International Journal of Reliability and Applications
Vol. 5, No. 4, pp.129-145, 2004

Forcing a Closer Fit in the Lower Tails of a Distribution for Better Estimating Extremely Small Percentiles of Strengths

Frank M. Guess, Ramón V. León and Weiwei Chen
Department of Statistics, Stokely Management Center,
University of Tennessee, Knoxville, TN 37996-0532 USA

Timothy M. Young*
Tennessee Forest Products Center, 2506 Jacob Drive,
University of Tennessee, Knoxville, TN 37996-4570 USA

Abstract. We use a novel, forced censoring technique that closer fits the lower tails of strength distributions to better estimate extremely smaller percentiles for measuring progress in continuous improvement initiatives. These percentiles are of greater interest for companies, government oversight organizations, and consumers concerned with safety and preventing accidents for many products in general, but specifically for medium density fiberboard (MDF). The international industrial standard for MDF for measuring highest quality is internal bond (IB, also called tensile strength) and its smaller percentiles are crucial, especially the first percentile and lower ones. We induce censoring at a value just above the median to weight lower observations more. Using this approach, we have better fits in the lower tails of the distribution, where these smaller percentiles are impacted most. Finally, bootstrap estimates of the small percentiles are used to demonstrate improved intervals by our forced censoring approach and the fitted model. There was evidence from the study to suggest that MDF has potentially different failure modes for early failures. Overall, our approach is parsimonious and is suitable for real time manufacturing settings. The approach works for either strengths distributions or lifetime distributions.

Key Words : *first percentile, lower percentiles, forced censoring for fitting better, strengths of materials, internal bond, tensile strength, probability plots.*

1. INTRODUCTION

Medium Density Fiberboard (MDF) is a superior engineered wood product with high reliability and grooving ability for unique designs. MDF provides greater qualities on

* Corresponding Author.
E-mail address: tmyoung1@utk.edu

consistency of finish and density, plus freedom from knots and natural irregularities compared to regular wood products. There are many examples of MDF being used in furniture, cabinets, shelving, flooring, molding, etc. Reliability of such products is important to all concerned.

Product "life" for MDF can be measured in terms of the strength to failure, as opposed to the time to failure. The strength or pounds per square inches to failure is a crucial reliability parameter of the product. It naturally allows the producer to make assurances to customers about the safe, useful "life" range of the product. One key measure of the quality or reliability is MDF's internal bond (IB), also called tensile strength, which is measured in pounds per square inch (or equivalent metric units) in destructive tests on sampled MDF until breakdown. Compare Young and Guess (2002) for how such data is stored and used in a real time data base with regression modeling to predict strength. See Guess, Walker, and Gallant (1992), Guess and Proschan (1988), and Guess, Hollander and Proschan (1986) for other measures of reliability than percentiles.

The lower percentiles are particularly of great interest and importance for companies, government oversight organizations, and consumers in specifying the product reliability of MDF. Some legal issues require careful monitoring and estimation of these lower percentiles in strengths of various products, e.g., MDF. Compare Kim and Kuo (2003), Kuo, Chien, Kim (1998), and Kuo, Prasad, Tillman, and Hwang (2000) for more on percentiles. Also, see Walker and Guess (2003) for strengths of container bottles using Kaplan and Meier graphs and nonparametric approaches. Guess, Edwards, Pickrell and Young (2003) view graphically and exploratorially this type of data, but did not provide confidence intervals for percentiles there. We compute in this paper such confidence intervals for lower percentiles, via parametric modeling and bootstrapping through a novel approach to fit the lower distributions better. New results on different IB data are presented, regarding these statistical distributions.

There is evidence from the study to suggest that MDF has potentially different failure modes for early failures. Probability plots show that naturally expected failure distributions, like the Weibull, do not fit the raw data satisfactorily overall. Even the distribution of overall best fit, such as the normal, provides poor estimates of the smaller percentiles. We induce censoring to weight lower observations more. All the observations no larger than the median are retained intact as exact failures, while observations beyond the median are censored at a forced value slightly larger than the median, but less than the next true observed failure above the median.

After applying the censoring technique, we have better goodness of fit in the lower tails, where the smaller percentiles are impacted the most. It shows that the Weibull distribution fits the lower, shorter-lasting MDF's better, while the overall strength is fitted by the normal distribution better. Recall Weibull's theory as a weakest link model for early failures; while, for overall failures, normality via the Central Limit Theorem (CLT) is more appropriate. The CLT normality comes by the total overall strength being typically the sum of many individual fiber strengths. Finally, bootstrap estimates of the small percentiles are used to support using the forced censoring technique by how it improves the fitted model's percentile confidence intervals. We use both percentile bootstraps and t bootstrap intervals.

Our approach can be used for many other applications beside strengths of materials and their lower percentiles. The forced censoring technique can be employed successfully for warranty or lifetime data analysis when estimates of new warranties are based on smaller percentiles. It could also be used in the study of time to submission for rebates or of times to return a product in marketing analysis.

Overall, our approach is parsimonious and is suitable for real time manufacturing settings. Also, it does not depend on the underlying distribution being Weibull, lognormal, or otherwise. The forced median censoring technique improves the data quality for lower percentile estimates. Probability plots help assess which underlying parametric model fits the best. The bootstrapping method adds further evidence to support this approach, as well as validating the confidence interval estimates.

We investigate the most important MDF product type, defined as "Type 1". The specifications of Type 1 are a density of 46 pounds per cubic foot (lbs/ft^3), thickness of 0.625 inches, and width of 61 inches. The dataset had 396 destructive tests till failure. Lower percentile confidence intervals are computed using normal-approximation maximum likelihood (ML) and bootstrap methods. More detailed references on these types of confidence intervals can be found in Meeker and Escobar (1998) with bootstrap intervals discussed by Davison and Hinkley (1997), Chernick (1999), Efron and Tibshirani (1993), and Efron (2003).

