

## Improved estimation of the lower percentiles of material properties

David J. Edwards · Frank M. Guess ·  
Timothy M. Young

Received: 4 November 2009 / Published online: 6 June 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Estimating lower percentiles in reliability for medium-density fiberboard is an important issue for manufacturers for better assessing and improving manufacturing processes, plus for guiding better product warranties while seeking lower costs. Since data may be sparse or costly in the lower tails, estimation of these percentiles may be difficult. Bootstrapping provides a helpful solution for interval estimation of lower percentiles when other approaches fail or are not as realistic. This computer intensive resampling technique estimates more accurately the true standard error of any population parameter, not just percentiles. Bootstrapping can be used for parametric models or indeed nonparametric settings when parametric models are not appropriate. This paper shows the usefulness of bootstrap methods to better assess the key quality metric of internal bond (IB or tensile strength) of medium-density fiberboard (MDF) in the critical lower percentiles when data are limited.

---

D. J. Edwards  
Department of Statistical Sciences and Operations Research,  
Virginia Commonwealth University, 4130 Harris Hall, Richmond,  
VA 23284, USA

F. M. Guess  
Department of Statistics, Operations, and Management Science,  
University of Tennessee, 335 Stokely Management Center,  
Knoxville, TN 37996-0532, USA

T. M. Young (✉)  
Forest Products Center, University of Tennessee, 2506 Jacob Drive,  
Knoxville, TN 37996-4570, USA  
e-mail: tmyoung1@utk.edu

## Introduction

There are many ways to measure reliability of a component subsystem or system product being manufactured. Compare the classic reliability references of Barlow and Proschan (1975, 1981) plus the more recent Kuo et al. (1998, 2000), and Meeker and Escobar (1998, 2004). Often, some key reliability measures are the mean or median time to failure (Hoffmeyer and Sorensen 2007). Kim and Kuo (2003) stress the importance of percentiles in optimizing system life in contrast to other classical approaches, see also Prasad et al. (2001). These lower percentiles may be of critical importance to manufacturers of engineered wood panels as such percentiles may represent product failure (Steiger and Arnold 2009).

In this article, bootstrapping methods as a useful approach to understanding reliability of manufactured medium-density fiberboard (MDF) are discussed. This study is an outcome of Edwards (2004) and builds upon the study of wood plastics composites discussed in Young et al. (2008). Bootstrapping's versatility allows this approach to be used on a wide range of engineering and manufacturing settings where standard approaches might yield misleading numbers.

In numerous reliability studies, it is of particular interest to estimate percentiles. In particular, interest usually lies in the estimation of the lower percentiles. These lower numbers are helpful for warranty analysis, understanding early failures during normal usage, improving the specification limits, reducing manufacturing costs, and avoiding costly product failure claims.

In this study, the authors focus on the needs of estimating percentiles of internal bond (IB) strengths of MDF measured in kilopascal (kPa), but the estimation procedure applies much more generally to various manufacturing settings, lifetimes, service response times, repair times, or any kind of response time (time to assemble a product, etc.) for improving reliability by more realistic assessment of uncertainty.

To be able to say that improvements have been made, one must be able to measure reliability expressed in percentiles that allow for statistical uncertainty inherent in real data. Knowing when to trust confidence intervals and when not to trust them are crucial for engineers and technical managers (Moses et al. 2003).

Historically, the problem of estimating percentiles was not in finding point estimators, but in finding standard errors and thus confidence intervals of percentiles. Serfling (1980) thoroughly and superbly examines the asymptotic distribution of the sample quantile. In particular, under mild requirements (i.e., smoothness of the distribution function), the sample quantiles are asymptotically normal. This is a useful result since by possessing asymptotic normality, asymptotic normal confidence intervals for the  $p$ th quantile can be constructed. Meeker and Escobar (1998) discuss the construction of such intervals for the location-scale distributions used commonly in reliability data analysis (i.e., normal, lognormal, Weibull). In particular, an asymptotic normal confidence interval for  $t_p$  is given by:

$$\hat{t}_p \pm z_{1-\alpha/2} \hat{\delta}_{\hat{t}_p} \quad (1)$$

where  $\hat{t}_p$  is the estimated  $p$ th quantile, and  $\hat{\delta}_{\hat{t}_p}$  is the standard error of the estimate approximated by:

$$\hat{s}_{i_p} = \hat{t}_p \left\{ \text{Var}(\hat{\mu}) + 2\Phi^{-1}(p)\text{Cov}(\hat{\mu}, \hat{\sigma}) + [\Phi^{-1}(p)]^2 \text{Var}(\hat{\sigma}) \right\}^{1/2}. \quad (2)$$

Equation (2) is obtained using the delta method;  $\hat{\mu}$  and  $\hat{\sigma}$  are the maximum likelihood estimates (MLEs) of the location and scale parameters, respectively, and  $\Phi^{-1}$  represents the inverse of the cumulative standardized location-scale distribution of interest.  $\text{Var}(\hat{\mu})$ ,  $\text{Var}(\hat{\sigma})$ , and  $\text{Cov}(\hat{\mu}, \hat{\sigma})$  are obtained from the inverse of the observed information matrix.

