

To the Graduate Council:

I am submitting herewith a thesis written by Yang Wang entitled “Reliability Analysis of Oriented Strand Board’s Strength with a Simulation Study of the Median Censored Method for Estimating of Lower Percentile Strength.” I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of requirements for the degree of Master of Science, with a major in Statistics.

Timothy M. Young, Major Professor

We have read this thesis
And recommend its acceptance:

Dr. Frank Guess

Dr. Ramón León

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the
Graduate School

(Original signatures are on file with official student records.)

**Reliability Analysis of Oriented Strand Board's Strength with a Simulation Study of
the Median Censored Method for Estimating of Lower Percentile Strength**

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Yang Wang
August 2007

DEDICATION

This thesis is dedicated to my parents
for their unconditional love and support since the day I was born.
Even though they are on the other side of the Earth, I feel their love and keep
them in my heart every day.
Without them, I could never have come this far.
Also I dedicate this thesis to my husband.
He always stands behind me and believes in me that I can achieve anything I
desire and work for.

PROFESSIONAL ACKNOWLEDGEMENTS

This work was supported by the United States Department of Agriculture, Special Wood Utilization Grants Program as administered by The University of Tennessee Institute of Agriculture Experiment Station, contract R11-2219-150, and the Tennessee Agricultural Experiment Station McIntire Stennis project TEN00MS-89. Special thanks for funding support goes to the University of Tennessee Forest Products Center, Dr. Timothy Rials, Professor and Director and Dr. J. Larry Wilson, Professor and former Interim Head, Department of Forestry, Wildlife and Fisheries. I would also like to thank the Committee Co-chairs and Research Associate Professor Timothy Young, Professor Dr. Frank Guess, and Professor Dr. Ramón León, for their guidance throughout this research project. These scholars have provided exceptional guidance, and it has been an honor to work with them. They have helped to provide wonderful growth in my graduate work and my future professional career as a statistician.

In addition, I appreciate Dr. Russell Zaretzki and Dr. Sharon Husch for their advice and constant encouragement. Special thanks to Dr. Nicolas André for his technical support on statistical software. I would like to thank Graduate Research Assistants Leslie Shaffer and Diane Perhac for proofreading papers, and thank senior PhD graduate student David Edwards for kindly sharing part of his Matlab codes. In addition, I thank the other members of the faculty and staff of the Department of Statistics and the Tennessee Forest Products Center. It has been good for my personal growth and has taught me the value of networking.

ABSTRACT

Oriented Strand Board (OSB), an engineered wood product, has gained increased market acceptance as a construction material. Because of its growing market, the product's manufacturing and performance have become the focus of much research. Internal Bond (IB) and Parallel and Perpendicular Elasticity Indices (EI), are important strength metrics of OSB and are analyzed in this thesis using statistical reliability methods.

The data for this thesis consists of 529 destructive tests of OSB panels. They were tested from July 2005 to January 2006. These OSB panels came from a modern OSB manufacture in the Southeastern United States with the wood furnish being primarily Southern Pine (*Pinus spp.*). The 529 records are for 7/16" thickness OSB strength, which is rated for roof sheathing (i.e., 7/16" RS).

Descriptive statistics of IB and EI are summarized including mean, median, standard deviation, Interquartile range, skewness etc. Visual tools such as histograms and box plots are utilized to identify outliers and improve the understanding of the data. Survival plots or Kaplan-Meier curves are important methods for conducting nonparametric analyses of life (or strength) reliability data and are used in this thesis to estimate the strength survival function of the IB and EI of OSB. Probability Plots and Information Criteria are used to determine the best underlying distribution or probability density function. The OSB data used in this thesis fit the lognormal distribution best for both IB and EI. One outlier is excluded for the IB data and six outliers are excluded for the EI data.

Estimation of lower percentiles is very important for quality assurance. In many reliability studies, there is great interest in estimating the lower percentiles of life or strength. In OSB, the lower percentiles of strength may result in catastrophic failures during installation of OSB panels. Catastrophic failure of 7/16" RS OSB, which is used primarily for residential construction of roofs, may result in severe injury or death of construction workers. The liability and risk to OSB manufacturers from severe injury or death to construction workers from an OSB panel failure during construction can result in extreme loss of market share and significant financial losses.

In reliability data, "multiple failure modes" is common. Simulated data of mixed distribution of the two two-parameter Weibull distribution is produced to mimic the multiple failure modes for the reliability data. A forced median censored method is adopted to estimate lower percentiles of the simulated data. Results of the simulation indicate that the estimated lower percentiles median censored method is relatively close to the true parametric percentiles when compared to not using the median censored method. I conclude that the median censoring method is a useful tool for improving estimation of the lower percentiles for OSB panel failure.

TABLE OF CONTENTS

CHAPTER	Page
CHAPTER I. INTRODUCTION	1
CHAPTER II. LITERATURE REVIEW	4
2.1 <i>Brief Overview of Engineered Wood</i>	<i>4</i>
2.2 <i>Introduction of Oriented Strand Board</i>	<i>8</i>
2.3 <i>Manufacturing Process of OSB</i>	<i>11</i>
2.4 <i>Mechanical Testing of OSB</i>	<i>13</i>
2.5 <i>Maximum Likelihood Parameter Estimation and Information Criteria</i>	<i>17</i>
2.6 <i>Reliability Data Analysis and Weibull distribution</i>	<i>18</i>
2.7 <i>Mixed two-parameter Weibull Distribution.....</i>	<i>21</i>
CHAPTER III. METHODS	25
3.1 <i>OSB Data Set.....</i>	<i>25</i>
3.2 <i>Internal Bond and Elasticity Indices</i>	<i>25</i>
3.3 <i>Simulated Data of Mixed Two-Parameter Weibull Distribution</i>	<i>26</i>
3.4 <i>Calculation of the Maximum Likelihood Estimator of Percentiles.....</i>	<i>27</i>
3.5 <i>Newton-Raphson Method</i>	<i>28</i>
3.6 <i>Forced Median Censored Method</i>	<i>29</i>
3.7 <i>Estimated Percentiles of Mixed Model</i>	<i>30</i>
CHAPTER IV. EXPLORING GRAPHICALLY AND STATISTICALLY RELIABILITY OF TENSILE STRENGTH (IB) IN OSB	32
CHAPTER V. EXPLORING GRAPHICALLY AND STATISTICALLY RELIABILITY OF STIFFNESS STRENGTH PARALLEL (EI) IN OSB.....	46
CHAPTER VI. PROTECTION OF CONTAMINATION FROM WEAK UNITS IN ESTIMATION OF SMALL PERCENTILES BY USING FORCED MEDIAN CENSORED DATA	55
CHAPTER VII. CONCLUSIONS	71

REFERENCES	75
APPENDIX	79
VITA	85

LIST OF TABLES

Table	Page
Table 1. OSB internal bond (IB) descriptive statistics.	33
Table 2. Selected model scores for the internal bond data, complete and excluding outliers.....	43
Table 3. OSB parallel EI descriptive statistics.....	47
Table 4. Selected model scores for the parallel EI data complete and excluding outliers.	53
Table 5 Comparison of the parametric 1st, 5th, 10th, 15th percentiles of the weak model and the strong model.....	61
Table 6. Comparison of 1 st percentile from the whole simulated data and median censored simulated data with the parametric 1st percentile having different levels of contamination from the weak Weibull distribution.	63
Table 7. Comparison of 5 th percentile from the whole simulated data and median censored simulated data with the parametric 5 th percentile having different levels of contamination from the weak Weibull distribution.	64
Table 8. Comparison of 10 th percentile from the whole simulated data and median censored simulated data with the parametric 10 th percentile having different levels of contamination from the weak Weibull distribution	65
Table 9. Comparison 15 th percentile from the whole simulated data and median censored simulated data with the parametric 15 th percentile with different levels of contamination from units from weak Weibull distribution.....	66

LIST OF FIGURES

Figure	Page
Figure 1. Illustration of OSB wood strands, panels and uses in construction.	9
Figure 2. OSB manufacturing process (Courtesy of Structural Board Association, Willowdale, Ontario, Canada.)	12
Figure 3. Applying tensile load to internal bond specimen. Also shown is failed specimen. http://forest.mtu.edu/research/woodprotection/research/mechanical.html	15
Figure 4. The stiffness test is applied to a specimen. See http://forest.mtu.edu/research/woodprotection/research/mechanical.html	16
Figure 5. The effect of the value of b on the Weibull failure rate(from the link	22
Figure 6. The bathtub curve of the mixed Weibull distribution with different shape parameter. See the link http://www.Weibull.com/hotwire/issue14/relbasics14.htm 23	23
Figure 7. OSB internal bond histogram and boxplot from JMP.	35
Figure 8. OSB internal bond probability plots from S-PLUS and SPLIDA-Smallest extreme value distribution.....	37
Figure 9. OSB internal bond probability plots from S-PLUS and SPLIDA-Weibull distribution	38
Figure 10. OSB internal bond probability plots from S-PLUS and SPLIDA-Normal distribution	39
Figure 11. OSB internal bond probability plots from S-PLUS and SPLIDA-lognormal	40

Figure 12. OSB internal bond probability plots from S-PLUS and SPLIDA-Largest extreme value	41
Figure 13. OSB internal bond probability plots from S-PLUS and SPLIDA-Frechet.....	42
Figure 14. Reliability Kaplan-Meier Plot of internal bond of OSB.....	45
Figure 15. OSB parallel EI histogram and boxplot from JMP.	48
Figure 16. OSB internal bond probability plots from S-PLUS and SPLIDA-Lognormal	49
Figure 17. OSB parallel EI probability plots from S-PLUS and SPLIDA-Loglogistic...	50
Figure 18. OSB parallel EI probability plots from S-PLUS and SPLIDA-Largest extreme value.....	51
Figure 19. Reliability Kaplan-Meier plot of parallel EI of OSB.	54
Figure 20. The histogram of simulated data from the mixed model with zero contamination from the weak model.....	56
Figure 21. The histogram of simulated data from the mixed model with 25% contamination from the weak model.....	57
Figure 22. The histogram of simulated data from the mixed model with 50% contamination from the weak model.....	58
Figure 23. The histogram of simulated data from the mixed model with 75% contamination from the weak model.....	59
Figure 24 The histogram of simulated data from the mixed model with 100% contamination from the weak model.....	60
Figure 25. a) Comparison of 1st percentile from the whole simulated data and median censored simulated data with the parametric 1st percentile having different levels of contamination from the weak Weibull distribution b) Comparison of 5 th percentile	

from the whole simulated data and median censored simulated data with the parametric 5th percentile having different levels of contamination from the weak Weibull distribution 68

Figure 26. a) Comparison of 10th percentile from the whole simulated data and median censored simulated data with the parametric 10th percentile having different levels of contamination from the weak Weibull distribution b) Comparison of 15th percentile from the whole simulated data and median censored simulated data with the parametric 15th percentile having different levels of contamination from the weak Weibull distribution 69

CHAPTER I. INTRODUCTION

In this thesis, statistical reliability ideas and tools are applied to help assess, manage, and improve the strength of the engineered wood product, Oriented Strand Board (OSB). This product has gained increased market acceptance as a construction material in almost all geographical areas of Canada and United States. Because of its growing market, the product's manufacturing and performances have become the focus of much research. Common physical and mechanical properties of OSB strength and product quality are Modulus of Rupture (MOR), Modulus of Elasticity (MOE), Internal Bond (IB), Parallel and Perpendicular Elasticity Indices (EI), Density and Thickness Swell (TS). In this thesis, the IB and EI of OSB are examined in the context of reliability methods.