Section 2 analyzes the primary Type 1 product for the complete data and forced censoring at the median. Section 3 presents bootstrap results including the confidence intervals for various parametric models for both the complete and the forced censoring cases. The statistical software S+ (http://www.insightful.com/products/default.asp) and a free add-on called Splida (http://www.public.iastate.edu/~splida/) are used throughout our paper, with some Matlab (http://www.mathworks.com/). Concluding remarks are in Section 4.

2. ANALYZING FORCED CENSORING DATA OF TYPE 1 PRODUCT TO FIT THE LOWER TAILS BETTER

The complete data set of 396 failures for the Type 1 product is initially fitted to several popular distributions of lifetime data. The qualities of the model fits are examined graphically on the respective probability plot in Figure 1. It is highly recommended by the authors to implement this exploratory step before making any further statistical inferences. By plotting the data, we can quickly identify underlying issues and proceed with the most appropriate strategies including median censoring. Recall that IB is measured here in pounds per square inches (psi).

Figure 1 displays that observed early failures deviated from the straight lines of parametric ML estimates. There are a few data points on the lower tail and mostly the upper tail that were not well captured by any of the distribution models, evidenced by both tails stretching outside the coverage of pointwise 95% confidence interval of ML estimated models. Later, it is important for quality goals that need both a specification number and pointwise confidence interval on the reliability. We notice that the amount of sampling variability at the extreme observations can be rather large, as suggested by the simultaneous confidence bands in Figure 2. See, for example, Section 3.8 of Meeker and Escobar (1998) for more details.

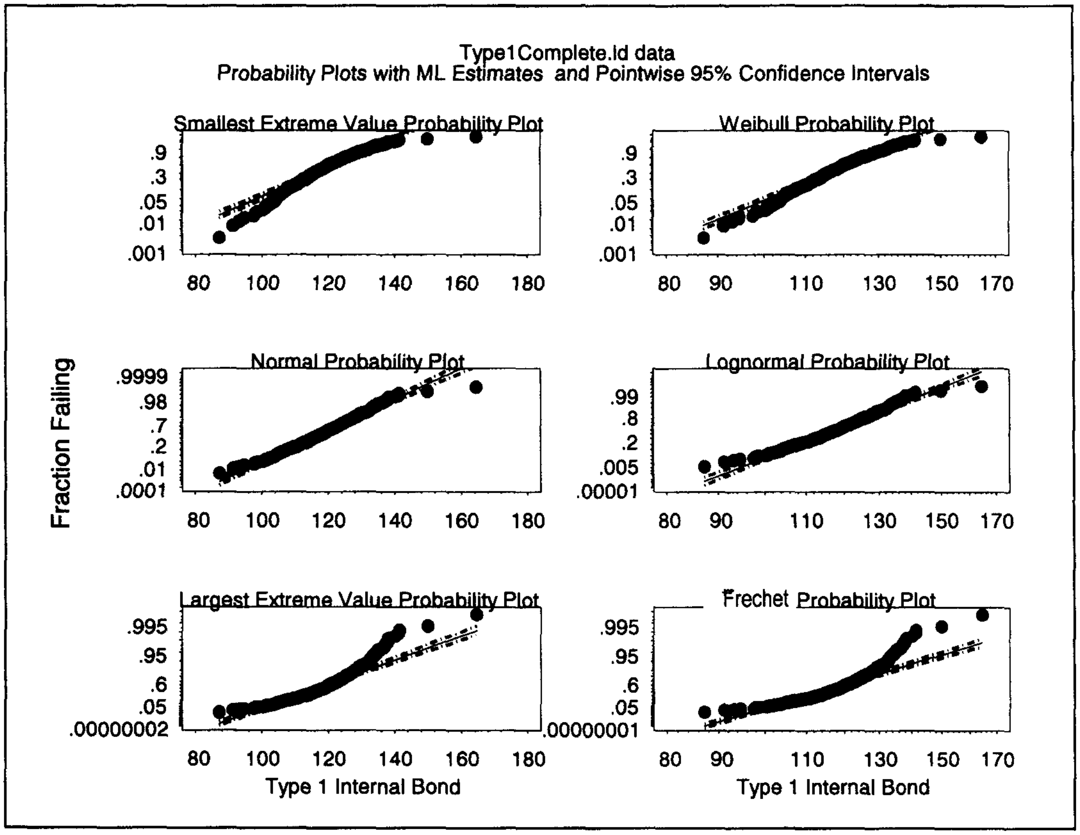

Figure 1. Complete Data Probability Plots with ML Estimates

It is indicated in Figures 1 and 2 that the ML estimated normal distribution model seems to be the best fit to the whole bulk of data, and that some curvature change exists no matter which model is fitted. The existence of such behavior in the data might be signs of potential different failure modes or mixture of subpopulations at the extremes or of outliers during the breakdown or measurement process (section 6.6, Meeker and Escobar, 1998). In these cases, a certain model, for example the normal distribution, may fit the majority of the data better than the other, but this is merely achieved by compromising the local approximation of failure modes toward extreme values, lower or upper. Or, the shape of an empirical failure model, such as Weibull, happens to be largely determined by the upper part of data, while the desired lower percentiles deviate from the observed data which are less influential.