When the sample size is sufficiently large, the asymptotic normal intervals can provide reasonable approximations. Even though these intervals are approximations, they are usually good enough for practice, provided the sample size is indeed large enough. However, data may not be plentiful, and in many manufacturing settings, parametric assumptions may be suspect or actually invalid, leading to a higher risk of inaccurate results. Asymptotic intervals are often criticized for not being as realistic for small or even moderate sample sizes. Bootstrapping provides an alternative strategy that can realistically inform the practitioner by a more accurate assessment of the variability inherent in a system or process.

## Methods

### MDF manufacturer dataset

The IB data are from a MDF manufacturer in North America and are sorted based on three key characteristics: density ( $\text{kg/m}^3$ ), thickness (mm), and width (mm). These three characteristics differentiate the MDF produced by the manufacturer for various applications. Since MDF in this particular study was produced in continuous length of sheets, length was not a crucial variable for the purposes here as indicated by the manufacturer. For the purpose of analysis, the MDF was separated into two main groups: Group I- standard density and Group II- high density. The *high density* type is MDF with densities of 753–769  $\text{kg/m}^3$ . The *standard density* type is MDF with densities of 721–737  $\text{kg/m}^3$ .

Since there were a number of MDF product types within each group produced by the manufacturer, two types were selected for a more detailed analysis: in particular, Type 1 (737  $\text{kg/m}^3$ , 15.9 mm thick, 1,550 mm wide) from Group I and Type 5 (769  $\text{kg/m}^3$ , 15.9 mm thick, 1,550 mm wide) from Group 2. These two MDF types were chosen since they are commonly used MDF product types and in order to allow for useful comparisons. Type 1, which had the most sales of the producer, had  $n = 396$  observations while Type 5, a higher valued product, had  $n = 74$  observations. This illustrates two extremes in the data.

### Bootstrap methods and confidence intervals

The fundamental idea behind the bootstrap is that the empirical bootstrap distribution provides an approximation to the theoretical sampling distribution of the statistic of interest. Meeker and Escobar (1998) contend that bootstrap methods,

“when used properly, can be expected to be more accurate than the normal-approximation methods and competitive with the likelihood-based methods.” Bootstrapping is a computer intensive statistical method where the basic idea is to simulate the sampling process a specified (usually large) number of times and obtain an approximate sampling distribution of interest. This empirical bootstrap distribution is then used to acquire characteristics (i.e., standard error, bias estimates, confidence intervals) with regard to the population parameter; see Chernick (1999) which is an excellent book on many bootstrap methods and their applications. Efron and Tibshirani (1993) provide an excellent introduction to the fundamental concepts and applications of bootstrapping. Also, DiCiccio and Efron (1996) are devoted to the construction of bootstrap confidence intervals.

### *Bootstrap sampling methods*

This study begins with the fully nonparametric bootstrap and adopts the notation of Martinez and Martinez (2002). In general, the basic nonparametric bootstrap procedure can be summarized as follows. For a given data set,  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  of size  $n$ , a population parameter is estimated nonparametrically, say  $\theta$ , by  $\hat{\theta}$ . For instance, the  $p$ th quantile is estimated as the  $(p/100)(n + 1)$ st observation in  $\mathbf{x}$ . It is then sampled with replacement (i.e., a unit is drawn from and then returned to the sample allowing for the possibility of being drawn again, repeating this process many times using simulation) from the original data set to obtain a bootstrap sample of the same size  $n$  as the original data denoted by  $\mathbf{x}^{*b} = (x_1^{*b}, x_2^{*b}, \dots, x_n^{*b})$ . This resampling with replacement is usually done a large number of times,  $B$ . For each bootstrap sample, a new estimate of  $\theta$  is calculated, denoted by  $\hat{\theta}^{*b}$  where  $b$  stands for the  $b$ th bootstrap estimate. The empirical bootstrap distribution of  $\hat{\theta}^*$ , is defined and used as an estimate to the true sampling distribution of  $\hat{\theta}$ . This method of sampling is helpful since it has the advantage of no distributional assumptions.

The completely parametric bootstrap, which requires the assumption of a parametric distribution, is described briefly in Efron and Tibshirani (1993), Meeker and Escobar (1998), and Chernick (1999). Meeker and Escobar (1998) point out that the parametric bootstrap has a disadvantage in reliability data problems. That is, the complete censoring process must be specified given that data from an assumed parametric distribution are simulated. This may seem to be unproblematic in simple examples where such specification is easy. For example, the strength data is complete. However, this can be more difficult for complicated systematic or random censoring. Thus, the fully parametric form of sampling is not emphasized in this paper.

