit of the model to the data. Distributions used for describing DBH data can be compared using model selection criteria based on loss of information such as Akaike (AIC) and Bayesian (BIC) information criteria. AIC and BIC allows us to compare models for the same model and they are given by  $AIC = -2\ell(\hat{\theta}) + 2p$  and  $BIC = -2\ell(\hat{\theta}) + p \log(n)$ , where  $\ell(\hat{\theta})$  is the logarithm of the likelihood function (log-likelihood) of the model with vector of parameters  $\theta$  evaluated at  $\theta = \hat{\theta}$ ,  $n$  is the size of the sample and  $p$  is the number of model parameters. For the case of BS, ShBS, BS- $t$  and ShBS- $t$  models, as mentioned,  $\ell(\theta)$  must be obtained by (4), (6), (8) and (10), respectively. AIC and BIC correspond to the log-likelihood function plus a component that penalizes such a function as the model has more parameters making it more complex. A model with a smaller AIC or BIC is better.

Differences between two values of the BIC are usually not very noticeable. Then, the Bayes factor (BF) can be used to highlight such differences, if they exist. Assume the data belongs to one of two possible models, according to probabilities  $P(\text{Data} \mid \text{Model 1})$  and  $P(\text{Data} \mid \text{Model 2})$ , respectively. Given probabilities  $P(\text{Model 1})$  and  $P(\text{Model 2}) = 1 - P(\text{Model 1})$ , the data produce conditional probabilities  $P(\text{Model 1} \mid \text{Data})$  and  $P(\text{Model 2} \mid \text{Data}) = 1 - P(\text{Model 1} \mid \text{Data})$ , respectively. The BF allows us to compare Model 1 (considered as correct) to Model 2 (to be contrasted with Model 1) and it is given by  $B_{12} = P(\text{Data} \mid \text{Model 1}) / P(\text{Data} \mid \text{Model 2})$ , which can be approximated by  $2 \log(B_{12}) \approx 2[\ell(\hat{\theta}_1) - \ell(\hat{\theta}_2)] - [d_1 - d_2] \log(n)$ , where  $\ell(\hat{\theta}_k)$  is the log-likelihood function for

the parameter  $\theta_k$  under the  $k$ th model evaluated at  $\theta_k = \hat{\theta}_k$ ,  $d_k$  is the dimension of  $\theta_k$ , for  $k = 1, 2$ , and  $n$  is the sample size. Notice that the above approximation is computed by subtracting the BIC value from Model 2, given by  $\text{BIC}_2 = -2\ell(\theta_2) + d_2 \log(n)$ , to the BIC value of Model 1, given by  $\text{BIC}_1 = -2\ell(\theta_1) + d_1 \log(n)$ . In addition, notice that if Model 2 is a particular case of Model 1, then the procedure corresponds to applying the likelihood ratio (LR) test. In this case,  $2 \log(B_{12}) \approx \chi_{12}^2 - \text{df}_{12} \log(n)$ , where  $\chi_{12}^2$  is the LR test statistic for testing Model 1 versus Model 2 and  $\text{df}_{12} = d_1 - d_2$  are the d.f.'s associated with the LR test, so that one can obtain the corresponding  $p$ -value from  $2 \log(B_{12}) \sim \chi^2(d_1 - d_2)$ , with  $d_1 > d_2$ . The BF is informative, because it presents ranges of values in which the degree of superiority of one model with respect to another can be quantified. An interpretation of the BF is displayed in Table 1.

TABLE 1: Interpretation of  $2 \log(B_{12})$  associated with the BF.

$2 \log(B_{12})$	Evidence in favor of Model 1
$< 0$	Negative (Model 2 is accepted)
$[0, 2)$	Weak
$[2, 6)$	Positive
$[6, 10)$	Strong
$\geq 10$	Very strong

## 2.6. Quantity and Quality of Wood

Because the DBH varies depending on the composition, density, geographic location and stand age, the diameter can be considered as an r.v. that we denote by  $T$ . As mentioned, information on the distribution of  $T$  in a forest plantation is an important element to quantify the products come from thinning and clearcutting activities. This information can help to plan the management and use of forest resources more efficiently. It is important to model the distribution of the DBH since this is the most relevant variable in determining the tree volume and then the forest production.

The forest volume quantification allows us to make decisions about the production and forest management, for example, to know when the forest should be harvested. However, the variable to maximize is diameter instead volume. Furthermore, the DBH is related to other variables such as cost of harvest, quality and product type. While the productivity is an important issue for timber industry, wood quality is also relevant in order to determine its use. Thus, volume and diameter distribution of trees determine what type of product will be obtained. For example, large diameter trees are used for saw wood and those of small diameter for pulpwood. This implies a financial analysis of forest harvest, i.e., how and when to harvest and what method to use. Studies from several types of climates and soils show trees growth as a function of the basal area. Making decisions using the forest basal area are related to pruning and thinning. These activities aim to improve tree growth and produce higher quality wood. The basal area of a tree is

the imaginary basal area at breast height (1.3 m above ground level) given by

$$B = \frac{\pi}{4} T^2$$

where  $B$  is the basal area and  $T$  the DBH.