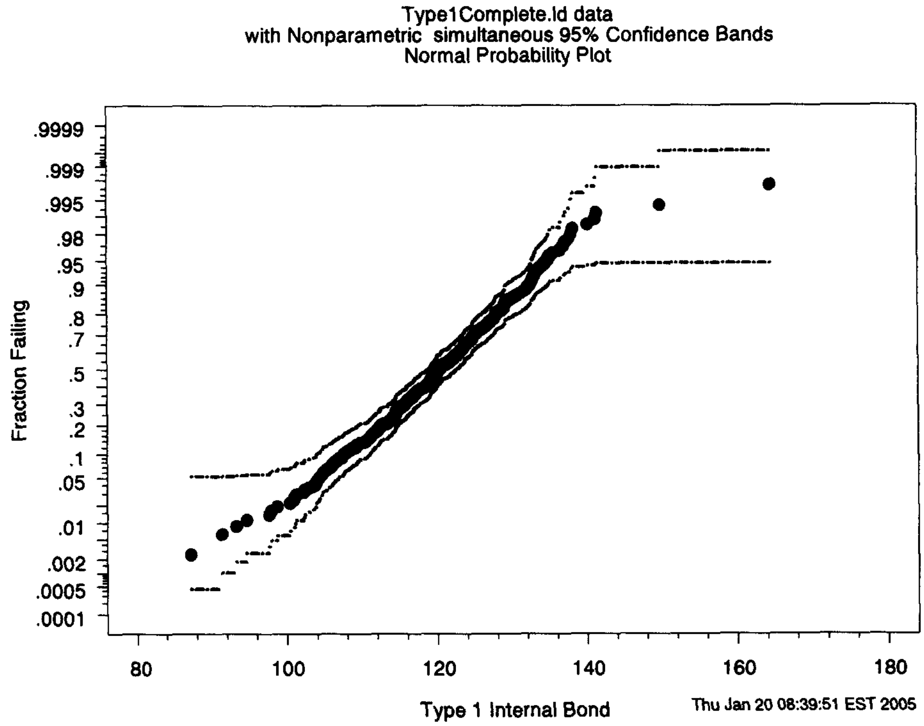

Figure 2. Normal Probability Plot for the Complete Data and Simultaneous Approximate 95% Confidence Bands

We will further present quantitative results in Section 3 that the first percentile (and lower percentiles) estimates using the complete data naively were generally unreliable, either too optimistic or overly conservative. This may lead to higher costs of manufacturing when product reliability is misjudged. With the existing data set that has included sufficient information for the small percentile estimates, it is a cost-efficient and statistically sound solution to reengineer and cleanse the data of potential outliers and reassess the pragmatic information quality for the lower percentiles. See English (1999), Huang, Lee, and Wang (1999), or Redman (2001).

Because the goodness of global model fit sacrifices the more important lower percentile estimates, we use a forced median-censoring technique to increase the model dependence on the lower tail information. Essentially, such a technique reengineers the data set so that the upper half of the complete data set are regarded as being censored at a forced value slightly larger than the median but less than the next true observed failure time. Hereby, these large observations are not as informative as the smaller observations in that their breakdown strengths are only known to be larger than the median. In other words, more weights are put on the observations of smaller values in fitting a model.

For the failure data of Type 1 product, we retain 198 observations on the lower tail while censoring the upper half of data (198 observations). This weighted data is fitted by selected models in Figure 3.

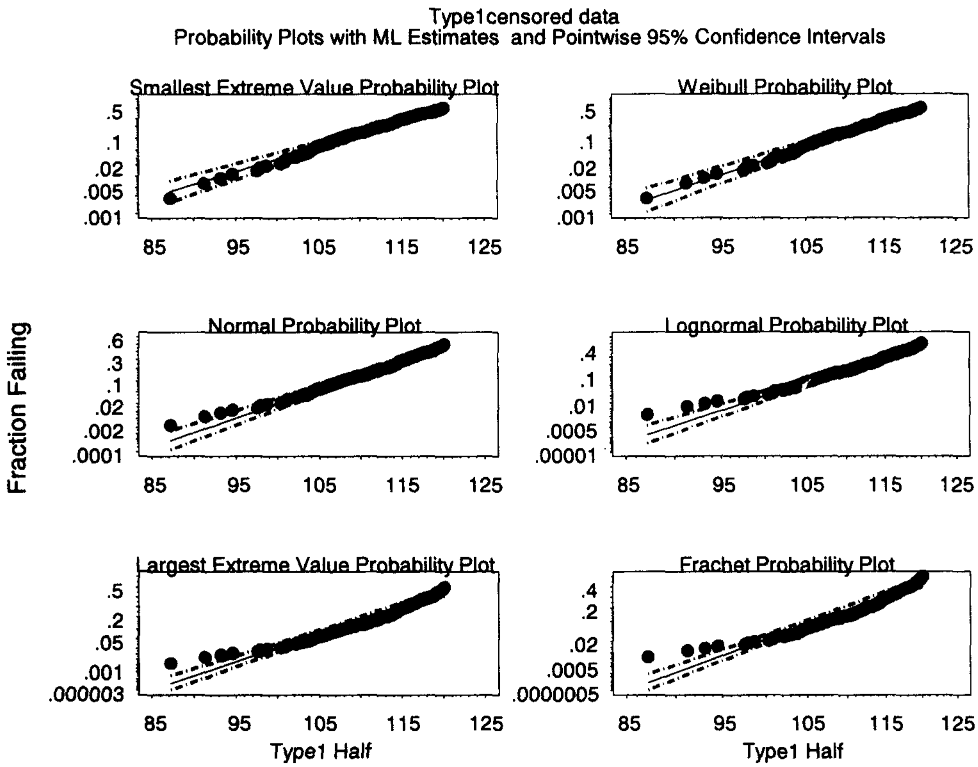

Figure 3. Probability Plots of Median Censored Data with ML Estimates

Upon censoring the upper half of the data, the fitted ML estimated lines of the Weibull and Smallest Extreme Value distributions (Figure 3) are able to capture the pattern of small extreme values more "closely" and more importantly, the data on the lower tail, than other models. The lowest data, which would be considered incorrectly as outliers if it were without median censoring (compare both Figures 1 and 2), now falls completely within the 95% confidence interval of a Weibull or S.E.V. model. For additional specific numbers, say 90 psi or a previous first percentile, for example, with improved, continuous quality goals, we really want and prefer to have pointwise confidence intervals for their new, improved reliability to report to management. When the interval fails to enclose the observed data, however, it is an appropriate conclusion that the data is not as consistent with the model hypothesis (Section 7.3.2, Meeker and Escobar, 1998). Thus, we suspect different underlying failure modes over the whole range of observed failures. The early failures are similar to the "infant mortality" for many manufacturing settings. Recall Weibull's weakest link model for the catastrophic effect of even a very small external force upon a certain portion of inferior products, here mostly the lower percentiles. However, the breakdown of the majority of MDF products is determined by a combined strength of individual fibers and bonding between the fibers, where the Central Limit Theorem is more suitable.