As an alternative method, Meeker and Escobar (1998) describe and illustrate applications of a “nonparametric” bootstrap sampling method for parametric inference, which is denoted, for the sake of simplicity, as NBSP, for nonparametric bootstrap sampling for parametric models. This sampling scheme does require parametric assumptions. However, rather than simulating random variates from an assumed parametric distribution, the authors sample with replacement from the original data. For each bootstrap sample of size  $n$ , MLEs are obtained based on the assumed parametric model. These MLEs are used to estimate the population

parameter of interest and form the bootstrap distribution. For instance, a parametric estimate of the  $p$ th percentile is given by  $\hat{t}_p = \exp[\hat{\mu} + \Phi^{-1}(p)\hat{\sigma}]$ , which requires the MLEs  $\hat{\mu}$  and  $\hat{\sigma}$ .

### *Bootstrap confidence intervals*

Different algorithms/methods are available for constructing bootstrap confidence intervals for population parameters. The authors emphasize the standard normal bootstrap confidence interval, bootstrap percentile interval, and bias-corrected bootstrap percentile interval. Much of the theoretical details are omitted. For those interested in the theoretical underpinnings and additional topics see, among others, Efron and Tibshirani (1993), DiCiccio and Efron (1996), and Davison and Hinkley (1997). The standard bootstrap confidence interval is given by:

$$[\hat{\theta} - z^{(\alpha/2)}s_{\hat{\theta}}, \hat{\theta} + z^{(1-\alpha/2)}s_{\hat{\theta}}] \quad (2.1)$$

where  $\hat{s}_{\hat{\theta}}$  is obtained by computing the standard deviation of the B bootstrap estimates of  $\theta$  and  $z^{(\alpha/2)}$  is the  $\alpha/2$ th quantile of the standard normal distribution. The necessary steps are provided in Algorithm 1 below. The algorithms that follow are given for the fully nonparametric case with the NBSP method alternatives shown in parentheses.

*Algorithm 1: standard bootstrap confidence interval:*

*Step 1.* From the original sample of size  $n$ , estimate the parameter(s) of interest (denoted by  $\hat{\theta}$ ). (For the NBSP method, obtain MLEs of the assumed parametric distribution and use them to estimate the parameter(s) of interest.)

*Step 2.* Sample with replacement from the original sample to create a bootstrap sample of size  $n$ .

*Step 3.* Estimate the parameter(s) of interest from the bootstrap sample to obtain  $\hat{\theta}^{*b}$ . (For the NBSP method, calculate the MLE's of the assumed parametric distribution based on the bootstrap sample and use them to estimate the parameter(s).)

*Step 4.* Repeat steps 2 and 3 a pre-specified  $B \geq 1,000$  times to form the bootstrap distribution.

*Step 5.* Calculate the standard deviation of the B bootstrap estimates ( $\hat{s}_{\hat{\theta}}$ ) and use this to estimate the standard error,  $s_{\hat{\theta}}$ .

*Step 6.* Use (2.1) to obtain the confidence interval.

Perhaps one of the most obvious ways to construct a confidence interval is to base it on the quantiles of the bootstrap distribution of estimates, which is known as the percentile method.

*Algorithm 2: bootstrap percentile confidence interval:*

*Steps 1, 2, 3, and 4.* Same as in Algorithm 1.

*Step 5.* Order the B bootstrap estimates,  $\hat{\theta}^{*b}$ .

*Step 6.* Determine the  $\alpha/2$ th and  $1 - (\alpha/2)$ th quantiles of the distribution of  $\hat{\theta}^*$  denoted by  $\hat{\theta}^{*(\alpha/2)}$  and  $\hat{\theta}^{*(1-\alpha/2)}$ , respectively.

*Step 7.* Form the  $1 - \alpha$  confidence interval as  $[\hat{\theta}^{*(\alpha/2)}, \hat{\theta}^{*(1-\alpha/2)}]$ .

Though the percentile method is easy to implement, Chernick (1999) points out that the percentile method works well if exactly 50% of the bootstrap distribution is less than  $\hat{\theta}$  which certainly might not hold and “in the case of small samples, the percentile method does not work well.” Fortunately, there are methods that help improve on the percentile method.