Product "life" for OSB can be measured in terms of the strength to failure, as opposed to the time to failure. The strength to failure is a crucial reliability parameter of the product. Estimation of the strength allows the producer to make assurances to customers about the safe, useful "strength" range of the product. We apply insightful statistical reliability tools to manage and seek improvements in the strengths of OSB. As a part of the OSB continuous manufacturing process, the product undergoes destructive testing at frequent time intervals during the manufacturing process to determine compliance with customers' specifications and OSB standards required by governing associations (e.g.,

American Plywood Association (APA), <http://www.apawood.org/>, Structural Board Association (SBS), <http://www.osbguide.com/about.html>). Workers in the destructive lab perform these tests on sample cross sections of the OSB panel to measure the tensile strength, also called internal bond (IB), in English units of pounds per square inch (or International units of kilograms per cubic meter) until failure. Additional stiffness tests of strength include parallel and perpendicular elasticity indices (EI), which are taken from cross sectional samples of the OSB panel in the parallel and perpendicular directions with respect to the orientation of the wood strands.

The lower percentiles of IB and EI may be of particular interest for the manufacturing company, oversight associations, residential construction companies, and consumers in specifying the product reliability of OSB. A novel technique called median censoring to weight lower observations is used. I study the forced median censored method in estimating the lower percentiles from simulated data of the mixture of two Weibull distributions. These simulations show the median censored method estimates lower percentiles very close to the true parametric percentiles. The median censored method provides strong improvements and protection in many places, while in other cases caution in its use is recommended. The simulations establish the median censoring method as a useful tool for improving estimation of the lower percentiles. Chapter II of the thesis is a literature review of OSB, the information criteria, maximum likelihood

estimation and the concept of the mixed model. Chapters III and IV focus on descriptive statistics and probability plots of the IB and EI of manufactured OSB. Chapter V explains the development of the simulated data, application of the median censored method, parametric percentile estimation from the mixed model and compares the parametric percentile with the percentile from the simulated data and median censored simulated data.

Chapter VI of the thesis is the conclusion. Suggestions for future research are also presented in this chapter.

CHAPTER II. LITERATURE REVIEW

2.1 Brief Overview of Engineered Wood

Wood products include paper, building, pallets, packaging, furniture, energy uses, and engineered wood. One of the important uses of wood is engineered wood (also known as wood composites.) Composites combine two or more materials -- typically a reinforcing material with a resin -- to produce car parts, building materials, and other products. (www.woodconsumption.org). Engineered wood includes a range of derivative wood products that are manufactured by binding together wood strands, particles, fibers, or veneers with adhesives to form composite materials. Engineered wood products are preferred over solid wood in many applications since it has comparative advantages. Some advantages are: 1) large panels and beams that span long lengths can be constructed using engineered wood; 2) small pieces of wood and wood residues can be used in many engineered wood products, especially particleboard and fiber-based boards (e.g., medium density fiberboard); 3) engineered wood can be designed to meet application-specific performance requirements (e.g., laminated veneer lumber used for long spans in roofing construction, oriented strand lumber as a substitute for solid wood studs, etc.); and 4) engineered wood products are often stronger and less prone to humidity-induced warping than equivalent solid woods.

Although engineered wood products may be more expensive than solid lumber, it may have some economic advantages, e.g., less expensive wood residue furnish than using standing trees. A current disadvantage that makes engineered wood products more expensive is the high weight-to-freight ratio, i.e., engineered wood panels are generally heavier than solid wood and are therefore more expensive to ship to market locations, especially given the current cost of diesel fuel. Engineered wood however, may be an environmentally wiser choice since it partially uses wood residues and also fully utilizes trees from natural forests more efficiently.

Some engineered woods (e.g., medium density fiberboard and particleboard) have urea-formaldehyde adhesives that may be toxic. In June of 2004, the formaldehyde issue and cancer emerged with a vengeance with the news that the International Agency for Research into Cancer (part of World Health Organization), had upgraded formaldehyde from category 2A (probably carcinogenic to humans) to Category 1 (carcinogenic to humans), see Sharp (2004). Reclassification was based on evidence of increased incidence of the relatively rare, naso-pharyngeal cancer among individuals exposed in the past to high levels of formaldehyde (Sharp 2004).

An example of the impending litigation potential of formaldehyde poisoning from wood composites is illustrated in Spake (2007). An estimated 275,000 Americans are living in

more than 102,000 travel trailers and mobile homes that FEMA purchased for \$2.6 billion after hurricane Katrina (Spake 2007). A class-action lawsuit was filed against FEMA and some trailer manufacturers in Louisiana in June 2006 on behalf of residents suffering from respiratory and flu-like illnesses they attribute to formaldehyde inside their trailers. According to the lab's report formaldehyde levels in the living room of one were more than three times the EPA's limit (Spake 2007).

OSB is bonded using Phenol-formaldehyde resins (PF), melamine-formaldehyde resin (MF) and Methylene diphenyl diisocyanate (MDI) resins. PF is commonly used for exterior exposure products. MF is used in exposed surfaces in comparatively costly designs. MDI are expensive and waterproof.

There are many types of engineered laminated wood. The veneer-based types include plywood, Laminated Veneer Lumber (LVL) and Stamina wood. Flakes or particle-type engineered woods include Oriented Strand Board (OSB), Oriented Strand Lumber and Particleboard. Fiber-based engineered wood include Insulation board, Homasote, Masonite, Medium Density Fiberboard (MDF) and Hardboard. In addition, there is a relatively new extruded engineered wood/plastic composite called Wood-Plastic Composite (WPC), http://en.wikipedia.org/wiki/Engineered_wood (2007), Perhac (2007).

Today, wood-based composite materials are common place and can be found in housing, cabinetry, flooring and furniture. Because of their wide application and highly competitive industry, reliability methods can add great value to improving product quality and lowering manufacturing costs and we focus on improving OSB product quality in this thesis.

Steele (2006) discusses the use of Mean Residual Life (MRL) functions, and more specifically, unique “function domain sets” confidence intervals. This different breed of confidence interval allows the practitioner to identify opportunities for quality improvement as well as to make novel statements about the process. Chen (2005) built upon the work of Edwards (2004) by exploring the use and effectiveness of estimating extremely small percentiles, or early failures, of strength measurements for MDF (i.e., IB). Chen (2005) observed that the distribution of strength failure data for IB does not follow a perfectly Normal distribution, and notes that forcing a Gaussian model on these data sets may lead to erroneous conclusions and profit loss. Chen (2005) proposes a forced censoring technique to closer fit the tails of strength distributions. The information obtained from these new fits may reduce the number of field failures, improve product safety, and even reduce the cost of destructive testing. More information on these reliability methods can be found in the published work of Chen et al. (2006) and Guess et al. (2004). Edwards (2004) also applies reliability techniques to

improve production quality and safety of MDF. Edwards (2004) is also concerned with the extremely small percentiles, or early failures, of MDF. Edwards (2004) discusses the applications of Akaike's Information Criteria or AIC (Akaike 1974) and Bozdogan's Information Complexity Criteria (ICOMP) (Bozdogan 1988) to the extremely small percentiles of MDF. Modeling these failures can be challenging given the small amounts of data in the tails of the MDF failure distributions. Given the small sample size Edwards (2004) discusses the use of bootstrap techniques to provide more accurate estimation of lower percentile strength data.

2.2 Introduction of Oriented Strand Board

Oriented Strand Board (OSB) is a structural engineered wood composite panel that is formed under heat and pressure by pressing mats that are formed using gravity and flake orientation of wood strands. Wood strands are approximately 0.030 inch in thickness on average, 2 inches in width on average and 4 inches in length on average. The mats are pressed under heat and pressure in both multi-opening (i.e., "day-light") and continuous presses. OSB is used in residential and non-residential construction for sheathing in walls, floors, and roofs (**Figure 1**). OSB is the most commonly used structural engineered wood panel in new residential housing construction in North America. It is the sheathing panel of choice in North America, since it is engineered with great uniformity and strength. See <http://www.osbguide.com/faqs/faq1.html>.2007



Figure 1. Illustration of OSB wood strands, panels and uses in construction.

The industry is currently experiencing unprecedented growth in North America and Europe in new mill startups and mill capacity expansion. Since 1990, new startups of mills have increased by 85% to 65 mills, while production capacity has increased by more than 100%, to a record 28 billion square feet per year (Adair 2005).

OSB is aggressively replacing plywood as the primary sheathing used in new construction in North America. Approximately 65% of the 43 billion square feet of construction sheathing used in 2005 consisted of OSB, while the remaining 35% consisted of plywood sheathing (Adair 2005). Plywood sheathing continues to decline in use for construction sheathing. Note 73% percent of all OSB sheathing produced is used in residential housing construction.

Residential housing construction in the U.S. is predicted to decline from a record of almost 2.0 million annual new housing starts in 2005 to approximately 1.8 million housing starts by 2010 (Adair 2005). This projected 10% decline in housing starts, in conjunction with recent OSB capacity expansion, implies OSB producers will face tremendous downward pressure on pricing. These business pressures will require OSB manufacturers to maintain a strong focus on reliability, quality, and cost. The reliability methods outlined in this paper can be used to improve the quality of OSB sheathing, plus help lower manufacturing costs by reducing raw material inputs and by proactively

preventing potential wasted OSB. See Guess and Proschan (1988), Guess, Hollander and Proschan (1986), Guess, Walker, and Gallant (1992), Young and Guess (1994), Young, and Guess (2002), Guess, León, Chen, and Young (2004), Guess, Zhang, Young, and León (2005) for discussions on various measures and approaches to understanding reliability and reliability as applied to engineered wood panels.

2.3 Manufacturing Process of OSB

In the general manufacturing process of OSB, debarked logs (i.e., removal of bark) are heated in soaking tanks and then sliced into thin wood elements. The strands are dried, blended with resin and wax, and formed into thick, loosely consolidated mats that pressed under heat and pressure into large panels. An overview of the OSB manufacturing process was showed (**Figure 2**).