The sum of the individual basal area of all trees in one hectare leads to the basal area per hectare. However, it is the volume which allows for the planning of various forestry activities. There are several formulae to determine the volume of logs using the mean diameter measured without bark, and the log length. Volume allows for the planning of silvicultural and harvesting activities. In general, the formula used for the volume of a tree is given by

$$V = F B H = \frac{\pi}{4} F T^2 H \quad (12)$$

where  $V$  is the tree volume,  $B$  its basal area,  $H$  its height and  $F$  the form factor, which is generally smaller than a value equal to one depending on the tree species.

## 2.7. Mortality and Tree Regeneration

The DBH is related to tree mortality, which is affected by stress factors such as light, nutrients, sunlight, temperature and water. The light and temperature can cause stress in minutes, whereas lack of water can cause stress in days or weeks. However, lack of nutrients in the soil can take months to generate stress. The mortality of a tree is similar to the material fatigue process described in Section 2.1, because the force of mortality of trees is growing rapidly in phase I, reaching a maximum and then decreases slowly until it is stabilized in phase II, which is consistent for almost all tree species.

Podlaski (2008) identified in a national park in Poland the following stress factors: (i) abiotic factors, such as severe weather (frost, hail, humidity, snow, temperature, wind), deficiency or excess of soil nutrients and toxic substances in air and soil, and (ii) biotic factors, such as bacteria (canker), fungi (dumping-off spots, root rots, rusts), insects and worms (nematodes), mycoplasma (elm phloem necrosis), parasitic plants (mistletoes) and viruses (elm mosaic). These factors caused the death of trees of the species *Abies alba*. From a theoretical point of view, the force of mortality of spruce could be more appropriately described by the h.r. of the BS distribution rather than using other distributions employed to model DBH. Podlaski (2008) indicated that mortality of spruce stand caused more openings within the stand and the canopy. Thus, with more spaces and gaps, trees of the species *Fagus sylvatica*, a kind that grows in temperate zones of the planet, tended to regenerate.

The regeneration process has been closely connected with the death of fir, whose speed in phase I also resulted in a rapid regeneration of beech, and the subsequent occurrence of understory vegetation in the stand. The decrease in the intensity of spruce mortality in phase II, as well as shading of soil by the understory, caused a gradual decrease in the intensity of the regeneration of beech. The stands generated by this process are characterized by a vertical structure of tree layers

of different heights. These layers correspond to multiple layers of canopy whose statistical distribution of the DBH is asymmetric and positively skewed, as in the BS model. Most of the spruce stands had diameters of approximately 0.15 m to 0.35 m. The interruption of the regeneration process resulted in the death of these stands, which had a DBH of less than 0.1 m. The necessary condition for the creation of stands with DBH distributions approximated by the BS model is the simultaneous death of fir at all levels of the stand with regeneration of beech, i.e., a death that considers the different forest layers and has a similar degree of seasonality in the subsequent occurrence of the understory.

### 3. Application

Next, we apply the methodology outlined in this article using real data of the DBH and a methodology based on BS models. First, we perform an exploratory data analysis (EDA) of DBH. Then, based on this EDA, we propose statistical distributions to model the DBH. We use goodness-of-fit methods to find the more suitable distribution for modeling the DBH data under analysis. Finally, we make a confirmatory analysis and furnish information that can be useful to make financial and forestry decisions.

#### 3.1. The Data Sets

The five DBH data sets to be analyzed are presented next. These data (all of them given in cm) are expressed in each case with the data frequency in parentheses and nothing when the frequency is equal to one.

**Giant paradise (*Melia azedarach* L.)** This is an exotic tree species originated from Asia and adapted to the province of Santiago del Estero, Argentina. Giant paradise produces wood of very good quality in a short time. We consider DBH data of giant paradise trees from four consecutive annual measurements collected since 1994 in 40 sites located at a stand in the Departamento Alberdi to the northwest of the province of Santiago del Estero, Argentina. Specifically, we use measurements collected at Site 7 due to the better conformation and reliability of the database (Pece et al. 2000). The data are: 16.5, 16.6, 17.8, 18.0 18.4, 18.5, 18.8, 18.9, 19.2, 19.3, 19.8, 20.3, 20.4, 20.6(2), 22.1, 22.2 23.5, 23.6, 26.7.

**Silver fir (*Abies alba*).** This is a species of tree of the pine family originated from mountainous regions in Europe. We consider DBH data of silver fir trees from 15 sites located at  wieta Katarzyna and  wiety Krzyz j forest sections of the  wietokrzyski National Park, in  wietokrzyskie Mountains (Central Poland). Specifically, we use measurements collected at Site 10 due to similar reasons to that from *Melia azedarach* (Podlaski 2008). The data are: 11(2), 12, 13, 14(5), 15(4), 16(5), 17(4), 18(4), 19(3), 20(8), 21(4), 22(3), 23(4), 24(5), 25(6), 26(5), 27(5), 28(2), 29(5), 30(2), 31(7), 32(3), 33(2), 34(4), 35, 36(2), 37(2), 39(2), 40(3), 41(2), 42, 43(2), 44(3), 46(3), 47(2), 48, 50(2), 51, 52, 53, 54, 55, 56, 57, 59, 61, 66, 70, 89, 97.