The bias-corrected percentile interval (or BC) was introduced in Efron (1981) and discussed further in Efron (1987). A bias-correction constant is defined as the amount of difference between the median of the bootstrap estimates  $\hat{\theta}^{*b}$  and the estimate,  $\hat{\theta}$ , from the original sample. Explicitly, the estimate of the bias-correction constant, denoted by  $\hat{z}_0$ , is defined as:

$$\hat{z}_0 = \Phi_{\text{NOR}}^{-1} \left( \frac{\#(\hat{\theta}^{*b} < \hat{\theta})}{B} \right) \quad (2.2)$$

where  $\Phi_{\text{NOR}}^{-1}$  represents the inverse cumulative standard normal distribution and # means “number of”. Then, a  $100(1 - \alpha)\%$  BC confidence interval for  $\theta$  is given by:

$$[\hat{\theta}^{*(\alpha_1)}, \hat{\theta}^{*(\alpha_2)}] \quad (2.3)$$

where  $\alpha_1$  and  $\alpha_2$  are the new quantities on which to base the percentile confidence interval endpoints. These quantities are defined as:

$$\alpha_1 = \Phi_{\text{NOR}}(2\hat{z}_0 + z^{(\alpha/2)}) \quad (2.4)$$

and

$$\alpha_2 = \Phi_{\text{NOR}}(2\hat{z}_0 + z^{(1-\alpha/2)}) \quad (2.5)$$

where  $\Phi_{\text{NOR}}$  is the cumulative standard normal distribution.

*Algorithm 3: bias-corrected percentile bootstrap confidence interval:*

*Steps 1, 2, 3, and 4.* Same as Algorithm 1.

*Step 5.* Calculate the bias-correction constant,  $\hat{z}_0$ , as given in (2.2).

*Step 6.* Determine the new cutoff percentages,  $\alpha_1$  and  $\alpha_2$ , as given in (2.4) and (2.5).

*Step 7.* Order the bootstrap estimates,  $\hat{\theta}^{*b}$ .

*Step 8.* Determine the  $\alpha_1$ th and  $\alpha_2$ th quantiles of the distribution of  $\hat{\theta}^*$  denoted by  $\hat{\theta}^{*(\alpha_1)}$  and  $\hat{\theta}^{*(\alpha_2)}$  respectively.

*Step 9.* Form the  $1 - \alpha$  confidence interval as given in (2.3).

## Results and discussion

For each method of sampling, the standard normal, percentile, and bias-corrected percentile bootstrap intervals were constructed and compared for the 1st, 10th, 25th, and 50th (median) percentiles for MDF product Types 1 and 5. These two types were chosen to aid in the illustration of the benefits and limitations of the bootstrap. Recall their respective sample sizes given above. For each method of sampling,  $B = 2,000$  bootstrap samples of the same size as the original sample were created.

In many cases, but not always, this should be a sufficient number of bootstrap samples to create the confidence intervals. The asymptotic normal confidence intervals will also be provided in order to compare with the bootstrap results.

Table 1 provides the 95% asymptotic normal confidence intervals for Type 1 MDF, while Table 2 shows the fully nonparametric 95% bootstrap confidence intervals. In the tables that follow, LCL stands for lower confidence limit and UCL stands for upper confidence limit. Figure 1 displays the nonparametric empirical bootstrap sampling distribution for each of the four quantiles. An initial look at the bootstrap sampling distributions shown in Fig. 1 indicates that the bootstrap distribution becomes narrower and more peaked as the percentiles increase from 1 to 50, reflecting smaller variability in the sampling distribution.

In Table 2, the intervals for the 1st percentile of Type 1 MDF are rather wide. They are, in fact, wider than the asymptotic normal intervals. This, again, is to be expected given the limited amount of data in the extreme lower tail of the IB data. These wide bootstrap intervals may provide early warnings on uncertainty to the MDF engineer or technical manager regarding the variability present in the destructive sampling process. These bootstrap intervals also provide MDF manufacturers with a defensible metric of quality near the manufacturer's lower

**Table 1** 95% Asymptotic normal confidence intervals for IB strength of Type 1 MDF

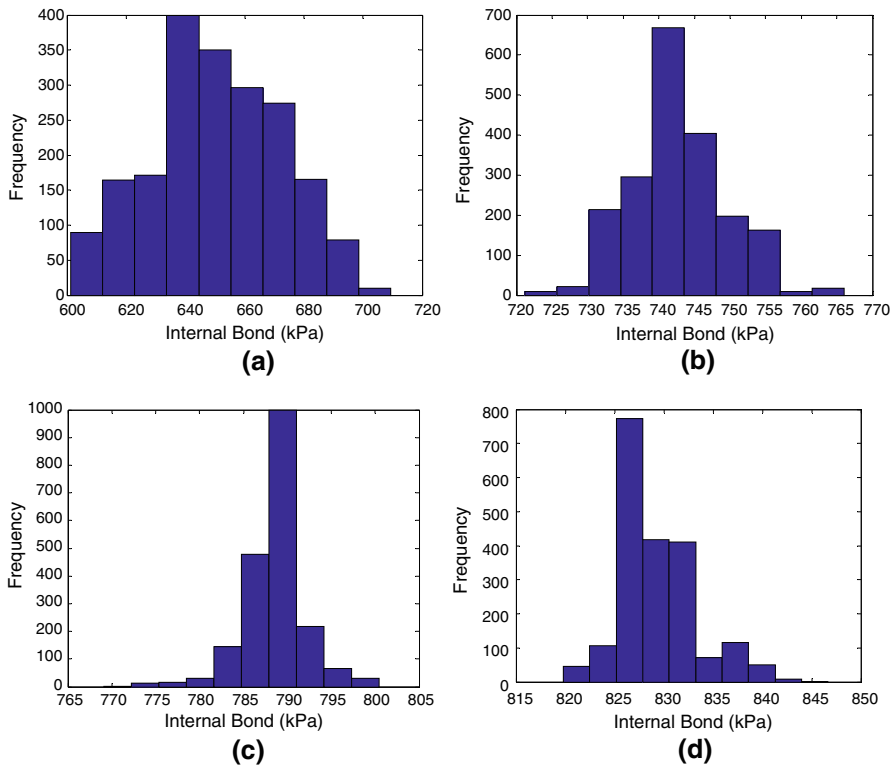
$p$	$\hat{t}_p = \text{quantile (kPa)}$	LCL	UCL
.01	670.7	657.8	683.6
.10	741.8	732.8	750.9
.25	783.1	775.7	790.6
.50	829.1	822.4	835.8