Typically, logs are debarked followed by a process called log conditioning. Log conditioning means that logs are soaked and /or heated in a log tank. Logs are sent to the stranding machine and long log disk or ring stranders are used as raw material to produce wood strands. Strands are stored in wet bins and then dried in a dryer. Many different types of dryers are available for the OSB manufacture such as a traditional triple-pass dryer, a single-pass dryer, a combination triple-pass/single-pass dryer or a three-section conveyor dryer. Dried strands are screened and sent to dry bins. The strand screener can

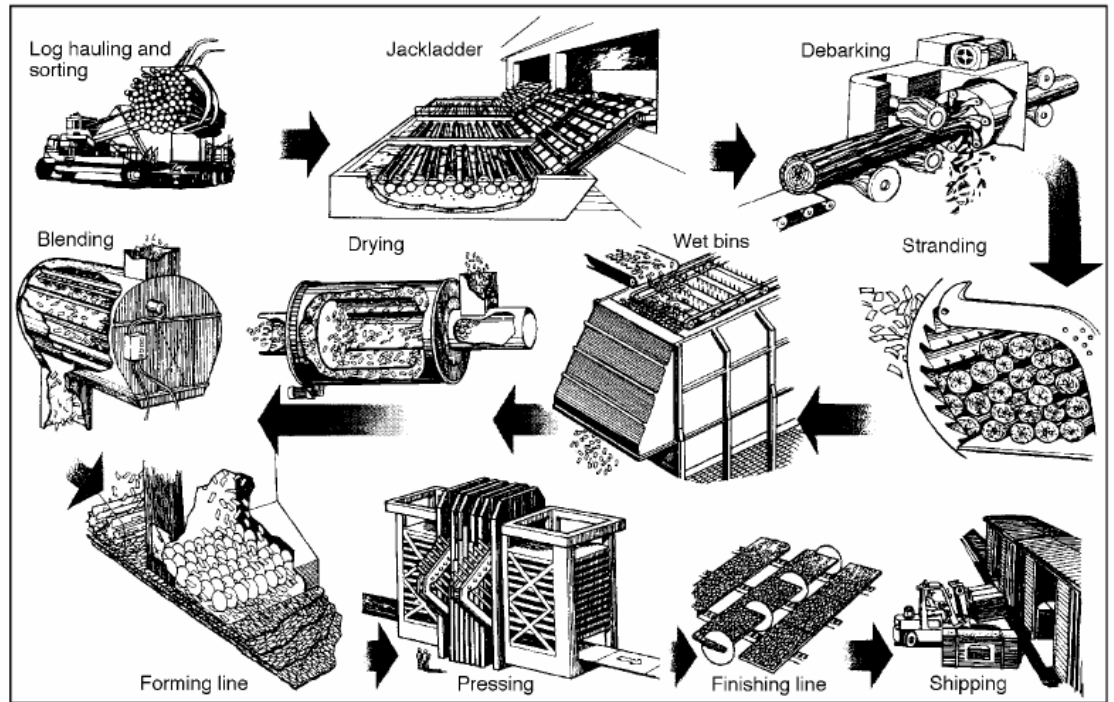


Figure 2. OSB manufacturing process (Courtesy of Structural Board Association, Willowdale, Ontario, Canada.)

be used to separate strands into three size fractions. Typically, different resin addition rates are used for face and core layers (i.e., higher addition rates in face, lower in core). Separate rotating blenders are used for adding the resin to face and core strands.

All mat formers use the long and narrow characteristic of the strand to place it between the spinning disks or troughs before it is ejected onto a moving screen or conveyor belt below the forming heads. Oriented layers of strands within the mat (face, core, face) are dropped sequentially, each by a different forming head. In hot pressing, the loose-layered mat of oriented strands is compressed under heat and pressure to cure the resin. The finished products are then subjected to a series of testing (Winandy, 2004)

2.4 Mechanical Testing of OSB

Tensile Strength

The internal bond (IB) test is commonly used for mat-formed engineered wood panels such as OSB, particleboard and MDF. IB strength is used as a quality control measure in plants and the test method is specified in standards such as the ASTM (American Society for Testing and Materials - ASTM <http://www.astm.org/cgi-bin/SoftCart.exe/index.shtml?E+mystore> 2007). IB strength is an important mechanical property of panel products, not only for industrial use but also for research and laboratory-scale tests.

Internal bond is a fundamental measure of the adhesive performance in wood composites. IB is a measure of the tensile strength that is calculated by pulling apart two inch by two inch OSB blocks using a destructive testing process (Figure 3). This strength is in large part determined by the effectiveness of the glue application in the composite manufacture, but is also affected by other variables such as flake size and density of the final composite. There is a lot of research on factors affecting IB strength, e.g., uniform distribution of resin, etc. Youngquist (1987), (<http://forest.mtu.edu/research/woodprotection/research/mechanical.html> 2007). The test methods comply with the standards of the Annual Book of ASTM Standards, 1989, D 1037 - Standard Methods of Evaluating the Properties of Wood-Base Fiber and Particle Panel Material.

Stiffness Strength

“Static bending refers to tests performed in which a bending stress is applied to the specimen to determine the stiffness, or modulus of elasticity of the specimen as well as the amount of force required to cause the specimen to fail, expressed as the modulus of rupture. The specimen size is dependent on the testing standard used, the material type, the original size, and intended end-use of the material being tested.” (Figure 4) <http://forest.mtu.edu/research/woodprotection/research/mechanical.html>



Figure 3. Applying tensile load to internal bond specimen. Also shown is failed specimen. <http://forest.mtu.edu/research/woodprotection/research/mechanical.html>



Figure 4. The stiffness test is applied to a specimen. See <http://forest.mtu.edu/research/woodprotection/research/mechanical.html>

2.5 Maximum Likelihood Parameter Estimation and Information Criteria

Many graphical statistical methods are used extensively in data analysis such as histograms, probability plots, etc. These methods provide a very direct impression to the viewer or reader through illustration, but are somewhat subjective and require personal assessment. Choosing a model based on information criteria is quantitative and objective and may reduce ambiguity that occurs from the viewing of a probability plot. Model selection given a data set is the task of selecting a statistical model from a set of potential models. In its most basic form, this is one of the fundamental tasks of scientific inquiry. Once the set of possible models are selected, the mathematical analysis allows us to determine the best of these models.

What is meant by “best” may be controversial in the scientific community. Good model selection techniques will balance goodness of fit and complexity. More complex models will be better able to adapt their shape to fit the data. Model selection methods include: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Deviance Information Criterion (DIC), etc.

Maximum Likelihood Parameter Estimation (MLE) is to determine the parameters that maximize the probability (likelihood) of the sample data. The method of maximum likelihood is considered more robust and yields estimators with good statistical properties.

MLE methods are versatile and are applicable to most models and different data types. In addition, MLE provides efficient methods for quantifying uncertainty through confidence bounds. Although the methodology for maximum likelihood estimation is simple, the implementation is mathematically intense. Today's computer power eliminates mathematical complexity as an obstacle.

Akaike's Information Criterion (AIC), developed by Hirotugu Akaike in 1971 and proposed in Akaike (1974), is a measure of the goodness of fit of an estimated statistical model. Akaike's Information Criterion (AIC), like other model-selection methods, takes the form of a lack of fit term (such as minus twice the log likelihood) plus a penalty term. In the general case, AIC has the following form:

$$AIC = -2 \log L(\hat{\theta}) + 2k \quad (1)$$

where $L(\hat{\theta})$ is the maximized likelihood function for a particular population parameter (either scalar or vector valued) and k is the number of parameters in the model. For example, if we consider the normal model with the parameters μ and σ^2 , then $k=2$.

2.6 Reliability Data Analysis and Weibull distribution

In today's technological world, we depend on, demand, and expect reliable products. When products fail, the results can be catastrophic: injury, loss of life and/or costly lawsuits can occur. This is extremely important for engineered wood product such as

OSB, plywood, MDF etc since they are widely used in residential and commercial construction. Shipping unreliable products can destroy a company's reputation in a very short time. Continual assessment of new product reliability and ongoing control of the reliability of everything shipped are critical necessities in today's competitive business arena. See <http://www.itl.nist.gov/div898/handbook/apr/section1/apr11.htm>.2007

Kaplan-Meier estimator (also known as the Product Limit Estimator) is one of important methods for conducting a nonparametric analysis for life/reliability data. It estimates the survival function from life-time data. An engineer might measure the time until failure of the products. A plot of the Kaplan-Meier estimate of the survival function is a series of horizontal steps of declining magnitude that, when a large enough sample is taken, approaches the true survival function for that population. The following is the formula of the survival function:

$$S(t) = \prod_{j=1}^t [(n-j)/(n-j+1)]^{\delta_{(j)}} \quad (2)$$

In the equation(2), $S(t)$ is the estimated survival function, n is the total number of cases, and $\prod$ denotes the multiplication(geometric sum) across all cases less than or equal to t ; $\delta_{(j)}$ is a constant that is either 1 if the j th case is uncensored(complete), and 0 if it is censored.

The Weibull distribution is a parametric analysis and widely used in reliability and

biomedical engineering because of goodness of fit to data and ease of handling (Weibull 1951) http://en.wikipedia.org/wiki/Waloddi_Weibull 2007. Inference on the quantiles of the distribution has been extremely important for reliability data.

The probability density function (pdf) of Weibull distribution with only shape and scale parameter is as below: for $x \geq 0$:

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha^\beta} x^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right], \quad (3)$$

The cumulative density function (cdf) is

$$F(x; \alpha, \beta) = 1 - e^{-(x/\alpha)^\beta} \quad (4)$$

for $x \geq 0$, and $F(x; \alpha; \beta) = 0$ for $x < 0$. The expected value and standard deviation of a Weibull random variable can be expressed as:

$$E(x) = \alpha \Gamma(1 + \beta^{-1}) \quad (5)$$

$$Var(x) = \alpha^2 [\Gamma(1 + 2\beta^{-1}) - \Gamma^2(1 + \beta^{-1})] \quad (6)$$

where Γ is the gamma function. The Weibull metrics include reliability function and failure rate function etc. The Weibull reliability function is given by

$$R(x) = e^{-(x/\alpha)^\beta} \quad (7)$$

The Weibull failure rate function is given by:

$$\lambda(x) = \frac{f(x)}{R(x)} = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} \quad (8)$$

The data from the population of mixed Weibull distribution represent multiple failure modes. During the early life or strength period, the substandard components fail, and are

moved from the population which cause a decreasing failure rate. The decreasing failure rate will stop when all the substandard components fail. The life characteristic is depicted by the Weibull distribution with $\beta < 1$. After all the substandard components fail and are removed, the components will fail by chance which will result in a relatively stable failure rate. The failure by chance can be explained by sudden, unpredictable stress Weibull distribution with a $\beta > 1$. Applications that have a stress level above those to which the product is designed. This is the second type of failure.

After specific length of operation time (or strength) accumulation, the failure rate will increase since as age (or stress) increases, more and more components will fail. This is the third type of failure that can be modeled by the normal distribution. This is the characteristic of the distribution where the value of β has a distinct effect on the failure rate. **(Figure 5)** These comprise the three sections of the classic "bathtub curve." A mixed Weibull distribution with one subpopulation with $\beta < 1$, one subpopulation with $\beta = 1$ and one subpopulation with $\beta > 1$ would have a failure rate plot that was identical to the bathtub curve. An example of a bathtub curve is showed **(Figure 6)**.