**Loblolly pine (*Pinus taeda* L.)** This variety of tree is one of several native pines at the Southeastern of the United States (US). The data set corresponds to DBH of 20 year old trees from a plantation in the Western Gulf Coast of the US (McEwen & Parresol 1991). The data are: 6.2, 6.3, 6.4, 6.6(2), 6.7, 6.8, 6.9(3), 7.0(2), 7.1, 7.2(2), 7.3(3), 7.4(4), 7.6(2), 7.7(3), 7.8, 7.9(4), 8.1(4), 8.2(3), 8.3(3), 8.4, 8.5(3), 8.6(4), 8.7, 8.8(2), 8.9(3), 9.0(4), 9.1(5), 9.5(2), 9.6, 9.8(3), 10.0(2), 10.1, 10.3.

**Ruñ (*Nothofagus alessandrii* Espinosa).** This is an endemic species of central Chile, which is at risk of extinction. This tree variety is the older species of the family of the *Fagaceae* in the South Hemisphere, i.e., these stands are the older formations in South America. The data set of DBH was collected close to the locality of Gualleco, Región del Maule, Chile (Santelices & Riquelme 2007). The data are: 16(2), 18(2), 20(2), 22, 24, 26(2), 28, 30(2), 32, 34.

**Gray birch (*Betula populifolia* Marshall).** This is a perennial species from the US that has its best growth during spring and summer seasons. Gray birch has a short life in comparison with other plant species and a rapid growth rate. During its maturity (around 20 years), gray birch reaches an average height of 10 m. The data used for this study correspond to DBH of gray birch trees that are part of a natural forest of 16 hectares located at Maine, US. This data set was chosen because its collection is reliable and the database is complete, so it allows an adequate illustration for the purpose of this study. The data are: 10.5(5), 10.6, 10.7, 10.8(3), 10.9, 11.0, 11.2, 11.3(5), 11.4, 11.5(3), 11.6(2), 11.7(3), 11.9(2), 12.0(3), 12.1(3), 12.2(2), 12.3, 12.4(3), 12.5(3), 12.6, 12.7(2), 12.8(3), 12.9(5), 13.0(7), 13.1(4), 13.2(2), 13.3(3), 13.5(2), 13.6(3), 13.7(5), 13.8(2), 14.0(3), 14.1(4), 14.2(3), 14.3, 14.4(2), 14.5(5), 14.6(3), 14.8(4), 14.9(3), 15.0, 15.1(3), 15.2, 15.3(2), 15.6(2), 15.7(2), 15.8, 15.9(2), 16.0(2), 16.1(2), 16.4, 16.5, 16.6(2), 16.7, 16.9(2), 17.0(2), 17.5(2), 17.8(2), 18.3, 18.4, 18.5, 19.2, 19.4(2), 19.9(2), 20.0, 20.3, 20.5, 21.3, 21.9, 23.1, 24.4, 26.0, 28.4, 39.3.

We call S1, S2, S3, S4 and S5 to the DBH data sets of the varieties of *Melia azedarach*, *Abies alba*, *Pinus taeda*, *Nothofagus alessandrii*, and *Betula populifolia*, respectively.

### 3.2. Exploratory Data Analysis

Table 2 presents a descriptive summary of data sets S1-S5 that includes median, mean, standard deviation (SD), CV and coefficients of skewness (CS) and kurtosis (CK), among other indicators. Figure 2 shows histograms, usual and adjusted for asymmetrical data boxplots (Leiva et al. 2011) and TTT plots for S1-S5. From Table 2 and Figure 2, we detect distributions with positive skewness, different degrees of kurtosis, increasing and IBT shaped h.r.'s and a variable number of atypical DBH data. Specifically, the TTT plot of the DBH presented in Figure 2 (fifth panel) shows precisely a h.r. as those that the tree DBH should theoretically have and that coincides with the h.r. of the BS fatigue models. In addition, minimum values for S1-S5 indicate to us the necessity for considering a shift parameter in the modeling. As a consequence, based on this EDA, the different BS models

presented in this paper seem to be good candidates for describing S1-S5, because they allow us to accommodate the different aspects detected in the EDA for these data sets. Particularly, BS- $t$  and ShBS- $t$  models allow us to accommodate atypical data in a robust statistically way. Also, BS distributions have a more appropriate h.r. to model such DBH data. This is a relevant aspect because DBH data have been widely modeled by the Weibull distribution. However, this distribution has a different h.r. to those that the tree DBH should theoretically have. Therefore, in the next section of model estimation and checking, we compare usual and shifted BS and Weibull models by means of a goodness-of-fit analysis in order to valuate whether this theoretical aspect is validated by the data or not.