LCL lower confidence limit, UCL upper confidence limit

**Table 2** Fully nonparametric 95% bootstrap confidence intervals for IB strength of Type 1 MDF

$p$	$\hat{t}_p = \text{quantile (kPa)}$	Interval type	LCL	UCL
.01	651.1	Standard	601.7	761.3
		Percentile	601.2	693.8
		Bias-corrected	601.2	684.5
.10	742.5	Standard	728.6	755.1
		Percentile	730.4	756.8
		Bias-corrected	730.2	754.5
.25	788.4	Standard	782.0	795.5
		Percentile	781.9	795.7
		Bias-corrected	777.7	793.9
.50	829.4	Standard	821.5	836.0
		Percentile	823.2	838.7
		Bias-corrected	822.5	838.4

LCL lower confidence limit, UCL upper confidence limit



**Fig. 1** Sampling distribution of percentiles for Type 1 MDF under the fully nonparametric bootstrap sampling method. **a** 1st, **b** 10th, **c** 25th, and **d** 50th

specification limit. As the percentiles increase and the “relative” IB data become more plentiful, the bootstrap confidence intervals (Tables 1 and 2) are more closely matching the asymptotic intervals. Also, it is useful to acknowledge that the three different methods for constructing the bootstrap confidence intervals when the data become plentiful yielded very similar results. Figure 1 yields plots reasonably close enough to normality for all of these three intervals to be in agreement.

In order to construct intervals based on the NBSP method, it was previously determined that the underlying parametric distribution for Type 1 MDF is better modeled by the normal than Weibull or lognormal. Table 3 and Fig. 2 show the confidence intervals and sampling distribution, respectively, for Type 1 MDF based on the NBSP sampling method.

The sampling distributions shown in Fig. 2 appear approximately normal for each of the percentiles. It is also observed that the intervals are similar to the asymptotic intervals as well as similar among themselves. Certainly, this shows the benefit of parametric assumptions.

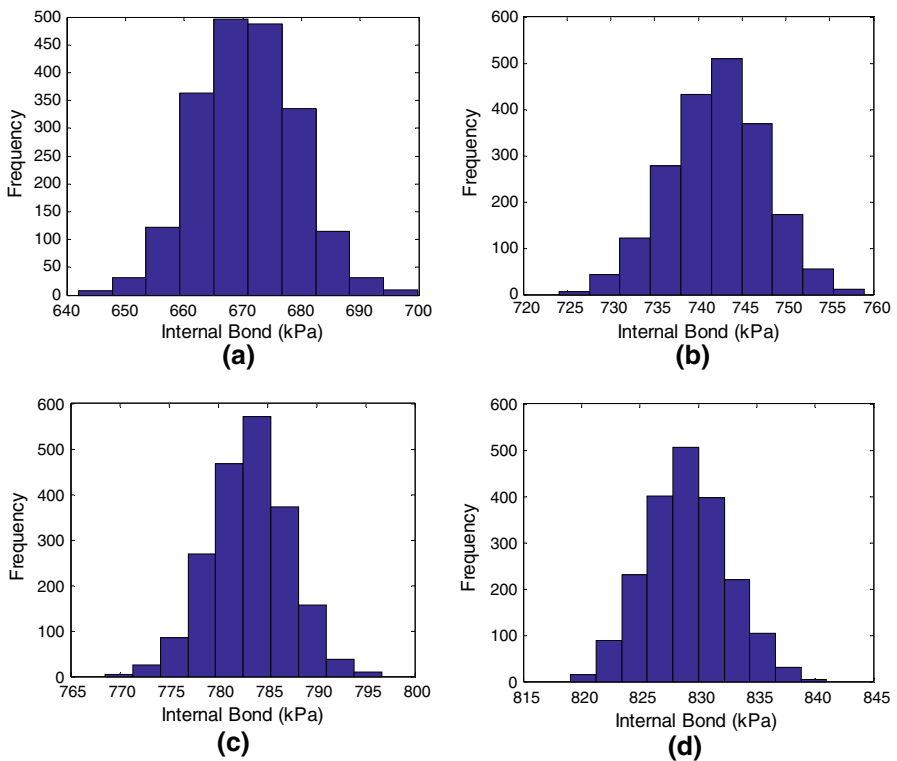
Table 4 provides the 95% asymptotic normal intervals for Type 5 MDF. Table 5 and Fig. 3 display the fully nonparametric intervals and sampling distributions, respectively, for Type 5 MDF percentiles. Notice the discrete nature and skewness