2.7 Mixed two-parameter Weibull Distribution

The Weibull distribution has been used to model times-to-failure data successfully. However, when a product has two or more failure modes for causes, the appropriate

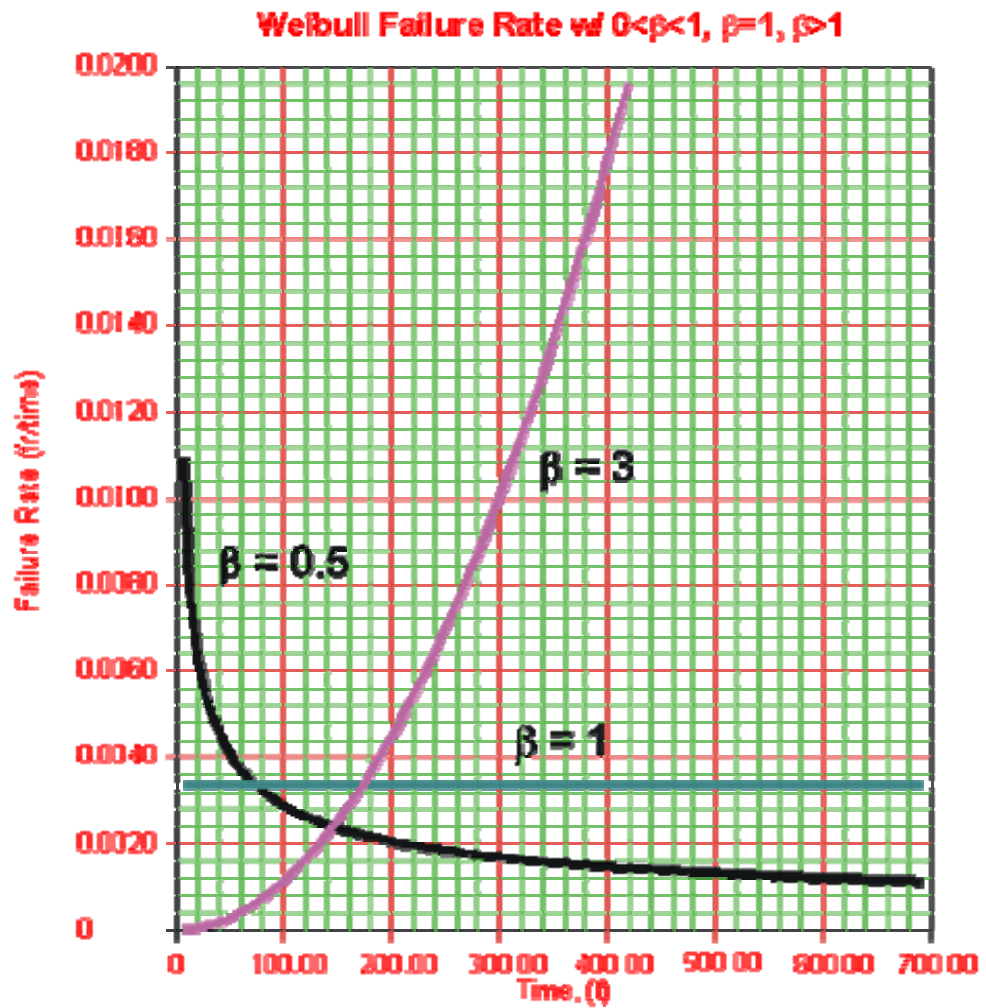


Figure 5. The effect of the value of b on the Weibull failure rate(from the link <http://www.Weibull.com/hotwire/issue14/relbasics14.htm>)

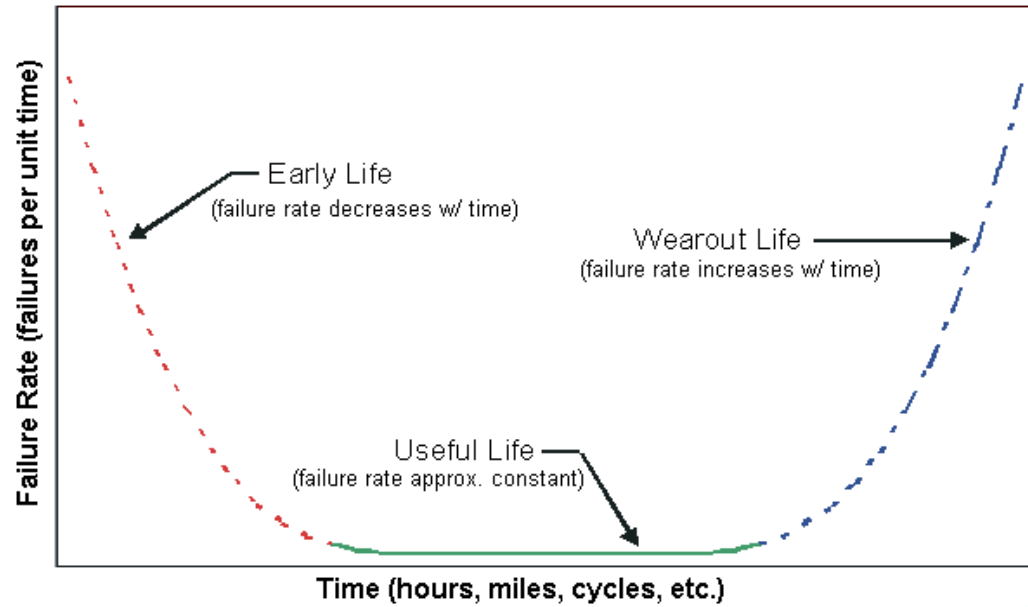


Figure 6. The bathtub curve of the mixed Weibull distribution with different shape parameter. See the link <http://www.Weibull.com/hotwire/issue14/relbasics14.htm>

mixed-Weibull distribution must be used. A mixed Weibull distribution represents a population that consists of several Weibull subpopulations. The mixed Weibull distribution (also known as a multimodal Weibull) is used to model data that do not fall on a straight line on a Weibull probability plot. The mixed Weibull distribution provides a good model for OSB products when the failure is caused by more than one failure mode. If the population consists of a mixture of two independent subpopulation with no correlation and each subpopulation has its own unique failure mode and distribution, then the cumulative density function for the mixed population can be expressed by the following:

$$F(x) = cF_1(x) + (1 - c)F_2(x) \quad (9)$$

where $F(x) = cdf$ of the mixed population, $F_j(x) = cdf$ of the j th subpopulation, $j=1, 2$, c =mixed weight, $c \in (0,1)$ usually, a subpopulation can be described by a single Weibull distribution. The cdf of the single Weibull distribution is described in the equation. The probability density function for mixed two Weibull distribution can be expressed by the following:

$$f(x) = cf_1(x) + (1 - c)f_2(x) \quad (10)$$

where $f(x) = pdf$ of the mixed population, $f_j(x) = pdf$ of the j th subpopulation, $j=1, 2$, c =mixed weight, $c \in (0,1)$, The pdf of the single Weibull distribution is described in equation (3).

CHAPTER III. METHODS

3.1 OSB Data Set

The data for this thesis consists of 529 destructive tests of OSB panels. They were tested from July 2005 to January 2006. These OSB panels came from a modern OSB manufacture in the Southeastern United States which used Southern Pine (*Pinus spp.*) wood furnish. The 529 records are for 7/16" thick OSB strength rated for roof sheathing (i.e., 7/16" RS).

3.2 Internal Bond and Elasticity Indices

In the previous literature review, I discussed the method for measuring IB strength and EI. In this chapter, I will give more details on this topic.

Composition board strength and physical properties has been very important for manufacturers since they are related to the safety of customers and construction workers. In addition, the product is expected to resist racking and shape distortion under high wind and earthquake forces. Many lawsuits are proposed because of OSB products' weakness in strength or/and physical properties in the field. Quality assurance is the highest priority for wood composite manufactures. A lot of parameters influence wood composite's strength and physical properties such as resin distribution, resin droplet size,

wood species, and furnish size and geometry (Youngquist 1986). IB test, a destructive test, provides information on how well the wood pieces have bonded together.

IB tests are performed according to ASTM D 1037-72 (American Society for Testing and Material 1981).

The Modulus of Elasticity is a measurement of the stiffness strength of a OSB panel. When a material is subjected to an external load, it may become distorted or strained. The material returns to its original dimensions when the load is removed providing that the loading is not too large to reach the rupture phase. Within the limits of elasticity, the ratio of the linear stress to the linear strain is termed the Modulus of Elasticity (EI). Specimens are tested by ASTM D6874-03 Standard Test Methods.

For more details of the IB and EI measurements, American Society for Testing and Material visit website <http://astm.org/cgi-bin/SoftCart.exe/index.shtml?E+mystore>. 2007

3.3 Simulated Data of Mixed Two-Parameter Weibull Distribution

MATLAB is used to produce the simulation data from the two Weibull distributions with different scale and shape parameters. One model with smaller value of median, percentile is called the “weak model” The second model is called the “strong model”.

The percentages of the data that randomly come from the “weak model” and “strong model” are designed in a systematic way in order to show the relationship of mixed leverage and the estimation of percentile.

3.4 Calculation of the Maximum Likelihood Estimator of Percentiles

The probability density function (p.d.f.) of the two parameter. Weibull distribution is given by the equation (11)

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha^\beta} x^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right], \quad (11)$$

where $\alpha > 0$ is the scale parameter (the characteristic life), $\beta > 0$ is the shape parameter.

Let $x_1, x_2, \dots, x_n$ be a random sample of size n observations (say failure times) which are distributed as Weibull with parameters α and β . The likelihood function of the same is given by equation (12)

$$L(\alpha, \beta \mid x_1, x_2, \dots, x_n) = \left(\frac{\beta}{\alpha}\right)^n \prod_{i=1}^n x_i^{\beta-1} \exp\left[-\left(\frac{\sum_{i=1}^n x_i}{\alpha}\right)^\beta\right] \quad (12)$$

Taking the log of equation (12), the log likelihood function is

$$l(\alpha, \beta) = \ln L(\alpha, \beta) = n \ln \beta - n \ln \alpha + (\beta - 1) \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \frac{x_i^\beta}{\alpha^\beta} \quad (13)$$

Differentiating equation (13), and equating the partial derivatives equal to zero, we obtain equation (14) and (15)

$$\frac{\partial \ln L}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^2} \sum_{i=1}^n x_i^\beta = 0 \quad (14)$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \ln x_i - \frac{1}{\alpha} \sum_{i=1}^n x_i^\beta \ln x_i = 0 \quad (15)$$

The maximum likelihood estimators $\hat{\alpha}, \hat{\beta}$ of the scale and shape parameters are the solution of the simultaneous equations:

$$\hat{\alpha} = \left[\left(\frac{1}{n} \right) \sum_{i=1}^n x_i^{\hat{\beta}} \right]^{\frac{1}{\hat{\beta}}} \quad (16)$$

$$\hat{\beta} = \frac{n}{\left(\frac{1}{\hat{\alpha}} \right)^{\hat{\beta}} \sum_{i=1}^n x_i^{\hat{\beta}} \ln x_i - \sum_{i=1}^n \ln x_i} \quad (17)$$

$\hat{\beta}$ can be obtained by Newton-Raphson or other suitable iterative methods since we can not obtain the MLE's of the two parameter Weibull in closed-form analytical expressions.

3.5 Newton-Raphson Method

Newton-Raphson method is used to find a root of a complicated function algebraically when you have some difficulty. Using some basic concepts of calculus, Newton-Raphson helps us have ways of numerically evaluating roots of complicated functions. It follows a set guideline to approximated one root, considering the function, its derivative, and an initial x-value. The Newton-Raphson method uses an iterative process to approach one root of a function. The specific root that the process locates depends on the initial,

arbitrarily chosen x-value:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (18)$$

where x_n is the current x-value, $f(x_n)$ represents the value of the function at x_n and $f'(x_n)$ is the derivative(slope) at x_n . x_{n+1} represents the next x-value that you are trying to find. Essentially, $f'(x)$, the derivative represents $f(x)/dx$. So x_{n+1} represents $x_n + dx$. The more iteration, the closer dx will be to zero. For more details, see the link <http://www.shodor.org/unchem/math/newton/index.html>. 2007

3.6 Forced Median Censored Method

A percentile is the value of a variable below which a certain percent of observations fall. The 25th percentile is also known as the first quartile; the 50th percentile as the median. In statistics, censoring occurs when the value of an observation is only partially known. The forced censored method means we exclude some observations from statistical analysis in the situation where these observations are collected already.

The data from the mixed Weibull models is ordered from the smallest to the largest. 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 (Chen 2005).

3.7 Estimated Percentiles of the Mixed Model

As mentioned previously, a percentile is the value of a variable below which a certain percent of observations fall. It is very important for quality assurance. In many reliability studies, interest centers around the estimation of lower percentiles. These lower numbers are useful for warranty analysis, understanding early failures during normal usage, and improving the specification limits.