TABLE 2: Descriptive summary of DBH for the indicated data set

Set	Median	Mean	SD	CV	CS	CK	Range	Minimum	Maximum	$n$
S1	19.55	20.09	2.53	12.58%	0.82	3.20	10.20	16.50	26.70	20
S2	27.00	30.68	14.85	48.42%	1.52	6.33	86.00	11.00	97.00	134
S3	8.20	8.19	1.01	12.37%	0.05	2.16	4.10	6.20	10.30	75
S4	24.00	24.00	5.95	24.80%	0.14	1.50	18.00	16.00	34.00	15
S5	13.70	14.54	3.61	24.85%	5.89	13.97	28.80	10.50	39.30	160

### 3.3. Model Estimation and Checking

As mentioned, the parameters of the BS, ShBS, BS- $t$ , ShBS- $t$  distributions can be estimated by the ML method adapted by a non-failing algorithm (Leiva et al. 2011). The estimation of the parameters of the BS distributions, as well as those of the usual and shifted Weibull distributions (as comparison), for S1-S5 are summarized in Table 3 together with the negative value of the corresponding log-likelihood function. In addition to the model selection criteria (AIC and BIC) presented in Section 2.1, the fit of the model to SI-S5 can be checked using the Kolmogorov-Smirnov test (KS). This test compares the empirical and theoretical c.d.f.'s (in this case of the BS and Weibull models). The p-values of the KS test, as well as the values of AIC, BIC and  $2\log(B_{12})$  are also provided in Table 3. Based on the KS test and BF results presented in Table 3, we conclude that the BS distributions fit S1-S5 better than Weibull distributions. All this information supports the theoretical justification given in Section 2.

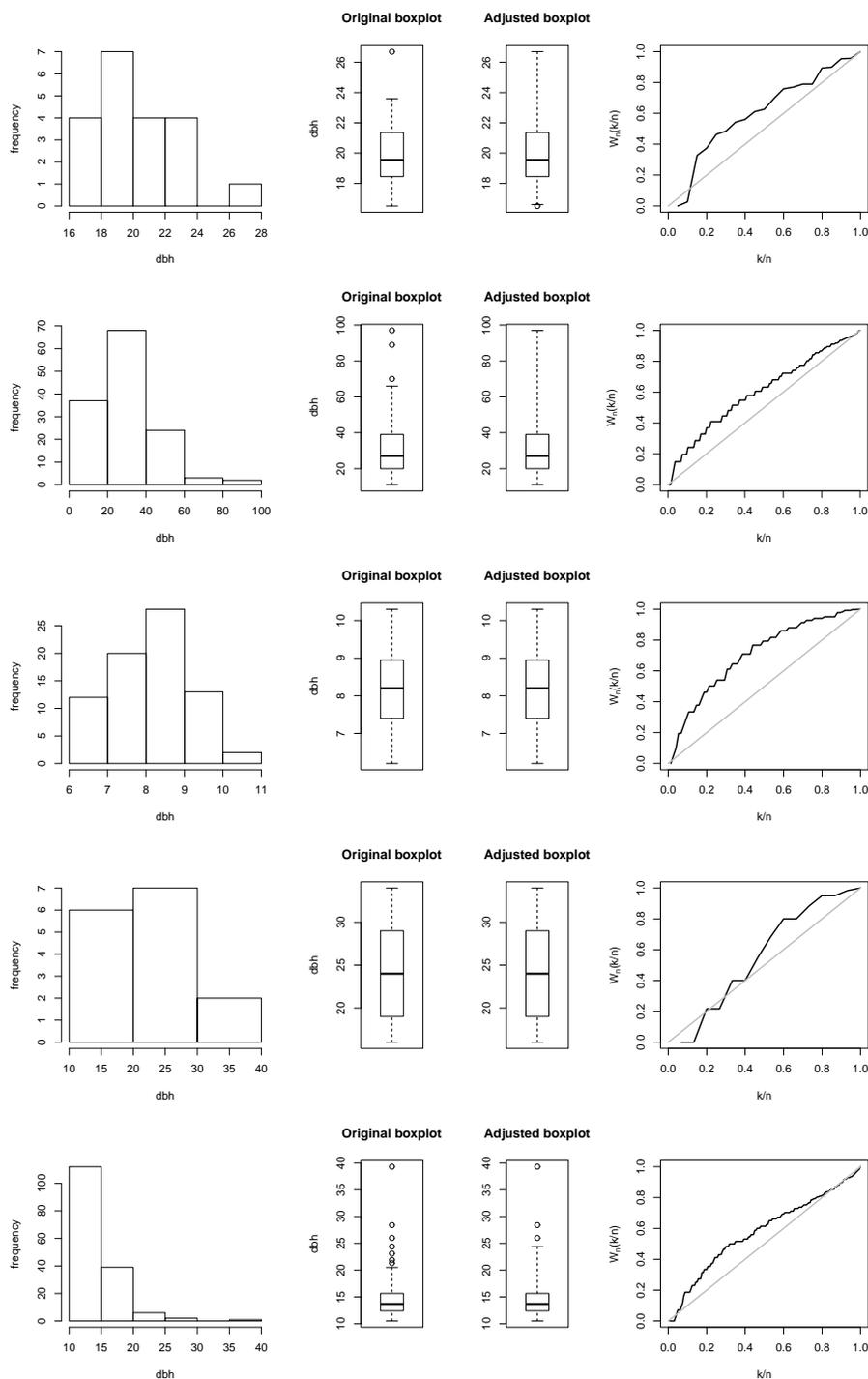


FIGURE 2: Histograms, usual and adjusted boxplots and TTT plots for S1 (first panel) to S5 (fifth panel).

TABLE 3: Indicators for the indicated data set and distribution.