Table 1 presents the loglikelihood and AIC scores of select models as the quantitative evidence for a different early failure mode than the normal model, which fits the complete data the best. The Akaike's Information Criterion (AIC) for model selection (Akaike, 1973) favors the model that minimizes AIC score based on the same information (median censoring or not). Therefore, the Weibull ML fit, also seen in Figure 4, is the best approximating model to the censored data set.

Table 1. Select Model Scores for the Complete and Censored Data

ML fit	With median censoring		W/O median censoring	
	Log likelihood	AIC	Log Likelihood	AIC
Weibull	-868.8	1741.6	-1518	3040
S.E.V.*	-869.4	1742.8	-1527	3058
normal	-871.5	1747	-1469	2942

* We use S.E.V. to denote the Smallest Extreme Value model where applicable in this paper.

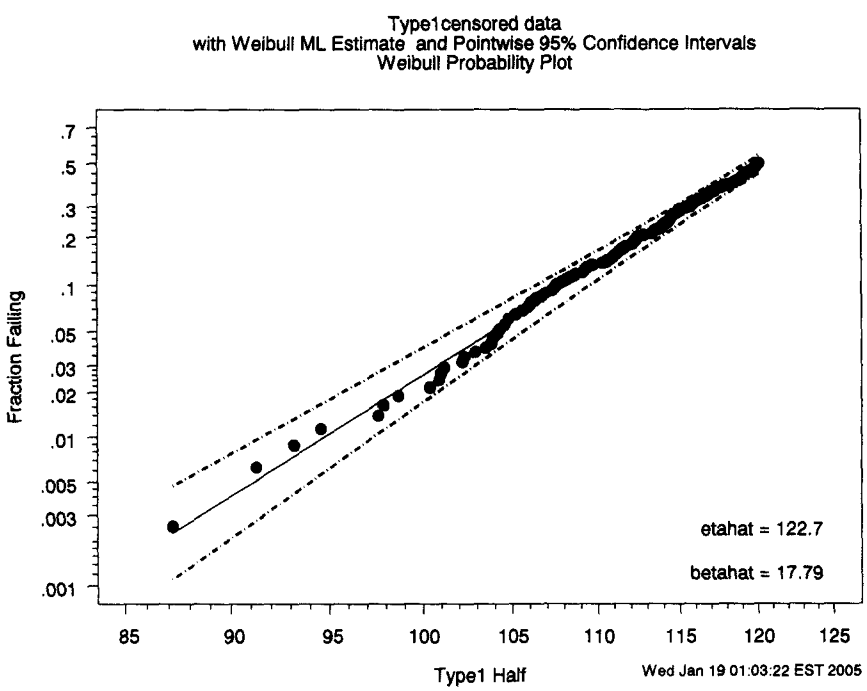

Figure 4. Median Censored Data on the Weibull Probability Plot

Figure 5 shows on the Weibull probability plot how the first percentile estimates are obtained from all three models in Table 1. The solid straight line and the corresponding 95% pointwise confidence bands show the Weibull ML fit, while the curve of normal ML fit deviates the most severely from the lower tail of observed failures. The difference between the Weibull and S.E.V. model on the first percentile is trivial. The S.E.V. model may be of interest if a conservative estimate is preferred in the practical context of reliability evaluation. It is noticeable that the S.E.V tends to produce overly underestimated results as the percentage (quantile) of interest becomes smaller than 1% (0.01).

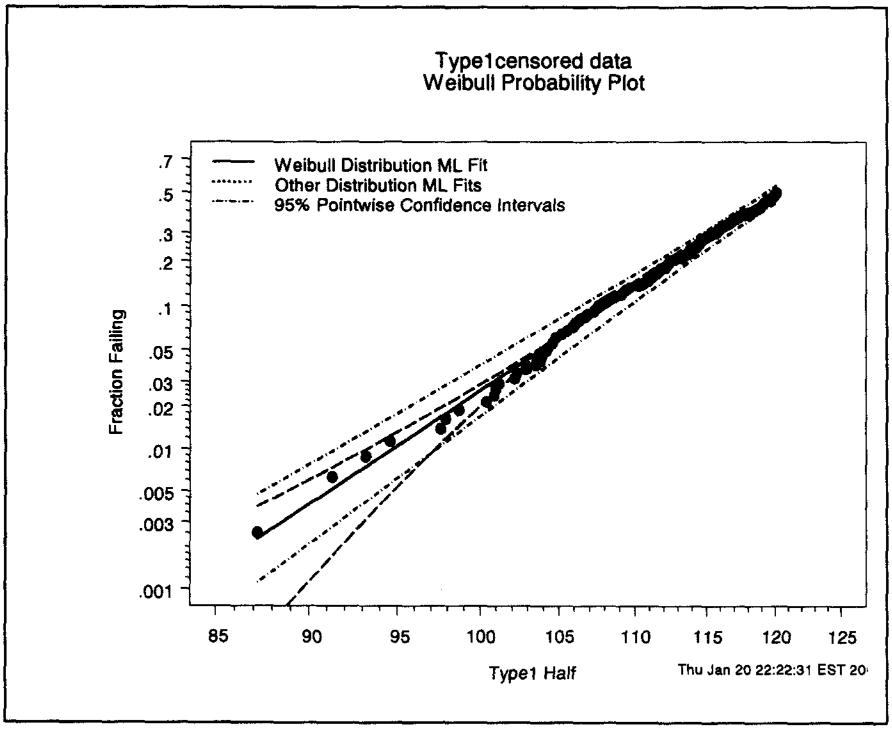

Figure 5. Estimating the First Percentiles from Select Models

Table 2. The First Percentile Normal-approximation Estimates of Select Models for the Censored Data