Equation (19) is the cumulative distribution function of mixed model with two Weibull distribution with two parameters:

$$\bar{F}(x) = c\bar{F}_1(x) + (1 - c)\bar{F}_2(x) \quad (19)$$

Note “c” represents the percentage of data coming from the first Weibull distribution with its two parameters. The “1-c” represents the percentage of data coming from the second Weibull distribution with two parameters. $F_1(x)$ and $F_2(x)$ represent the individual cumulative distribution function of a specific Weibull model with two parameter as the below:

$$F(x; \alpha, \beta) = 1 - e^{-(x/\alpha)^\beta} \quad (20)$$

where $x \geq 0$, and $F(x; \alpha; \beta) = 0$ for $x < 0$. $\beta > 0$ is the shape parameter, $\alpha > 0$ is the scale parameter. $F(x)$ corresponds to the percentile probability that fit the definition of the cumulative distribution function. (Kececioglu 1998) The cumulative distribution

function (cdf) completely describes the probability distribution of a real-valued random variable X . For every real number x , the cdf of X is given by

$$x \rightarrow F_X(x) = P(X \leq x), \quad (21)$$

where the right-hand side represents the probability that the random variable X takes on a value less than or equal to x . When we calculate the parametric value of the percentile from the mixed model, the $F_X(x)$ is corresponding to the percentile. For example, 10th percentile corresponds to the solution of x when 0.1 in the above equation. Since we know the two Weibull parameters of the two Weibull distributions, and the alpha as the mixing level, we also know the probability of the left side of equation (19); we can solve the numeral equation.

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 with some Matlab (<http://www.mathworks.com>) in the analysis for the thesis. JMP[®] (<http://www.jmp.com>, a SAS[®] division), a statistical discovery software platform with scripting, is also used in the analysis for the thesis. Tutorials on the use of both software for reliability applications can be found at Professor Ramón V. León's course webpage at <http://web.utk.edu/~leon/>.

CHAPTER IV. EXPLORING GRAPHICALLY AND STATISTICALLY RELIABILITY OF TENSILE STRENGTH (IB) IN OSB

Our analysis of OSB begins with descriptive statistics of the internal bond strength. The data set was obtained from a modern OSB manufacturer located in the southeastern United States. The OSB manufacturer uses southern pine species wood and phenol resin during the manufacturing process. Our first step is calculating means, medians, percentiles, etc; our second step is producing box plots and histograms of the OSB strength data. We want to understand means, medians, percentiles, box plots, and histograms of the strengths of OSB. See, for example, Guess, Walker, and Gallant (1992) for how different measures of reliability can be used. Compare Guess, Edwards, Pickrell and Young (2003) for work on the modern engineered wood of medium density fiberboard.

A summary of descriptive statistics of internal bond that characterizes the location, variability, and shape of this data set was showed. (**Table 1**) The mean and median are location statistics. The standard deviation, coefficient of variation, and interquartile range (IQR) are variability statistics. The shape of the data can be further characterized by skewness and kurtosis. Skewness measures the direction and degree of asymmetry. A

Table 1. OSB internal bond (IB) descriptive statistics.

Statistics	IB (psi)
Mean	49.7
Median	48.5
Standard Deviation	11.3
Coefficient of Variation	22.7
IQR	15.6
Min	15.3
Max	90.0
Skewness	0.43
Kurtosis	0.02

positive value indicates skewness (long tailed) to the right while a negative value indicates skewness to the left. The value of 0.43 in this data suggests a mild positive skewness.

The histogram is an effective graphical technique for showing both the skewness and kurtosis of the data set. The histogram in **Figure 7** indicates that the internal bond is neither symmetrical nor normally distributed. Boxplots are one of the more used visual tools for summarizing a set of data measured on an interval scale. They are often utilized to show the shape of the distribution, its central value, variability, and outliers. In a boxplot graph, any points outside the whisker and the box are possible outliers. The boxplot indicates a potential outlier that is far less than the other observations in the data. (**Figure 7**) This possible outlier is highlighted as a thicker point to the left side of the box plot.

Probability plots are one of the most commonly used graphical techniques in the analysis of reliability data, because they are powerful visual tools that clearly demonstrate how a particular data set fits a specific candidate probability distribution(s). The data are ordered and then plotted against the theoretical order statistics for a desired distribution. If the data set “conforms” to a particular distribution, the points will form a straight line.

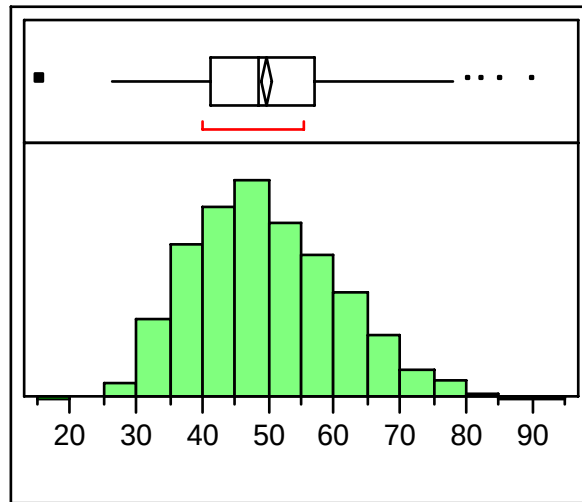


Figure 7. OSB internal bond histogram and boxplot from JMP.

Simultaneous confidence bands, along with pointwise confidence intervals (see Meeker and Escobar, 1999) can provide objective assessments of deviation from the line. Data points outside the confidence bands are shown to deviate from the candidate probability distribution in question. **(Figure 8 to 13)** See Chapter 6 of Meeker and Escobar (1998) for further information. Smallest extreme value (SEV), lognormal, largest extreme value (LEV), Frechet, normal, and Weibull probability plots were produced for this OSB internal bonding data using S-PLUS and SPLIDA.

From **Figure 11**, we see that, except for an outlier, the lognormal probability plot appears to fit the data best. The other probability plots have both outlier(s) plus departure in the lower or the upper tail. After the lognormal, according to Akaike's Information Criterion (AIC), the normal and loglogistic distributions fit the data best followed by the logistic, Weibull, Frechet, and SEV distributions. The probability plots provide a visual, subjective method for assessing the underlying distribution for the different product types. The lognormal distribution was determined to be a reasonable fit compared to Weibull, normal, plus five other distributions (we do not show all distributions here to save space).

Table 2 presents the log likelihood and AIC scores of select models. These serve as quantitative evidence of the best-fitting distribution model. The AIC for model selection (Akaike, 1973, Bozdogan, 2000) favors the model that minimizes AIC scores based on

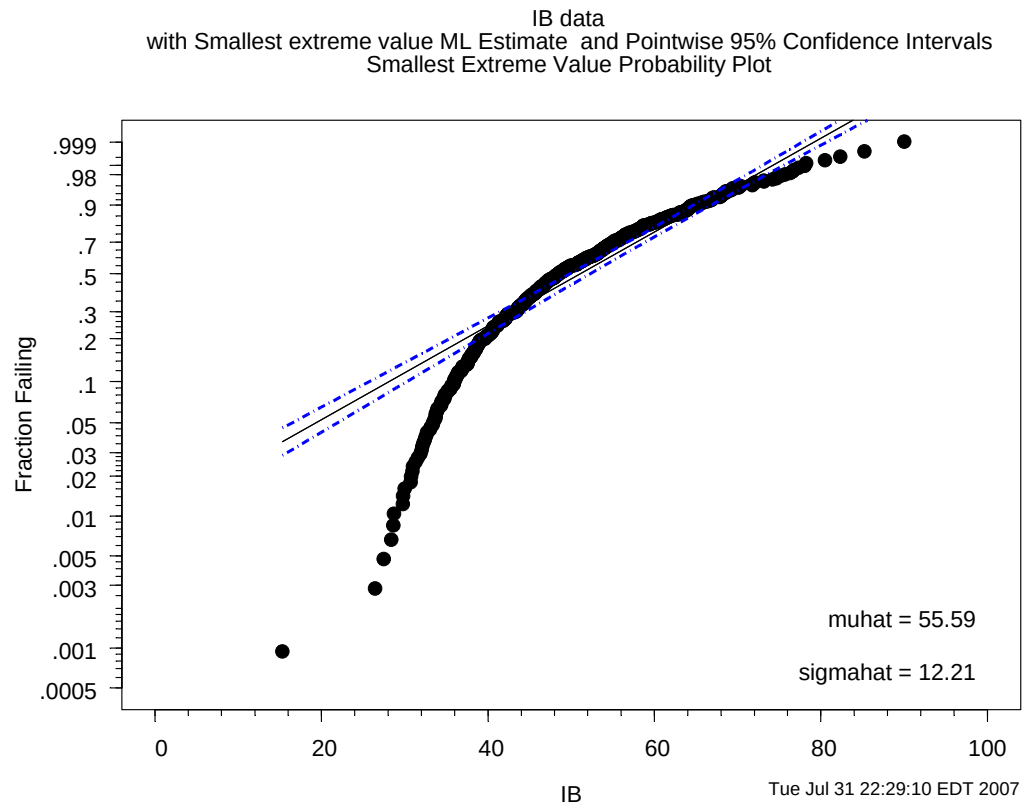


Figure 8. OSB internal bond probability plots from S-PLUS and SPLIDA-Smallest extreme value distribution

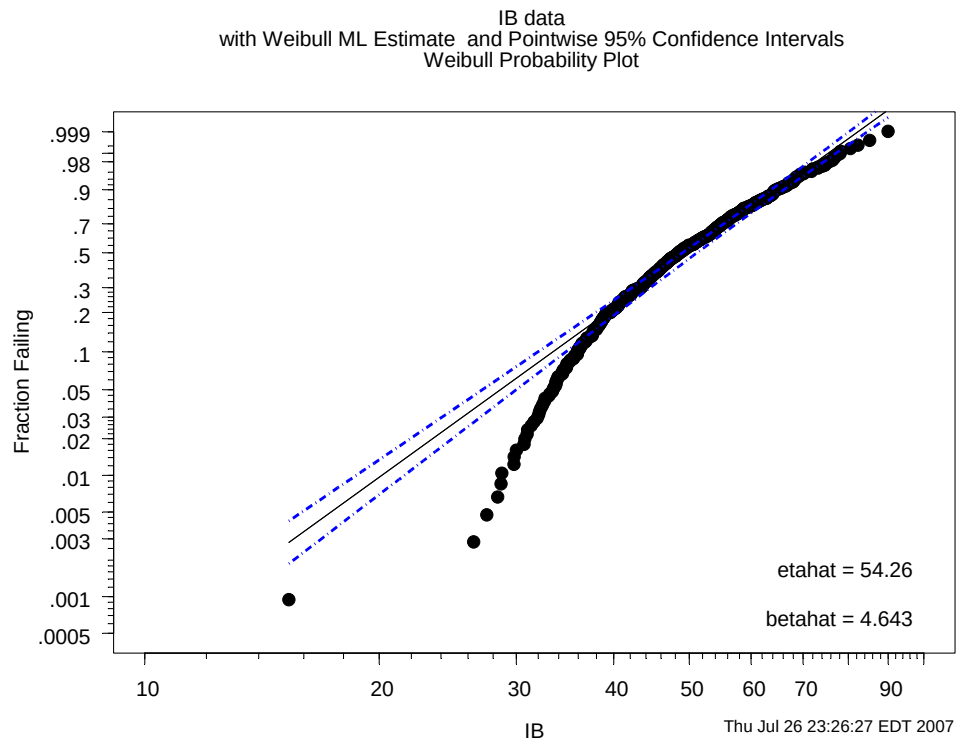


Figure 9. OSB internal bond probability plots from S-PLUS and SPLIDA-Weibull distribution