**Table 3** NBSP 95% bootstrap confidence intervals for IB strength of Type 1 MDF

$p$	$\hat{t}_p = \text{quantile (kPa)}$	Interval type	LCL	UCL
.01	670.9	Standard	654.1	686.9
		Percentile	654.3	687.0
		Bias-corrected	652.5	685.5
.10	742.0	Standard	731.4	752.1
		Percentile	731.2	752.5
		Bias-corrected	730.8	751.8
.25	783.3	Standard	775.3	790.9
		Percentile	775.4	790.8
		Bias-corrected	774.9	790.5
.50	829.0	Standard	822.2	835.9
		Percentile	822.4	835.9
		Bias-corrected	822.6	836.3

LCL lower confidence limit, UCL upper confidence limit



**Fig. 2** Sampling distribution of percentiles for Type 1 MDF under the NBSP method. **a** 1st, **b** 10th, **c** 25th, and **d** 50th

**Table 4** 95% Asymptotic normal confidence intervals for IB strength of Type 5 MDF

$p$	$\hat{t}_p = \text{quantile (kPa)}$	LCL	UCL
.01	1,037.4	994.4	1,080.5
.10	1,140.0	1,109.8	1,170.2
.25	1,199.6	1,174.8	1,224.3
.50	1,265.7	1,243.4	1,288.1

LCL lower confidence limit, UCL upper confidence limit

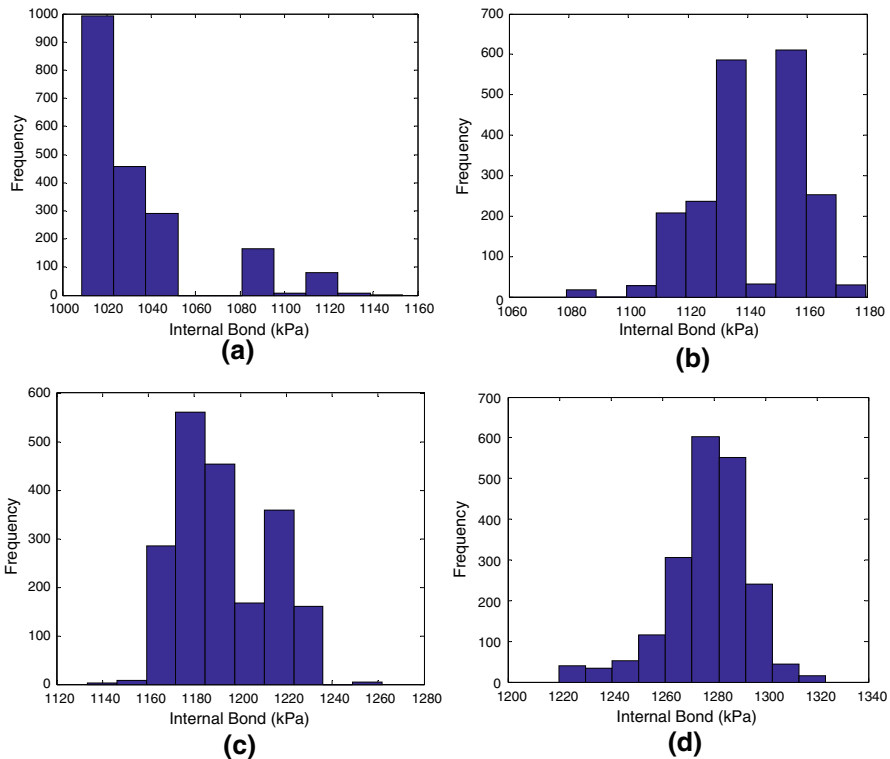
**Table 5** Fully nonparametric 95% bootstrap confidence intervals for IB strength of Type 5 MDF

$p$	$\hat{t}_p = \text{quantile (kPa)}$	Interval type	LCL	UCL
.01	1,030.0	Standard	956.0	1,069.0
		Percentile	1,008.7	1,110.9
		Bias-corrected	1,008.7	1,039.1
.10	1,140.0	Standard	1,098.8	1,168.2
		Percentile	1,110.1	1,164.5
		Bias-corrected	1,085.2	1,160.3
.25	1,191.9	Standard	1,148.9	1,227.0
		Percentile	1,160.4	1,225.9
		Bias-corrected	1,160.4	1,224.5
.50	1,277.4	Standard	1,248.9	1,309.7
		Percentile	1,232.8	1,303.1
		Bias-corrected	1,231.0	1,302.1

LCL lower confidence limit, UCL upper confidence limit

of the histogram for the 1st percentile. As can be seen in Table 5, the bias-corrected interval takes this into account whereas the normal and percentile intervals do not.