Indicator	BS	BS- <i>t</i>	ShBS	ShBS- <i>t</i>	ShWeibull	Weibull
S1						
$\hat{\alpha}$	0.118	0.117	0.374	0.443	1.587	7.736
$\hat{\beta}$	19.950	19.934	6.161	3.260	4.810	21.228
$\hat{\nu}$	-	87	-	1	-	-
$\hat{\gamma}$	-	-	13.498	16.498	15.900	-
$-\ell(\hat{\theta})$	45.534	45.533	44.752	43.367	44.811	48.565
AIC	95.069	97.067	95.503	94.733	95.622	101.131
BIC	97.061	100.053	98.490	98.716	98.609	103.121
$2 \log(B_{12})$	-	2.992	1.429	1.655	1.548	6.060
KS p-value	0.806	0.829	0.986	0.963	0.882	0.385
S2						
$\hat{\alpha}$	0.452	0.448	0.590	0.588	1.440	2.193
$\hat{\beta}$	27.840	27.803	21.301	21.171	22.453	34.760
$\hat{\nu}$	-	100	-	100	-	-
$\hat{\gamma}$	-	-	5.666	5.778	10.358	-
$-\ell(\hat{\theta})$	525.820	525.889	524.255	524.306	524.772	540.361
AIC	1055.640	1057.777	1054.511	1056.612	1055.544	1084.721
BIC	1061.436	1066.472	1063.204	1068.203	1064.238	1090.518
$2 \log(B_{12})$	-	5.036	1.768	6.768	2.802	29.082
KS p-value	0.899	0.912	0.959	0.815	0.828	0.129
S3						
$\hat{\alpha}$	0.124	0.123	0.124	0.124	2.514	8.952
$\hat{\beta}$	8.125	8.127	8.125	8.127	2.635	8.636
$\hat{\nu}$	-	100	-	100	-	-
$\hat{\gamma}$	-	-	0.000	0.000	5.850	-
$-\ell(\hat{\theta})$	107.038	107.180	107.038	107.180	105.798	108.609
AIC	218.076	220.360	218.076	222.360	217.596	221.219
BIC	222.711	227.312	222.711	227.312	224.548	225.853
$2 \log(B_{12})$	-	4.601	-	4.601	1.837	3.142
KS p-value	0.876	0.874	0.876	0.874	0.918	0.840
S4						
$\hat{\alpha}$	0.245	0.2445	0.394	0.585	2.837	4.685
$\hat{\beta}$	23.298	23.304	14.625	9.207	16.568	26.282
$\hat{\nu}$	-	100	-	1	-	-
$\hat{\gamma}$	-	-	8.240	15.995	9.300	-
$-\ell(\hat{\theta})$	47.327	47.377	47.292	44.460	47.098	47.515
AIC	98.656	100.754	100.584	96.920	100.195	99.0294
BIC	100.070	102.878	102.708	99.752	104.084	100.446
$2 \log(B_{12})$	0.319	3.126	2.956	-	4.332	0.694
KS p-value	0.933	0.934	0.858	0.936	0.894	0.852
S5						
$\hat{\alpha}$	0.208	0.151	0.727	0.563	1.502	3.467
$\hat{\beta}$	14.230	13.817	3.774	4.232	4.749	15.920
$\hat{\nu}$	-	4	-	8	-	-
$\hat{\gamma}$	-	-	9.761	9.439	10.180	-
$-\ell(\hat{\theta})$	399.776	389.438	380.330	378.912	386.075	448.921
AIC	803.553	816.853	766.659	765.826	778.152	901.842
BIC	809.702	794.102	775.886	778.125	787.376	907.992
$2 \log(B_{12})$	33.817	18.216	-	2.239	11.490	132.107
KS p-value	0.052	0.400	0.530	0.773	0.467	< 0.001

Due to space limitations, in order to visualize the model fit to the DBH data, we only focus on S5. In addition, we only depict three plots corresponding to the shifted versions of the BS, BS- $t$  and Weibull distributions, which are those that fit the data better. Comparison between the empirical (gray line) and ShBS, ShBS- $t$  and ShWeibull theoretical (black dots) c.d.f.'s are shown in Figure 3. Histograms with the estimated ShBS, ShBS- $t$  and ShWeibull p.d.f. curve are shown in Figure 4. Probability plots with “envelopes” based on the BS, BS- $t$  and Weibull distributions for S5 are shown in Figure 5. The term “envelope” is a band for the probability plot built by means of a simulation process that facilitates the adjustment visualization. For example, for the BS distribution, this “envelope” is built using an expression given in (3). From Figure 5, we can see the excellent fit that the ShBS- $t$  model provides to S5 and the bad fit provided by the ShWeibull model. Then, once the ShBS- $t$  model has been considered as the most appropriate within the proposed distributions to model S5, we provide information that can be useful to make economical and forestry decisions based on this model and the methodology given in this study.