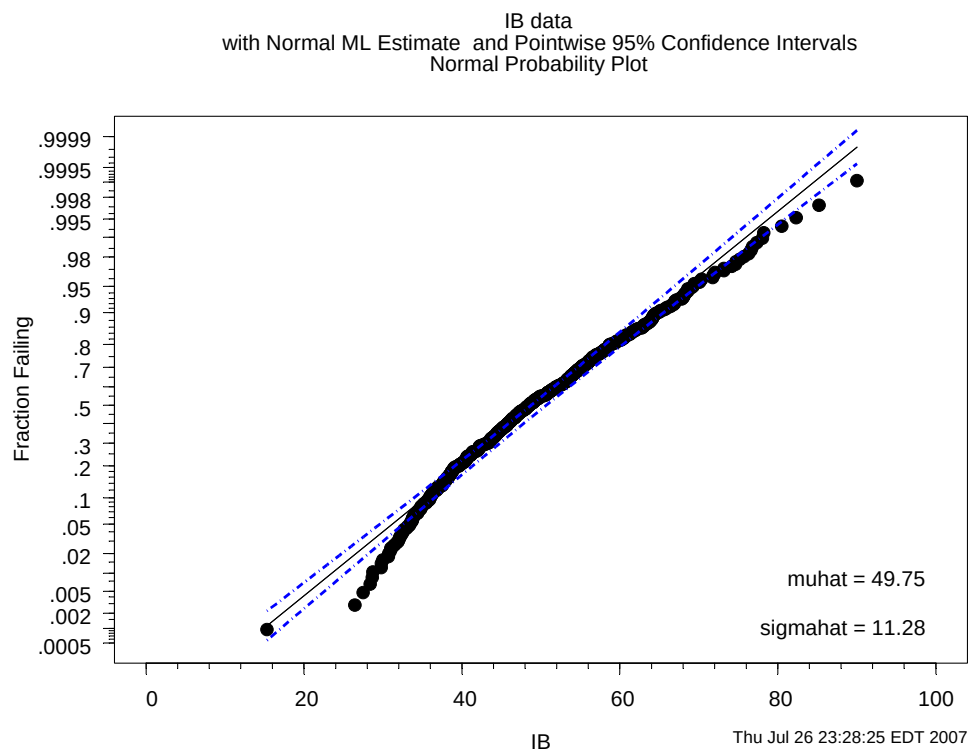


Figure 10. OSB internal bond probability plots from S-PLUS and SPLIDA-Normal distribution

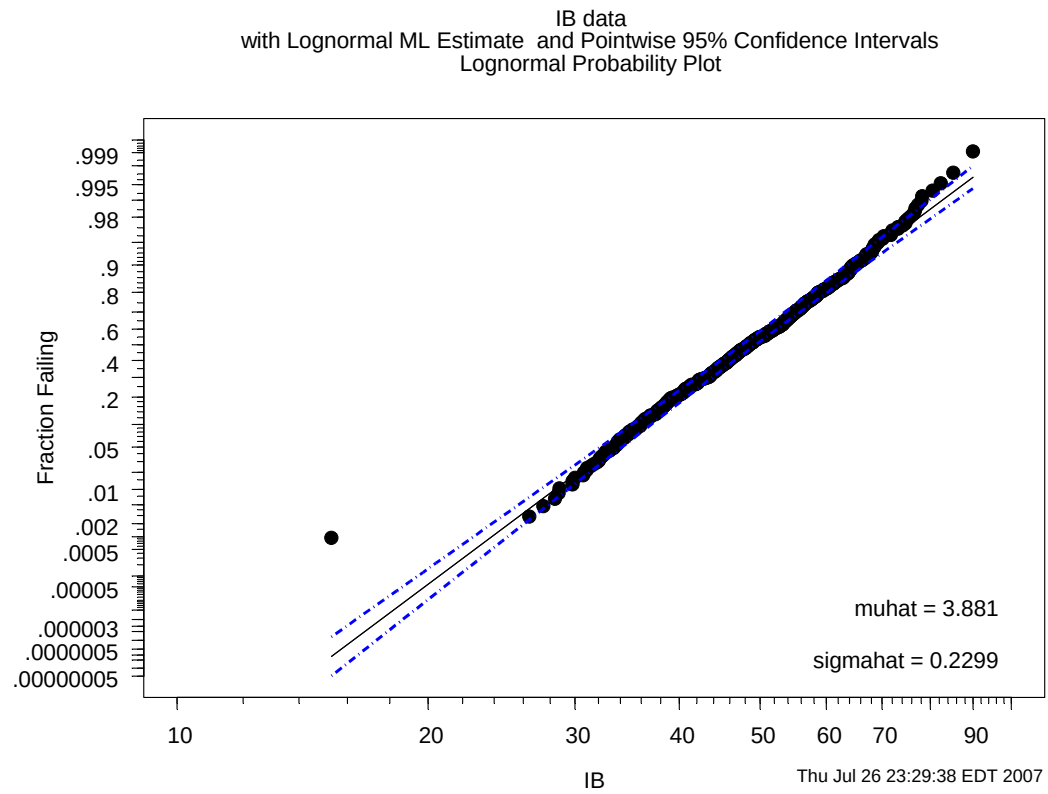


Figure 11. OSB internal bond probability plots from S-PLUS and SPLIDA-lognormal

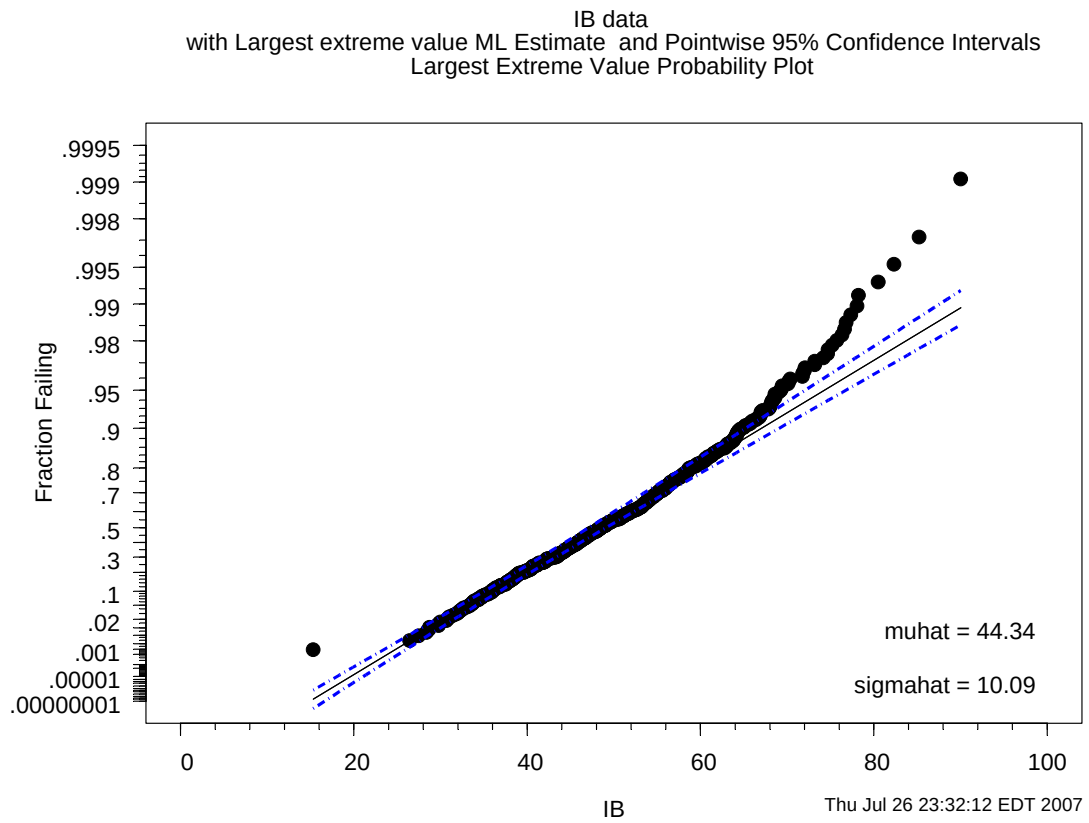


Figure 12. OSB internal bond probability plots from S-PLUS and SPLIDA-Largest extreme value

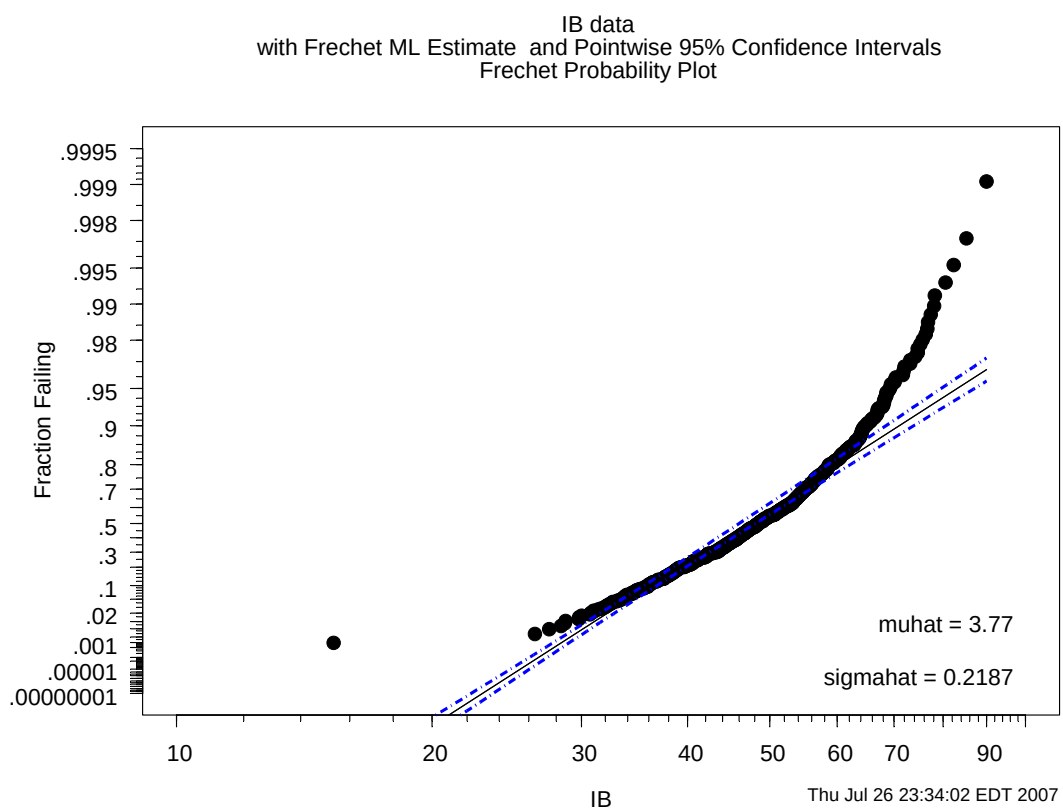


Figure 13. OSB internal bond probability plots from S-PLUS and SPLIDA-Frechet

Table 2. Selected model scores for the internal bond data, complete and excluding outliers.

Model fit	Complete data		Data with one outlier excluded	
	Log likelihood	AIC	Log likelihood	AIC
Lognormal	-2026	4056	-2011	4026
LEV	-2035	4074	-2017	4038
Loglogistic	-2032	4068	-2022	4048
Normal	-2032	4068	-2024	4052
Logistic	-2038	4080	-2031	4066
Logistic	-2038	4080	-2031	4066
Weibull	-2047	4098	-2040	4084
Frechet	-2082	4168	-2048	4100

the same data. Therefore, the lognormal fit is the best approximating model for the data. The AIC scores from the data without the outlier are smaller than the complete data, which is more empirical evidence that observation number 144 is indeed an outlier.