The sampling distributions shown provide an example of a limitation of the fully nonparametric bootstrap. When the sample size is relatively small, as is the case of Type 5 MDF, the sampling distributions appear more discrete. Practitioners are advised that when these histograms are discrete or appear “snaggle-toothed”, as in Fig. 3a and b, to increase the resampling size to, say,  $B = 5,000$ . If the histogram no longer has a “snaggle-toothed” appearance, then the larger resampling size has helped. However, if the sampling distribution still maintains a “snaggle-toothed” appearance, then practitioners are advised not to use the fully nonparametric approach for constructing bootstrap confidence intervals. The NBSP method intervals for Type 5 MDF are shown in Table 6. The sampling distributions appear very similarly to those in Fig. 3 and will not be shown in order to conserve space. The intervals are in greater agreement with the asymptotic intervals and with each other than with the fully nonparametric case, and the sampling distributions appear normally distributed for each percentile.



**Fig. 3** Sampling distribution of percentiles for Type 5 MDF under the fully nonparametric bootstrap sampling method. **a** 1st, **b** 10th, **c** 25th, and **d** 50th

## Summary and conclusions

This paper has given the reader an opportunity to briefly explore the basic ideas surrounding bootstrap methods, the construction of bootstrap confidence intervals, and how it can be applied to the estimation of percentiles (especially lower) from real manufacturing data using costly (due to primarily human labor along with material lost) destructive testing on IB. The approach is broader than just improving manufacturing assessment of reliability (or quality or safety specification) percentiles; plus it allows for less restrictive assumptions.

For a sufficiently large sample size, as is the case for Type 1 MDF, the fully nonparametric bootstrap sampling distributions appear continuous and are roughly normally distributed. It is relatively a matter of preference as to which of the bootstrap interval types are used. Indeed, they provide very similar results. However, it is clear that some care should be taken when examining the 1st percentile. When the sample size is large, nonparametric sampling is an appropriate choice and can be used more confidently.

Conversely, when the sample size is much smaller, as is the case for Type 5 MDF, and when sampling is done using the fully nonparametric method, the

**Table 6** NBS 95% bootstrap confidence intervals for IB strength of Type 5 MDF

$p$	$\hat{t}_p = \text{quantile (kPa)}$	Interval type	LCL	UCL
.01	1,038.1	Standard	991.7	1,080.0
		Percentile	994.7	1,083.0
		Bias-corrected	989.9	1,079.2
.10	1,140.0	Standard	1,106.1	1,172.1
		Percentile	1,108.2	1,172.7
		Bias-corrected	1,107.7	1,171.9
.25	1,199.4	Standard	1,172.3	1,225.9
		Percentile	1,172.5	1,225.5
		Bias-corrected	1,171.8	1,225.3
.50	1,265.6	Standard	1,243.7	1,287.8
		Percentile	1,242.6	1,287.0
		Bias-corrected	1,242.3	1,286.2

*LCL* lower confidence limit, *UCL* upper confidence limit

bootstrap sampling distributions can be irregular and often do not resemble a normal distribution. Furthermore, the three methods discussed for constructing bootstrap confidence intervals do not yield similar results using the nonparametric bootstrap. This may complicate the interpretation of such intervals and requires considerations other than those recommended for the large sample case.

If no distributional assumptions can be made, it is recommended that the practitioner use the bias-corrected percentile intervals as a first choice. Doing so can still produce accurate results for the median or lower quartile using the nonparametric method with a small sample size. However, the authors would recommend not using bootstrap confidence intervals for the lower percentiles and instead resort to another approach. As an alternative, one can use kernel smoothing to better estimate lower percentile in smaller samples (see Polansky 2000). Some others might even consider doing a Bayesian approach, if expert assessments warrant. Ideally, the best answer to estimating lower percentiles realistically is to have a larger sample. Next, three alternatives are suggested to get around this difficulty of needing a larger sample size when cost is prohibitive.