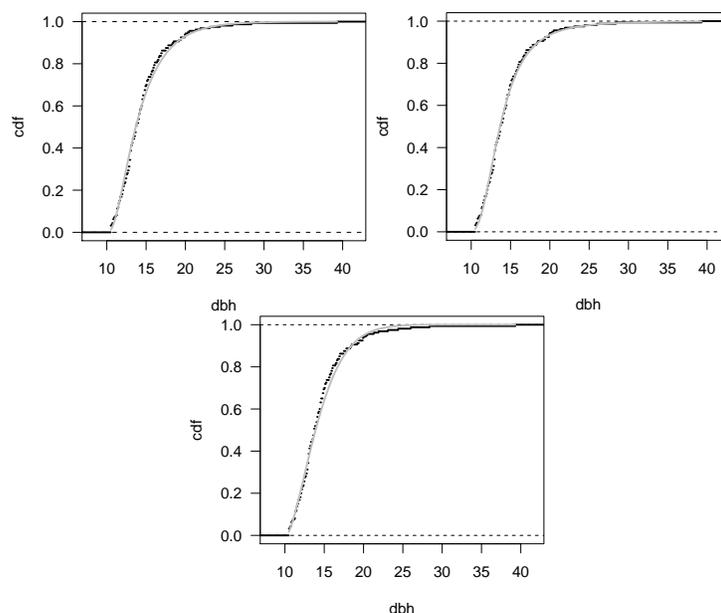


FIGURE 3: Empirical (bold) and theoretical (gray) c.d.f.'s for S5 using the ShBS, ShBS- $t$  and ShWeibull distributions.

### 3.4. Financial Evaluation

We select S5 for carrying out a financial analysis. In this case, the ShBS- $t$  distribution is considered as the best model. Then, we propose a forest production problem to illustrate the methodology presented in this article. Once the ShBS- $t$  model parameters are estimated, we determine the mean volume per tree in a

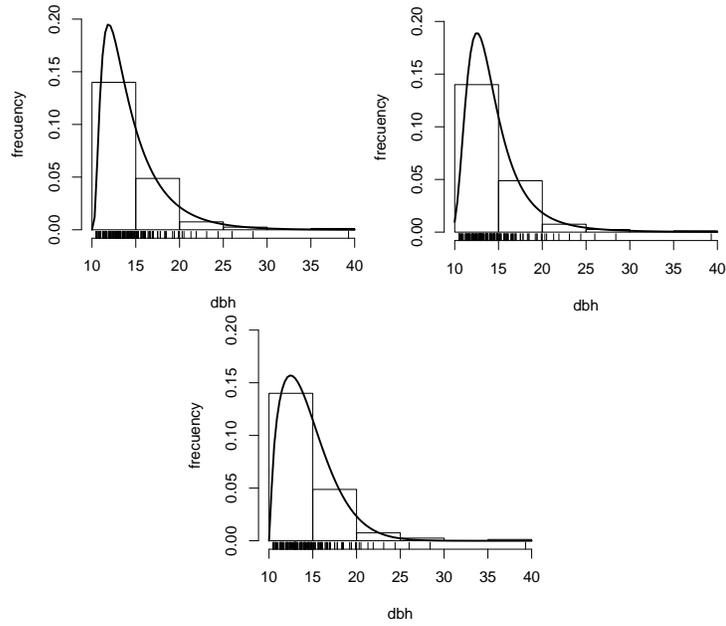


FIGURE 4: Histogram with ShBS, ShBS- $t$  and ShWeibull p.d.f.'s for S5.

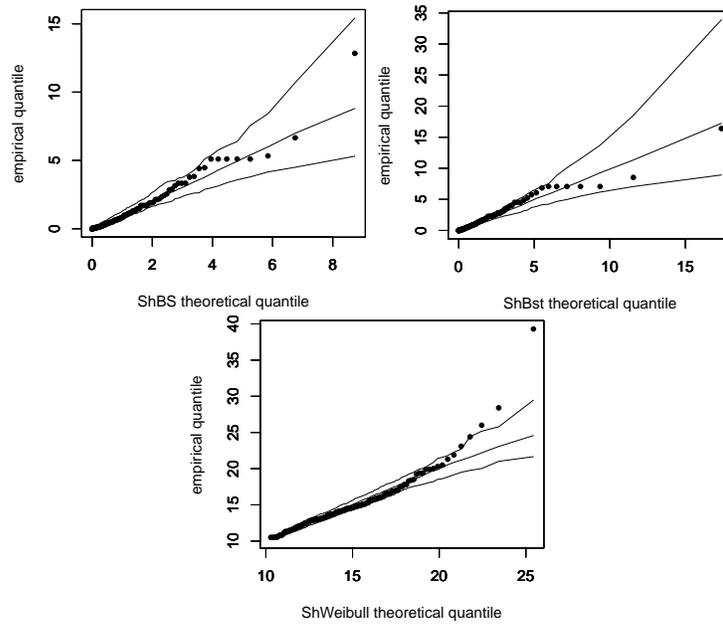


FIGURE 5: Probability plots with envelopes for S5 using the ShBS, ShBS- $t$  and ShWeibull distributions.

stand by using (12) that leads to  $E[V] = (250/3) \pi E[T^2]$ , recalling that  $T$  is the DBH,  $H$  the known height of the tree equal to 10 m (1000 cm, because the data are expressed in cm) and  $F$  the form factor being it equal to 1/3 due to the birch case, which has conical shape, with the DBH equivalent to the diameter at the base of the cone. Using the expected value and variance of  $T$  given in (11), we get the expected volume as

$$E[V] = \frac{250}{3} \pi \left[ \beta^2 \left\{ 1 + \alpha^2 \left( 2A + \frac{5}{4} B \alpha^2 + \frac{A^2 \alpha^2}{4} \right) \right\} + \gamma \{ \gamma + \beta(2 + A\alpha^2) \} \right]$$