The reliability/survival function assesses the probability that the product will survive beyond a specified “time” or “pressure.” In our data, pressure to failure is measured. Kaplan-Meier plots (also called the Product Limit graphs) are one of the most popular survival plots. The Kaplan-Meier plot is a simple way of computing the survival curve in spite of troublesome data challenges, such as censored data. As we can see from the **Figure 14**, survival of OSB declines as pressure increases. For example, the probability that IB will be greater than 50 psi is approximately 0.50, while the probability that IB is greater than 65 psi is approximately 0.10. Statistically, 5% of OSB failed before a pressure of 33 psi and 95% of OSB failed before a pressure of 68 psi. The Kaplan-Meier plot for this OSB data set indicates that pressure to failure decreases at increasing rates between 35 and 65 psi.

The practitioner may use the Kaplan-Meier plots as an exploratory tool to estimate the effects of different wood and resin inputs for new product development, by comparing plots for the different rates. These comparisons may be very helpful for minimizing raw material inputs, while maintaining product reliability and reducing sources of variation.

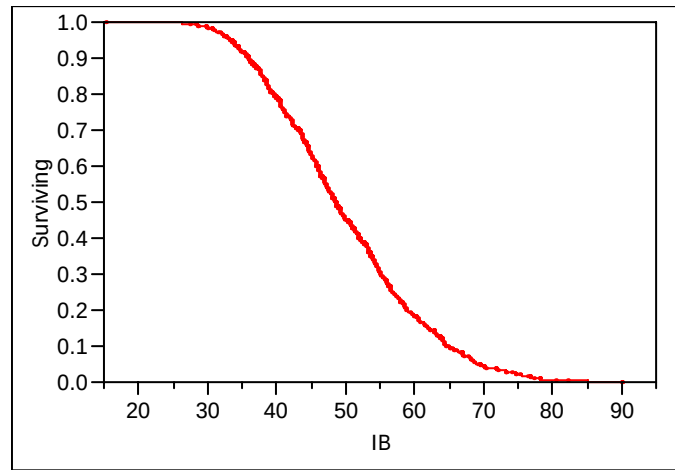


Figure 14. Reliability Kaplan-Meier Plot of internal bond of OSB.

CHAPTER V. EXPLORING GRAPHICALLY AND STATISTICALLY RELIABILITY OF STIFFNESS STRENGTH PARALLEL (EI) IN OSB

Table 3 is the summary of OSB parallel EI descriptive statistics. The skewness is 0.90, which indicates a positive skewness. The Kurtosis is 2.69, which indicates a “peaked” distribution.

The histogram and box plot are used to show the shape of the data. In Figure 5, the histogram appears slightly non-symmetrical, having an obvious right tail when compared to the normal distribution. **(Figure 15)** The box plot is a graphic that displays the center portions of the data and some information about the range of the data. There are three points on the left side of the whisker plus more than 10 points on the right side of the whisker, with an obvious gap between points on the right side. The distance of this gap is almost the value of the IQR. Therefore, we consider the three points on the far right side of the whisker and the three points on the left side of the whisker as outliers. In the following probability plots, the data excluding these six outliers are used. Probability plots for the lognormal, loglogistic, and largest extreme value distributions are shown in **Figure 16 to 18**. The lognormal distribution appears to be the best fit for the Parallel EI, followed by the loglogistic, and next the largest extreme value distributions, respectively.

Table 3. OSB parallel EI descriptive statistics.

Statistics	Parallel EI (psi)
Mean	58212
Median	57860
Standard Deviation	4205.7
Coefficient of Variation	7.22
IQR	4831
Min	45903
Max	79499
Skewness	0.90
Kurtosis	2.69

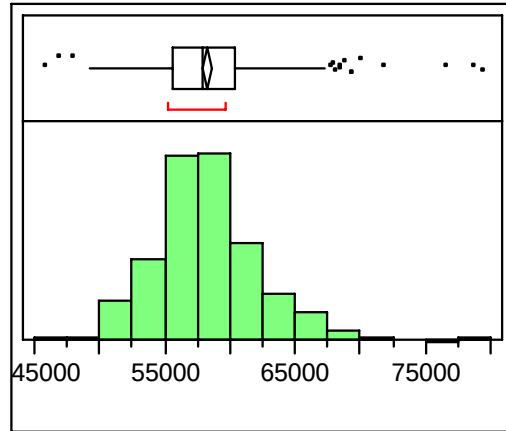


Figure 15. OSB parallel EI histogram and boxplot from JMP.

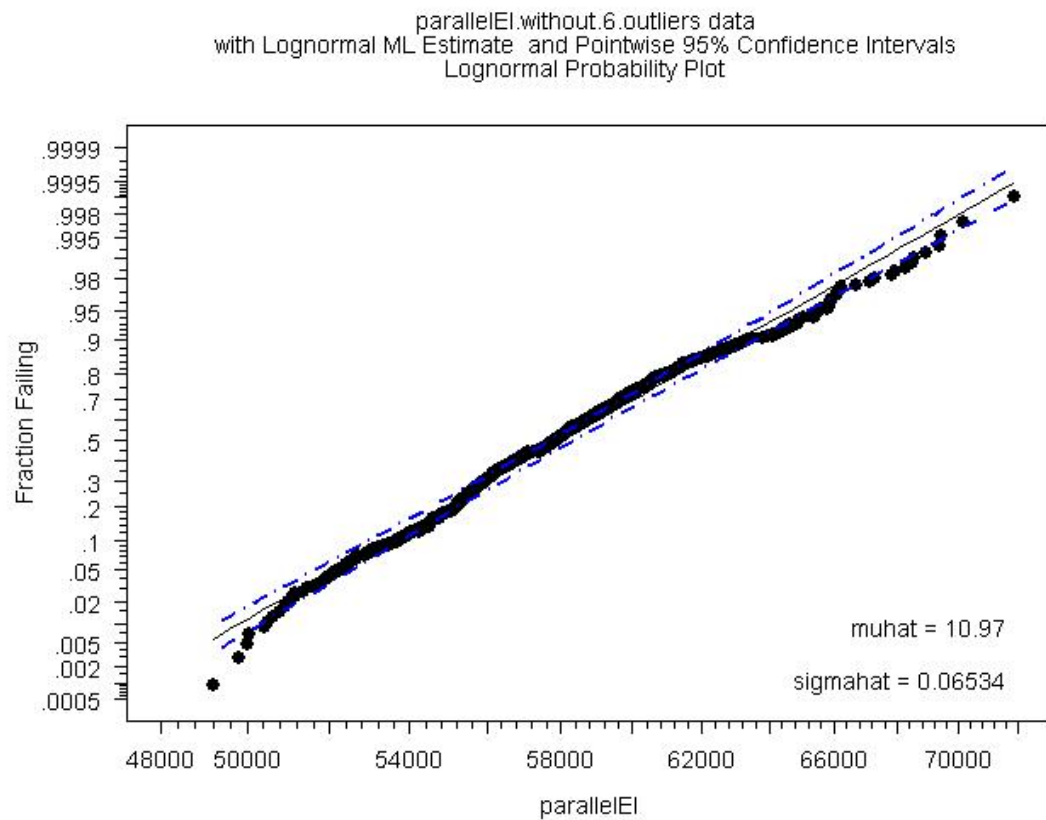


Figure 16. OSB internal bond probability plots from S-PLUS and SPLIDA-Lognormal

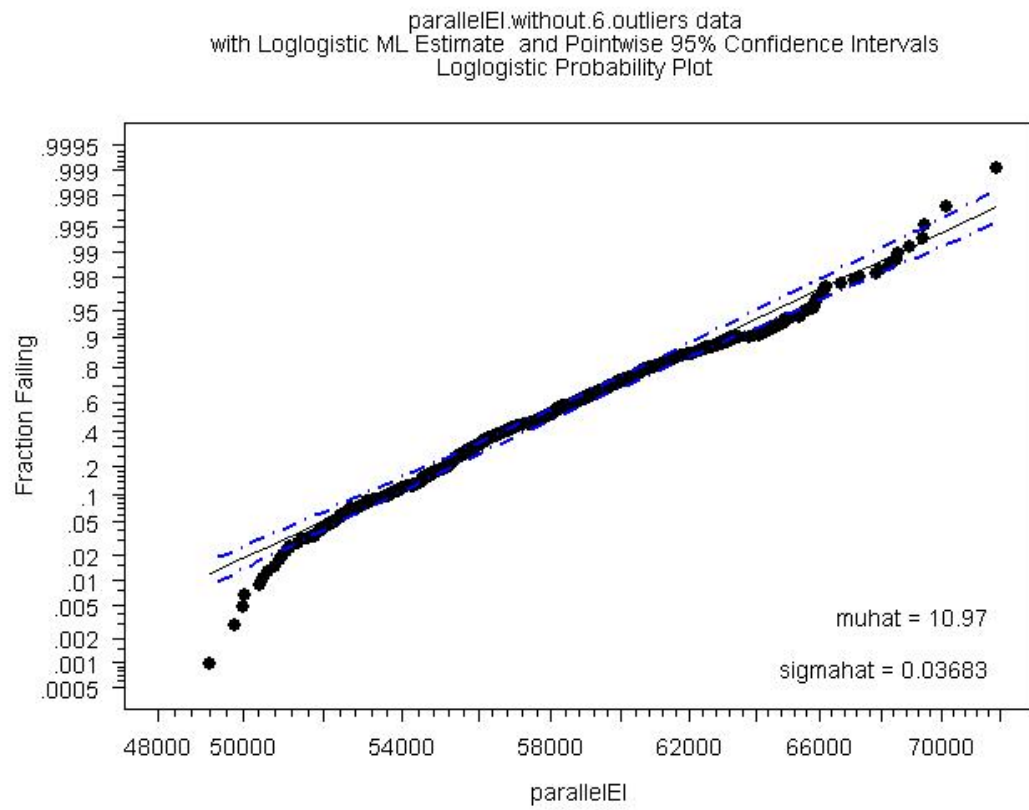


Figure 17. OSB parallel EI probability plots from S-PLUS and SPLIDA-Loglogistic

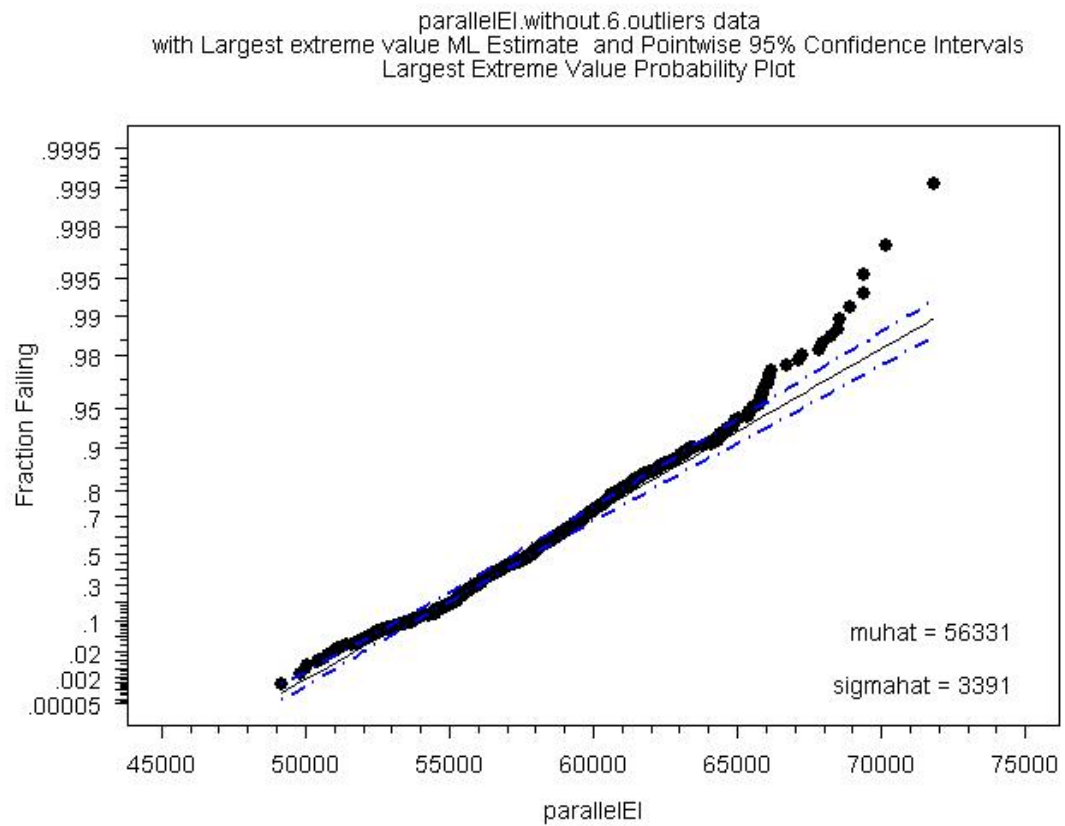


Figure 18. OSB parallel EI probability plots from S-PLUS and SPLIDA-Largest extreme value