ML fit	p	Pencentile	Std_Err_	95%_Lower	95%_Upper
Weibull	0.01	94.746	1.47018	91.908	97.672
S.E.V.	0.01	93.255	1.75203	89.821	96.689
Normal	0.01	97.262	1.16150	94.986	99.539

Splida also computes the asymptotic normal-approximation confidence intervals while generating the "probability plot with parametric ML fit", which is a macro in the Splida menu. Table 2 presents the 95% confidence intervals generated based on the Weibull, S.E.V., and normal ML fits. The S.E.V. model gives the most conservative estimate, while the normal model is too optimistic because the data is unduly fitted. See Section 7.3.3 and 8.4 of Meeker and Escobar (1998) for more details on the normal assumption of log-percentile in this estimation method. Meeker and Escobar (1998) comment, "with moderate-to-large samples (the normal approximation) are useful for preliminary confidence intervals" and "quick, useful, and adequate for exploratory work". Other alternatives of estimating confidence intervals, including a simple nonparametric estimation and various bootstrap methods, are discussed in the next section of this paper. We report more and final results of the first percentile confidence intervals in Table 3 of Section 3.

3. USING BOOTSTRAP CONFIDENCE INTERVALS TO DEMONSTRATE IMPROVEMENTS IN INTERVALS BY THE FORCED CENSORING TECHNIQUE

The novel technique of forced median censoring has shown its capability in helping detect possibly different failure modes and improving the model fit, as well as percentile estimates on the lower tail. However, there are some potential weaknesses, both theoretic and practical, in the approach thus far. Figure 2 has suggested that the sampling variability at the extremes can be rather large so that the ML fit plots may give the false impression in model comparisons (Section 6.4.1 Meeker and Escobar, 1998). The entropic information model selection criterion such as AIC affirms our conclusions drawn from probability plotting; yet, the normal-approximation confidence interval still has its theoretic shortcomings, for example, the normal assumption of transformed data may not be the case especially when the sample size is not large. In this section, we rely on the bootstrap method to further demonstrate the information quality improvements from applying the forced median censoring technique, and will provide more accurate confidence intervals, in support of a practitioner's work and presentation to higher management. Table 3 presents the 95% confidence intervals of the first percentile for both complete and median censored data, using the approximate and bootstrap method, both nonparametric and parametric. We will explain more below.

The main idea of the bootstrap method is to simulate the repeated sampling process, reduce the sampling variations in the data, and compute intervals from the simulated distribution of needed statistics without having to making any assumptions on the sampling distribution. The following are three standard steps: 1.) generate a resampled data set, called bootstrap sample, repeatedly for a large number of times, 2.) compute the desired statistic for each bootstrap sample, and 3.) extract information from the distribution of the statistics obtained in 2.), which is the simulated sampling distribution of the population statistic.

For step 1.), the resampling method can be either parametric or nonparametric. See Section 9.2.2 of Meeker and Escobar (1998). We choose the nonparametric bootstrap sampling scheme for all of our bootstrap samples. There are B = 2000 bootstrap samples, each consisting of 396 failures resampled with replacement from the actual data cases, bound with their respective original censoring information where applied. For step 2.), the statistic (first percentile here) for each bootstrap sample can be computed both parametrically and nonparametrically, as the first column of Table 3 specifies "nonparametric", "Weibull", etc.

To avoid confusion of terminology in step 1.), we stress again that all the resampling schemes in this paper are assumed to be nonparametric. The term "nonparametric" we use, for example in Table 3, specifically refers to the "totally nonparametric bootstrap method" (compare Martinez and Martinez 2002 and their notation which we borrow here): not only is the resampling scheme nonparametric in the "totally nonparametric method," but the population parameter θ (here the first percentile) is also calculated nonparametrically as $\hat{\theta}$; the same nonparametric computation of estimate of θ repeats to each bootstrap sample, producing the empirical bootstrap distribution of $\hat{\theta}^*$, where $\hat{\theta}^{*b}$ is the *b*th bootstrap estimate. All the other confidence intervals in Table 3, which are not labeled under the "nonparametric model assumption", are obtained in the parametric way: a ML estimated model is used to generalize the sample data and further statistical inference is computed based on the model parameters. For other different general details on asymptotic normality of percentiles, see Serfling (1980).

Under each model assumption, there are different confidence interval methods, noted by "interval method" as the last column of Table 3, to construct a confidence interval for the desired statistic, namely the first percentile. "Normal-approximation" refers to the pointwise normal-approximation confidence intervals under the nonparametric model assumption (Section 3.4.2, Meeker and Escobar, 1998), or to a log-percentile normal-approximation confidence interval under respective parametric model assumptions (Section 7.3.3, Meeker and Escobar, 1998). When using bootstrap method, one can select either "bootstrap-t" or "bootstrap-percentile" method to compute the confidence intervals from the simulated sampling distribution of bootstrap step 3.). If appropriately used, the bootstrap-t confidence intervals can be expected to usually be more accurate than the normal-approximation ones. The mathematical descriptions of these confidence intervals can be found, for example, in Section 3.6, 7.3.3, and 9.3, respectively, of Meeker and Escobar (1998), or compare Edwards, Guess, Young (2004). Splida has provided GUI macros to compute all but the nonparametric bootstrap confidence intervals for the first percentile. We wrote the MATLAB code to compute the bootstrap-t and bootstrap-percentile confidence intervals under the nonparametric model assumption. Recall our previous comments and references earlier in this section.