The stand considered in this study only produces native wood that can be sold to sawmills at a price of US\$250 (international price in US dollars) per cubic meter. This stand of Maine, US, had in the spring of 2004 an amount of 3327 trees, of which 160 (4.8%) were of the gray birch variety. Thus, the estimated expected economical value for gray birch wood of this forest (stand) based on the ShBS- $t$  model is

$$\text{US\$}0.25 \times \widehat{E[V]} \times 160 = 10000 \pi \left[ \widehat{\beta}^2 \left\{ 1 + \widehat{\alpha}^2 \left( 2A + \frac{5}{4} B \widehat{\alpha}^2 + \frac{A^2 \widehat{\alpha}^2}{4} \right) \right\} + \widehat{\gamma} \{ \widehat{\gamma} + \widehat{\beta}(2 + A\widehat{\alpha}^2) \} \right] \quad (13)$$

being its estimation based on the proposed methodology and S5 of US\$7,342,267.

## 4. Concluding Remarks

In this paper, we have presented, developed, discussed and applied a statistical methodology based on Birnbaum-Saunders distributions to address the problem of managing forest production. Specifically, we have linked a fatigue model to a forestry model through Birnbaum-Saunders distributions. This linkage has been possible because the hazard rate of this distribution has two clearly marked phases that coincide with the force of mortality of trees. This mortality is related to the diameter at breast height of trees. We have modeled the distribution of this diameter because this variable is the most relevant in determining the basal area of a tree. For its part, the basal area allows the volume of a tree to be determined setting thus the production of a forest. Finally, we have shown the applicability of this model using five real data sets, obtaining for one of them financial information that may be valuable in forest decision making. The unpublished data used in the economical evaluation corresponded to the diameter at breast height of 10 m height mature gray birch trees collected in 2004, which are part of the inventory of a natural forest of area 16 hectares of different species located at Maine, US.

## Acknowledgements

The authors wish to thank the Editor-in-Chief, Dr. Leonardo Trujillo, and two anonymous referees for their constructive comments on an earlier version of

this manuscript which resulted in this improved version. Also, we wish to thank the colleagues Celia Benítez (Argentina), Rafal Podlaski (Poland) and Rómulo Santelices (Chile) for kindly providing us some of the data sets analyzed in this paper. This research work was partially supported by grant FONDECYT 1120879 from the Chilean government and by Universidad de Talca, Chile.

[Recibido: diciembre de 2011 — Aceptado: junio de 2012]

## References

- Aarset, M. V. (1987), 'How to identify a bathtub hazard rate', *IEEE Transaction on Reliability* **36**, 106–108.
- Azevedo, C., Leiva, V., Athayde, E. & Balakrishnan, N. (2012), 'Shape and change point analyses of the Birnbaum-Saunders- $t$  hazard rate and associated estimation', *Computational Statistics and Data Analysis* **56**, 3887–3897.
- Bailey, R. & Dell, T. (1973), 'Quantifying diameter distributions with the Weibull function', *Forest Science* **19**, 97–104.
- Birnbaum, Z. & Saunders, S. (1968), 'A probabilistic interpretation of miner's rule', *SIAM Journal of Applied Mathematics* **16**, 637–652.
- Birnbaum, Z. & Saunders, S. (1969), 'A new family of life distributions', *Journal of Applied Probability* **6**, 319–327.
- Bliss, C. & Reinker, K. (1964), 'A lognormal approach to diameter distributions in even-aged stands', *Forest Science* **10**, 350–360.
- Borders, B., Souter, R., Bailey, R. & Ware, K. (1987), 'Percentile-based distributions characterize forest stand tables', *Forest Science* **33**, 570–576.
- Clutter, J. & Bennett, F. (1965), Diameter distributions in old-field slash pine plantation, Report 13, US Forest Service.
- Díaz-García, J. & Leiva, V. (2005), 'A new family of life distributions based on elliptically contoured distributions', *Journal of Statistical Planning and Inference* **128**, 445–457. (Erratum: *Journal of Statistical Planning and Inference*, **137**, 1512–1513).
- Ferreira, M., Gomes, M. & Leiva, V. (2012), 'On an extreme value version of the Birnbaum-Saunders distribution', *Revstat-Statistical Journal* **10**, 181–210.
- García-Güemes, C., Cañadas, N. & Montero, G. (2002), 'Modelización de la distribución diamétrica de las masas de *Pinus pinea* de Valladolid (España) mediante la función Weibull', *Investigación Agraria-Sistemas y Recursos Forestales* **11**, 263–282.
- Gavrilov, L. & Gavrilova, N. (2001), 'The reliability theory of aging and longevity', *Journal of Theoretical Biology* **213**, 527–545.