First, study the outliers and classify as due to measurement error or statistical variation. One might do bootstrapping in a way that takes into account the outliers in a data set or determine whether they are truly not representative, thus can be eliminated. A second approach is to estimate the lower percentiles for IB using the multiple regression equation in Young and Guess (2002) for estimating IB or to use a quantile regression approach as in Young et al. (2008), taking advantage of co-variables, when they are available (see also Parajo et al. 1994; André et al. 2008). These modeling approaches may yield more helpful estimates on the lower percentiles. Although these approaches may save sample size, cost, and time from destructive tests, they would require continuous validation of the models. Alternatively, engineering judgment and experiences could be incorporated into a helpful Bayesian approach to get more realistic estimates on the lower percentiles

when the data are small. A third approach would be to sample using the NBSP method which may be a more defensible choice when the sample size is small, provided there is some confidence in the underlying parametric model. Although requiring parametric assumptions, this method is useful in constructing intervals for the extreme lower percentiles.

In summary, the sample size is key for bootstrapping methods. Chernick (1999) tells us that “the main concern in small samples is that with only a few values to select from, the bootstrap sample will under represent the true variability as observations are frequently repeated and the bootstrap samples themselves repeat.” This does not mean that the bootstrap should not be used with small sample sizes. Rather, much greater care should be taken when analyzing the accuracy of results, using the helpful checks in the histogram plots to see whether “snaggle-toothed” histograms appear or not. It is recommended that in the case of constructing confidence intervals, more than 1,000 bootstrap samples should be generated. This number should be increased even more when the sample size is small. Bootstrapping can be used in many other manufacturing settings and on numerous other reliability parameters besides the lower percentiles targeted for improvements.

**Acknowledgments** The authors would like to thank Dr. Halima Bensmail for proof reading over a close to final draft of this paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- André N, Cho HW, Baek SH, Jeong MK, Young TM (2008) Enhanced prediction of internal bond strength in a medium density fiberboard process using multivariate methods and variable selection. *Wood Sci Technol* 42:521–534
- Barlow RE, Proschan F (1975) *Statistical theory of reliability and life testing: probability models*. Holt Rinehart and Winston, New York, NY
- Barlow RE, Proschan F (1981) *Statistical theory of reliability and life testing: probability models*. To Begin With, Silver Spring, MD
- Chernick MR (1999) *Bootstrap methods: a practitioner’s guide*. Wiley, New York, NY
- Davison AC, Hinkley DV (1997) *Bootstrap methods and their application*. Cambridge University Press, New York, NY
- DiCiccio TJ, Efron B (1996) Bootstrap confidence intervals. *Stat Sci* 11(3):189–212
- Edwards DJ (2004) An applied statistical reliability analysis of the internal bond of medium density fiberboard. Masters Thesis. Department of Statistics. University of Tennessee at Knoxville
- Efron B (1981) Nonparametric standard errors and confidence intervals. *Can J Stat* 9:139–172
- Efron B (1987) Better bootstrap confidence intervals. *J Am Stat Assoc* 82(397):171–185
- Efron B, Tibshirani R (1993) *An introduction to the bootstrap*. Chapman & Hall, New York, NY
- Hoffmeyer P, Sorensen JD (2007) Duration of load revisited. *Wood Sci Technol* 41(8):687–711
- Kim KO, Kuo W (2003) Percentile life and reliability as performance measures in optimal system Design. *IIE Trans* 35:1133–1142
- Kuo W, Chien WTK, Kim T (1998) *Reliability, yield, and stress burn-in*. Kluwer Academic Publishers, Norwell, MA
- Kuo W, Prasad VR, Tillman FA, Hwang CL (2000) *Optimal reliability design: fundamentals and applications*. Cambridge University Press, Cambridge, UK

- Martinez WL, Martinez AR (2002) Computational statistics handbook with Matlab. Chapman and Hall/CRC, Boca Raton, LA
- Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. Wiley, New York, NY
- Meeker WQ, Escobar LA (2004) Reliability: the other dimension of quality. *Qual Technol Quant Manag* 1(1):1–25
- Moses DM, Prion HGL, Li H, Boehner W (2003) Composite behavior of laminated strand lumber. *Wood Sci Technol* 37(1):59–77
- Parajo JC, Alonso JL, Vazquez D (1994) Effect of selected operational variables on the susceptibility of NaOH-pretreated pine wood to enzymatic-hydrolysis—a mathematical approach. *Wood Sci Technol* 28(4):297–307
- Polansky AM (2000) Stabilizing bootstrap-T confidence intervals for small samples. *Can J Stat* 28(3):501–526
- Prasad VR, Kuo W, Kim KO (2001) Maximization of a percentile life of a series system through component redundancy allocation. *IIE Trans* 33:1071–1079
- Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York, NY
- Steiger R, Arnold M (2009) Strength grading of Norway spruce structural timber: revisiting property relationships used in EN 338 classification system. *Wood Sci Technol* 43(3–4):259–278
- Young TM, Guess FM (2002) Mining information in automated relational databases for improving reliability in forest products manufacturing. *Int J Reliab Appl* 3(4):155–164
- Young TM, Perhac DG, Guess FM, León RV (2008) Bootstrap confidence intervals for percentiles of reliability data for wood plastic composites. *For Prod J* 58(11):106–114