There is no significant difference in the nonparametric confidence intervals of first percentile between the complete and median censored data, or bootstrap and non-bootstrap method, because the nonparametric method only makes use of the data points local to the first percentile. These nonparametric confidence intervals are much wider, however, than the ones obtained under parametric model assumptions. Although these nonparametric intervals can serve as fairly broad, robust comparisons for intervals obtained by other methods, they do not allow for practical precision of more importance in the real world.

Table 3. 95% Confidence Intervals of the First Percentile Computed Under Various Model Assumptions and With/Without Median Censoring Technique

Model assumption	With median censoring 95%_Lower	95%_Upper	W/O median censoring 95%_Lower	95%_Upper	Interval Method
Nonparametric	**87.2**	**98.7**	**87.2**	**98.7**	Normal-Approximation
Nonparametric	**86.647**	**100.035**	**86.151**	**101.242**	Bootstrap-t
Nonparametric	**87.200**	**100.630**	**87.200**	**100.676**	Bootstrap-Percentile
Weibull	**91.908**	**97.672**	**87.969**	**91.601**	Normal-Approximation
Weibull	**91.834**	**97.392**	**88.085**	**97.164**	Bootstrap-t
Weibull	**91.836**	**97.711**	**78.134**	**92.051**	Bootstrap-Percentile
S.E.V.	**89.821**	**96.689**	**81.305**	**86.346**	Normal-Approximation
S.E.V.	**89.878**	**96.358**	**80.456**	**94.572**	Bootstrap-t
S.E.V.	**89.808**	**96.347**	**64.647**	**87.956**	Bootstrap-Percentile
Normal	**94.986**	**99.539**	**95.402**	**99.147**	Normal-Approximation
Normal	**94.363**	**99.672**	**94.552**	**99.607**	Bootstrap-t
Normal	**94.175**	**99.771**	**94.741**	**99.739**	Bootstrap-Percentile

Because the parametric model is built to best generalize a whole bulk of data and extract information in terms of a few parameters, the computation of normal-approximation confidence interval under a parametric model may come quick and conditionally useful only at the cost of a local approximation, especially at the extremes. Such an approach may be correct when the model fit is good globally over the data range; however, when the globally good fit disagrees with the local data, the estimates become very unreliable. In the case of Type 1 product, the complete data set includes outliers and multiple failure modes. The normal-approximation confidence intervals from the Weibull and S.E.V. ML fits tend to severely underestimate the lower tail, compared to the generally more accurate bootstrap estimates (Meeker and Escobar, 1998). The gap between the bootstrap and normal-approximation confidence intervals ranges from a few to more than twenty pounds per square. Not surprisingly, due to the speculations of overall physical breakdowns in section 2, the normal ML fit may seem to produce close confidence intervals between the bootstrap and non-bootstrap results, by consistently ignoring the smallest extreme values and fitting the majority. The consequence, therefore, is that the normal ML fit tends towards overestimation the lowest percentile.

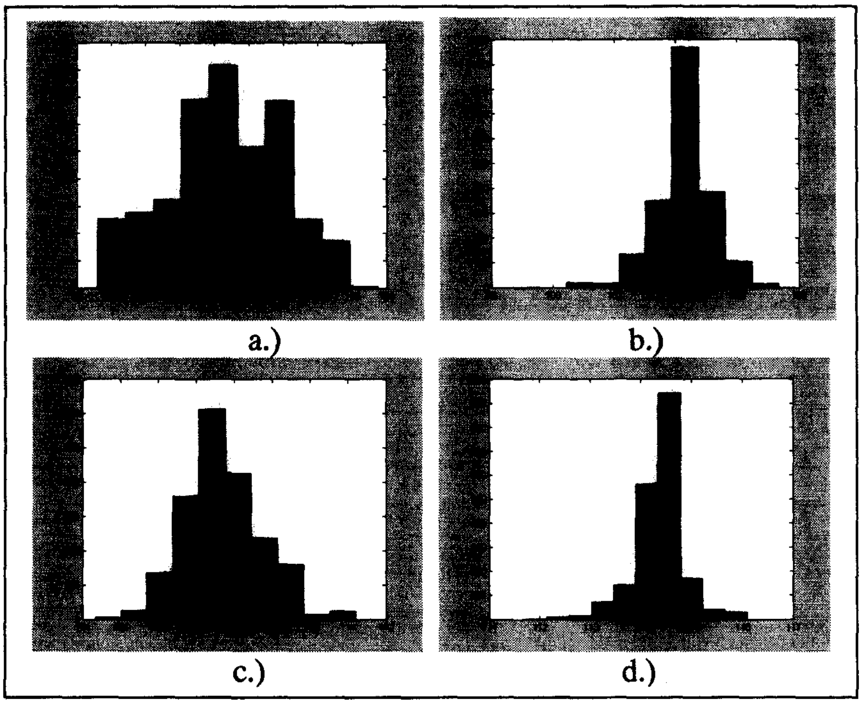

Figure 6. The Histograms of a.) 1st, b.) 5th, c.) 10th, and d.) 25th Percentile Nonparametric Estimates from Bootstrap Samples of the Complete Data for Type 1 Product