- Guiraud, P., Leiva, V. & Fierro, R. (2009), 'A non-central version of the Birnbaum-Saunders distribution for reliability analysis', *IEEE Transaction on Reliability* **58**, 152–160.
- Hafley, W. & Schreuder, H. (1977), 'Statistical distributions for fitting diameter and height data in even-aged stands', *Canadian Journal of Forest Research* **7**, 481–487.
- Johnson, N., Kotz, S. & Balakrishnan, N. (1995), *Continuous Univariate Distributions*, Vol. 2, Wiley, New York.
- Leiva, V., Athayde, E., Azevedo, C. & Marchant, C. (2011), 'Modeling wind energy flux by a birnbaum-saunders distribution with unknown shift parameter', *Journal of Applied Statistics* **38**, 2819–2838.
- Leiva, V., Barros, M., Paula, G. & Sanhueza, D. (2008), 'Generalized Birnbaum-Saunders distributions applied to air pollutant concentration', *Environmetrics* **19**, 235–249.
- Leiva, V., Sanhueza, A. & Angulo, J. (2009), 'A length-biased version of the Birnbaum-Saunders distribution with application in water quality', *Stochastic Environmental Research and Risk Assessment* **23**, 299–307.
- Leiva, V., Vilca, F., Balakrishnan, N. & Sanhueza, A. (2010), 'A skewed sinh-normal distribution and its properties and application to air pollution', *Communications in Statistics - Theory and Methods* **39**, 426–443.
- Lenhart, J. & Clutter, J. (1971), Cubic foot yield tables for old-field loblolly pine plantations in the Georgia Piedmont, Report 22, US Forest Service.
- Li, F., Zhang, L. & Davis, C. (2002), 'Modeling the joint distribution of tree diameters and heights by bivariate generalized Beta distribution', *Forest Science* **48**, 47–58.
- Little, S. (1983), 'Weibull diameter distributions for mixed stands of western conifers', *Canadian Journal of Forest Research* **13**, 85–88.
- Maltamo, M., Puumalainen, J. & Päivinen, R. (1995), 'Comparison of beta and Weibull functions for modelling basal area diameter in stands of *Pinus sylvestris* and *Picea abies*', *Scandinavian Journal of Forest Research* **10**, 284–295.
- Marchant, C., Leiva, V., Cavieres, M. & Sanhueza, A. (2013), 'Air contaminant statistical distributions with application to PM10 in Santiago, Chile', *Reviews of Environmental Contamination and Toxicology* **223**, 1–31.
- McEwen, R. & Parresol, B. (1991), 'Moment expressions and summary statistics for the complete and truncated Weibull distribution', *Communications in Statistics - Theory and Methods* **20**, 1361–1372.
- McGee, C. & Della-Bianca, L. (1967), Diameter distributions in natural yellow-poplar stands, Report 25, US Forest Service.

- Meyer, H. (1952), 'Structure, growth, and drain in balanced uneven-aged forests', *Journal of Forestry* **50**, 85–92.
- Miner, M. A. (1945), 'Cumulative damage in fatigue', *Journal of Applied Mechanics* **12**, 159–164.
- Nelson, T. (1964), 'Diameter distribution and growth of loblolly pine', *Journal of Applied Mechanics* **10**, 105–115.
- Palahí, M., Pukkala, T. & Trasobares, A. (2006), 'Modelling the diameter distribution of *Pinus sylvestris*, *Pinus nigra* and *Pinus halepensis* forest stands in Catalonia using the truncated Weibull function', *Forestry* **79**, 553–562.
- Pece, M., de Benítez, C. & de Galíndez, M. (2000), 'Uso de la función Weibull para modelar distribuciones diamétricas en una plantación de *Melia azedarach*', *Revista Forestal Venezolana* **44**, 49–52.
- Podlaski, R. (2006), 'Suitability of the selected statistical distribution for fitting diameter data in distinguished development stages and phases of near-natural mixed forests in the Swietokrzyski National Park (Poland)', *Forest Ecology and Management* **236**, 393–402.
- Podlaski, R. (2008), 'Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution', *Canadian Journal of Forest Research* **18**, 518–527.
- Rennolls, K., Geary, D. & Rollinson, T. (1985), 'Characterizing diameter distributions by the use of the Weibull distribution', *Forestry* **58**, 57–66.
- Santelices, R. & Riquelme, M. (2007), 'Forest mensuration of *Nothofagus alessandri* of Coipué provenance', *Bosque* **28**, 281–287.
- Schmelz, D. & Lindsey, A. (1965), 'Size-class structure of old-growth forests in Indiana', *Forest Science* **11**, 258–264.
- Schreuder, H. & Hafley, W. (1977), 'A useful bivariate distribution for describing stand structure of tree heights and diameters', *Biometrics* **33**, 471–478.
- Vilca, F. & Leiva, V. (2006), 'A new fatigue life model based on the family of skew-elliptical distributions', *Communications in Statistics - Theory and Methods* **35**, 229–244.
- Vilca, F., Santana, L., Leiva, V. & Balakrishnan, N. (2011), 'Estimation of extreme percentiles in Birnbaum-Saunders distributions', *Computational Statistics and Data Analysis* **55**, 1665–1678.
- Wang, M. & Rennolls, K. (2005), 'Tree diameter distribution modelling: introducing the logit-logistic distribution', *Canadian Journal of Forest Research* **35**, 1305–1313.
- Zutter, B., Oderwald, R., Murphy, P. & Farrar, R. (1986), 'Characterizing diameter distributions with modified data types and forms of the Weibull distribution', *Forest Science* **32**, 37–48.