The bootstrap estimates are, to a certain extent, resistant to the influence of outliers, but not unconditionally. Even though empirically better than the approximate method, the bootstrap confidence intervals computed from the complete data might be as misleading in the complete data case. During step 3 of the bootstrap procedures, a histogram of the statistics from bootstrap samples can be drawn out as a simulation of the true sample distribution of the statistic. Such bootstrap histograms can warn us of potentially false structure in the complete data or reassure us in the censoring case of their likely usefulness. Figure 6 from the complete data shows much more variations in the first percentile nonparametric estimates of bootstrap samples, compared to the other percentiles, which corresponds to Figure 2 normal plot and causes the estimates of lowest percentiles to be difficult as discussed previously. Figure 7, also generated from the complete data, further shows a strong sign of ambiguity lying in the estimation of first percentile from Weibull ML fit of bootstrap samples. There are apparently two peaks in the histogram-simulated distribution of bootstrap first percentile estimates, caused potentially by different failure modes, or even possibly two different-shaped Weibull's over different failure range, that are previously speculated in this paper. Outliers in the data could be another reason that affected the bootstrapping histograms. The bootstrap estimates reaffirm that a simple complete data ML fit is insufficient to capture the failure mode of Type 1 product and produce reliable estimates of the lowest percentiles.

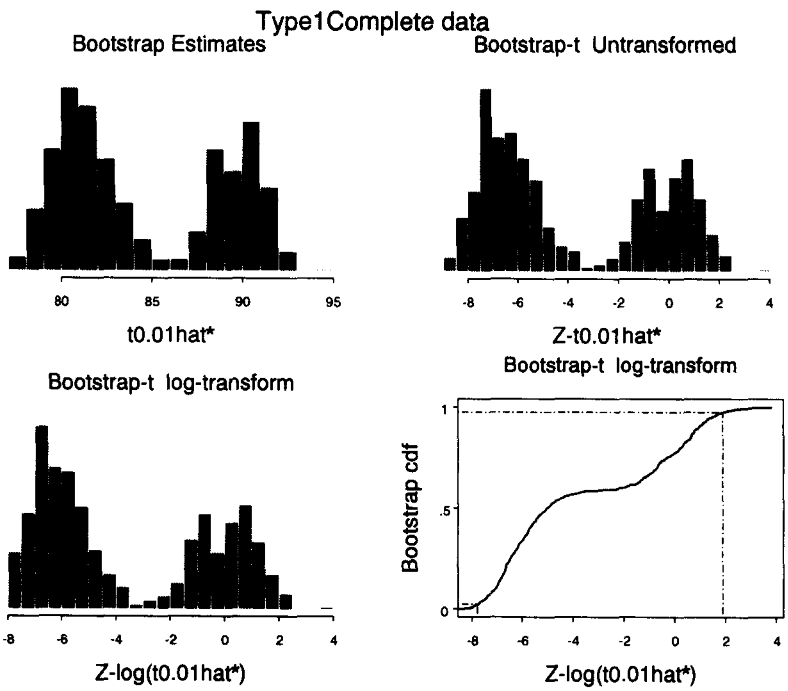

Figure 7. The Histograms of First Percentile Weibull ML Estimates from Bootstrap Samples of the Complete Data for Type 1 Product

As a comparison, the histograms of bootstrap estimate on the median censored data in Figure 8 show no such bimodal patterns. Also, note carefully the scale is different in Figure 8 for the normal to not be as spread out as the other previous Figures. If we look at the computed confidence intervals from the median censored data in Table 3, all three types of estimation methods, normal-approximation, bootstrap-t, and bootstrap-percentile, produce very close results under the a simple model assumption. We slightly favor the Weibull model because the S.E.V. has the tendency of underestimating the data, and because Weibull model is further supported by the information model selection criterion. On different occasions the choice between the Weibull and S.E.V. fit may depend on whether a more accurate or conservative estimate is preferred.

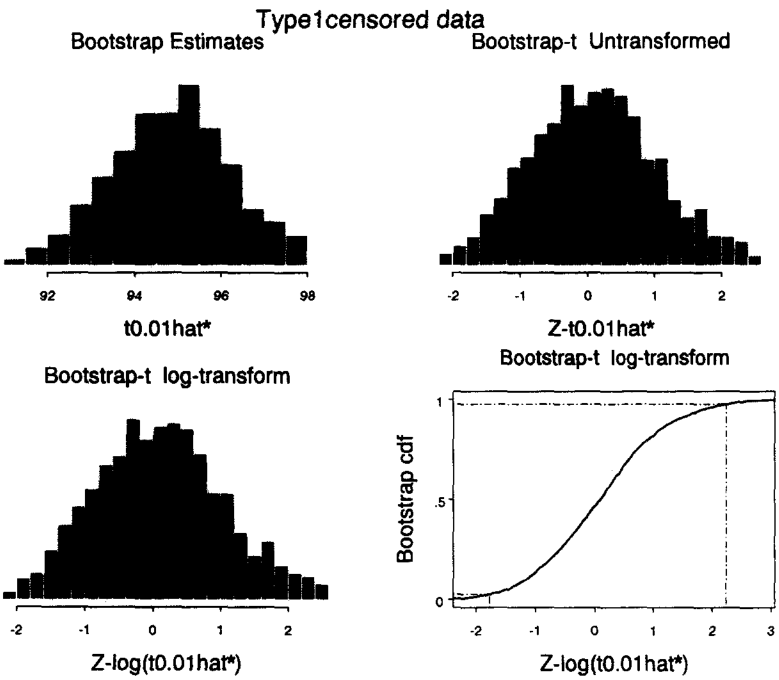

Figure 8. The Histogram of First Percentile Weibull ML Estimates from Bootstraps Samples of the Median Censored Data for Type 1 Product

In presentations to higher management, it is reassuring to have different methods of confidence interval estimation agreeing so closely. Though we generally trust the bootstrap-t estimates more, the ML fit normal estimates are close to them because of the improvement in data quality by the median censoring technique. From a practitioner's point of view, even if a bootstrap-t macro or a computationally-intensive environment is unavailable, the conventional ML fit approach can still be acceptable as long as the median censoring technique has been applied. Such a conclusion also helps the tasks that demand online feedback or timely solutions.

The bootstrap method supports the methodology of the median censoring technique: the data is reengineered by different weights so that a simple model can fit the observed data very well; moreover, higher quality information of the lower percentiles are protected from the influence of overall failure complexity as well as upper outliers in the complete data.

4. SUMMARY AND CLOSING REMARKS

The observed complexity within the complete data set of Type 1 MDF product failures causes great difficulties with estimating the lower percentile. The nonparametric methods poorly manage the benefits of available information. However, simply fitting a parametric model to the primitive, complete data is also problematic for its inadequacy in weighting the most crucial information. The resulting estimates of either approach are not as precise and may lead to higher costs for poor product information or product reliability.

Rather than building a complicated model to match every portion of the observed data, or being misled to unnecessarily collect more data at even higher cost, we introduce a new technique of median censoring which places more weight to only the lower tail data where it is critical for the smallest percentile estimates. It has been shown both graphically and quantitatively that after the data quality is improved, a simple as well as empirical failure model, Weibull, fits the lower tail exceptionally well and produces consistently reliable estimates of the small percentiles.

Probability plots and ML fits are very supportive of the median censoring technique. What is additionally crucial is the confirmation provided by bootstrapping. We have shown that not only is the median censoring technique supported, but empowered by the bootstrap method. The bootstrap simulated sampling distribution reveals different failure modes existing in the complete data set, and that the median censoring technique resolves the bimodality difficulty in the ML fit. The high degree of agreement between the normal-approximation C.I. and the bootstrapped C.I. is strong evidence that the median censoring technique is superior.

Finally, we caution practitioners that as straightforward as the practice seems to be by fitting a commonly known or accepted model to the raw lifetime data, it is dangerous and costly to draw any immediate or convenient inference merely from that type of preliminary analysis. We suggest that the data structure be examined via various probability plots first. If these plots suggest deviations from the ML fit or possible outliers or curvatures, it is advised to apply the forced median censoring technique to put more weights on the part of data of best interest. Then, refit a parametric model for better estimates of small percentiles. It is important that the bootstrap method be used to validate the model and improve the estimates. Under limited situations, the model fitting methods without bootstrapping may perform just fine and render quick and satisfactory results because of the critically improved data quality by the median censoring technique. Overall, our approach to analyzing complex real-world lifetime data is empirically successful in its applications. This approach is also applicable to lifetime data, and for smaller sample sizes that industries sometimes encounter when starting up new mills and developing new products.

ACKNOWLEDGEMENTS

Timothy M. Young, Weiwei Chen, and Frank M. Guess gratefully acknowledge funding in 2004 from the United States Department of Agriculture, Special Wood Utilization Grants Program as administered by The University of Tennessee Institute of Agriculture Experiment Station, contract R11-2219-109.

REFERENCES

Akaike, H. (1973). Information Theory and An Extension of the Maximum Likelihood Principle. In B.N. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory*, Academiai Kiado, Budapest, 267-281.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press, Cambridge, U.K. and New York, NY.

Chernick, M. R. (1999). *Bootstrap Methods: A Practitioner's Guide.* Wiley, New York, NY.

Edwards, D.J., Guess, F.M., and Young T.M., (2004). Bootstrap Confidence Intervals for Percentiles of Reliability of Modern Engineered Wood, submitted. Email: tmyoung1@utk.edu or fguess@utk.edu directly for copies.

Efron, B. (2003). Second Thoughts on the Bootstrap. *Statistical Science*, **18** (2), 135-140.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, New York, NY.

English, L.P. (1999). *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits.* Wiley, New York, NY.

Guess, F.M., Edwards D.J., Pickrell T.M. and Young T.M. (2003). Exploring Graphically and Statistically the Reliability of Medium Density Fiberboard. *International Journal of Reliability and Application*, **4** (4), 97-109.

Guess, F., Hollander, M., and Proschan, F. (1986). Testing Exponentiality Versus a Trend Change in Mean Residual Life. *Annals of Statistics*, **14** (4), 1388-1398.

Guess, F. and Proschan, F. (1988). Mean Residual Life: Theory and Applications. *Handbook of Statistics: Quality Control and Reliability*, **7**, 215-224.

Guess, F., Walker, E., and Gallant, D. (1992). Burn-in to Improve which Measure of Reliability? *Microelectronics and Reliability*, **32**, 759-762.

Huang, K-T., Lee, Y. L. and Wang, R. Y. (1999). *Quality Information and Knowledge.* Prentice Hall, New York, NY.

Kim, K. O. and Kuo, W. (2003). Percentile Life and Reliability as Performance Measures in Optimal System Design. *IIE Transactions*, **35**, 1133-1142.

Kuo, W., Chien, W. T. K., and Kim, T. (1998). *Reliability, Yield, and Stress Burn-In.* Kluwer Academic Publishers, Norwell, MA.

Kuo, W., Prasad, V. R., Tillman, F. A., and Hwang, C. L. (2000). *Optimal Reliability Design: Fundamentals and Applications*. Cambridge University Press, Cambridge, UK.

Martinez, W. L. and Martinez, A. R. (2002). *Computational Statistics Handbook with Matlab*. Chapman and Hall/CRC, Boca Raton, LA.

Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. Wiley, New York, NY.

Redman, T. C. (2001). *Data Quality: The Field Guide*. Butterworth-Heinemann Digital Press, Boston, MA.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York, NY.

Walker, E. and Guess, F. (2003). Comparing Reliabilities of the Strength of Two Container Designs: A Case Study. *Journal of Data Science*, **1**, 185-197.

Young, T.M. and Guess, F.M. (2002). Mining Information in Automated Relational Databases for Improving Reliability in Forest Products Manufacturing. *International Journal of Reliability and Applications*, **3** (4), 155-164.