These probability plots are consistent with the AIC scores shown in **Table 4** for the data with the six outliers excluded. Note that **Table 4** illustrates the log likelihood and AIC scores of select models. These scores provide quantitative evidence for choosing the best distribution for the data. Before the six outliers are removed, the AIC score of loglogistic model had the lowest score, while the lognormal was third lowest among some of the most common models for reliability data. After the outliers were removed, the lognormal model of stiffness strength had the lowest AIC score, which is consistent with the AIC lognormal model selected for the IB data of the previous section. The overall AIC scores became smaller after the outliers were removed, providing additional evidence that these observations are indeed outliers.

Compare Walker and Guess (2003). The Kaplan-Meier plot is illustrated again in **Figure 19** to capture the reliability function of Parallel EI of OSB. From Kaplan-Meier estimates, 95% of the Parallel EI of OSB can survive at 52219 psi, while 5% of the Parallel EI of OSB can survive at 65435 psi. Half of the Parallel EI of OSB can survive at 57856 psi. This information is helpful for OSB manufacturers and end users. In addition, two different groups of Parallel EI of OSB can be plotted together. By comparing the Kaplan-Meier curves from different groups of OSB, manufacturers may improve product quality, while minimizing raw material inputs and reducing sources of variation.

Table 4. Selected model scores for the parallel EI data complete and excluding outliers.

Model fit	Complete data		Data with 6 outliers excluded	
	Log likelihood	AIC	Log likelihood	AIC
Lognormal	-5151	10306	-5052	10108
Loglogistic	-5140	10284	-5053	10110
LEV	-5158	10320	-5057	10118
Logistic	-5147	10298	-5058	10120
Normal	-5164	10332	-5059	10122
SEV	-5301	10606	-5143	10290
Weibull	-5259	10522	-5159	10322
Frechet	-5173	10350	-5066	10136

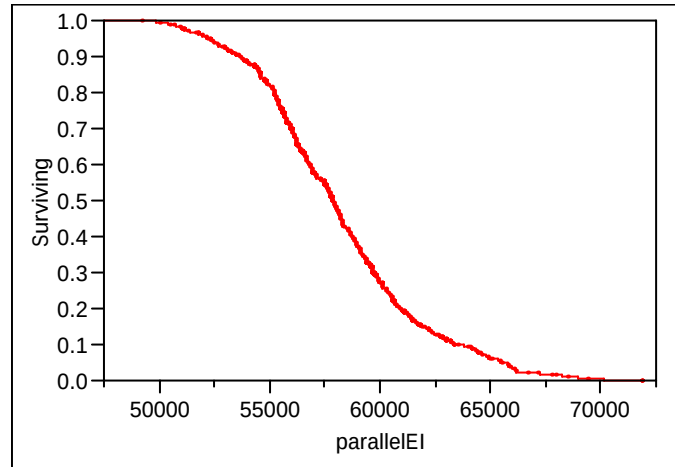


Figure 19. Reliability Kaplan-Meier plot of parallel EI of OSB.

CHAPTER VI. PROTECTION OF CONTAMINATION FROM WEAK UNITS IN ESTIMATION OF SMALL PERCENTILES BY USING FORCED MEDIAN CENSORED DATA

The simulated random data are from two Weibull distributions. The first Weibull distribution has a shape parameter of 27 and a scale parameter of 95. The second Weibull distribution has a shape parameter of 15 and a scale parameter of 135. We refer to the first Weibull distribution as “weak” and the second one as “strong” since the percentiles are much larger in the strong distribution, indicating stronger units over the weak distribution. This is why the scale parameters for the two Weibull distributions are set as 95 and 135 separately, since scale parameters are roughly equal to the value of the median. Their Weibull distribution formulas are

$$F_1(x) = 1 - e^{-(x/95)^{27}} \quad (22)$$

$$F_2(x) = 1 - e^{-(x/135)^{15}} \quad (23)$$

Histograms of the simulated data from the mixed model were used as a powerful visual tool to explore changes of the simulated data with different mixture leverages. (**Figure 20 to 24**)

The parametric 1st, 5th, 10th, 15th percentiles of the weak model and strong model are summarized in **Table 5**. The 1st, 5th, 10th, and 15th parametric percentiles, MLE estimated

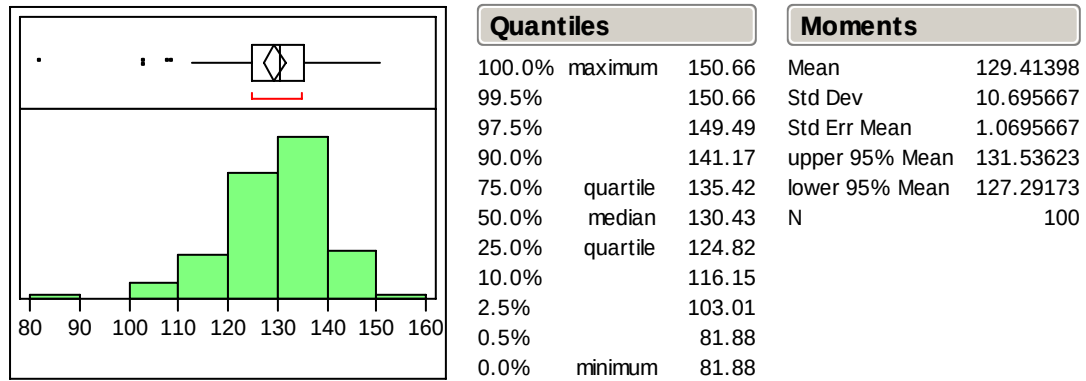


Figure 20. The histogram of simulated data from the mixed model with zero contamination from the weak model.

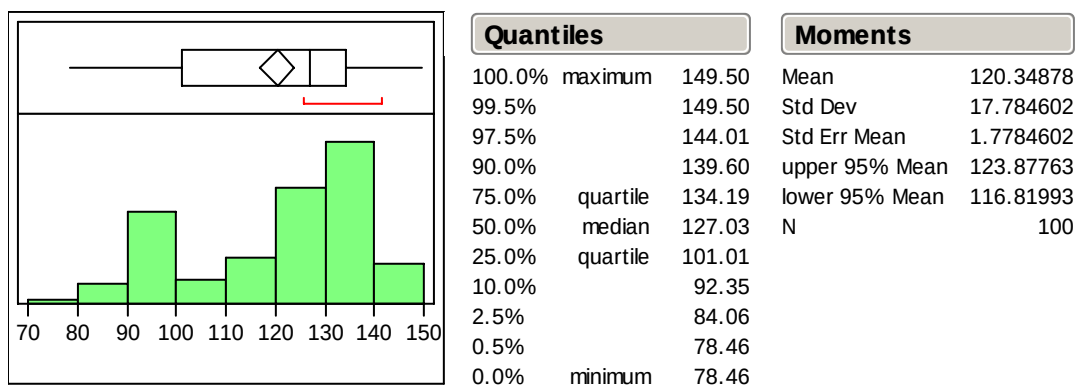


Figure 21. The histogram of simulated data from the mixed model with 25% contamination from the weak model.

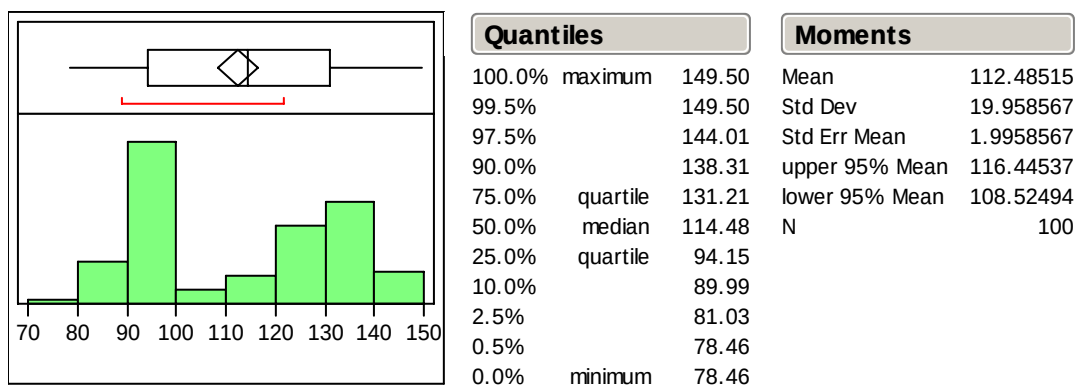


Figure 22. The histogram of simulated data from the mixed model with 50% contamination from the weak model.

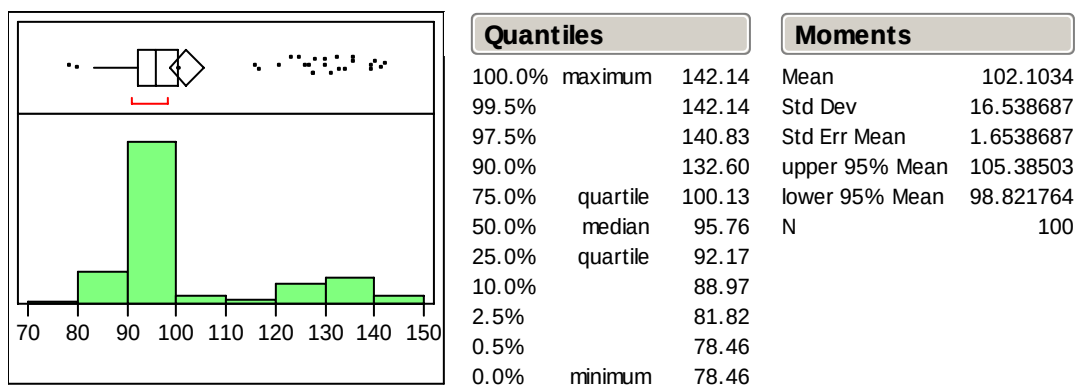


Figure 23. The histogram of simulated data from the mixed model with 75% contamination from the weak model.

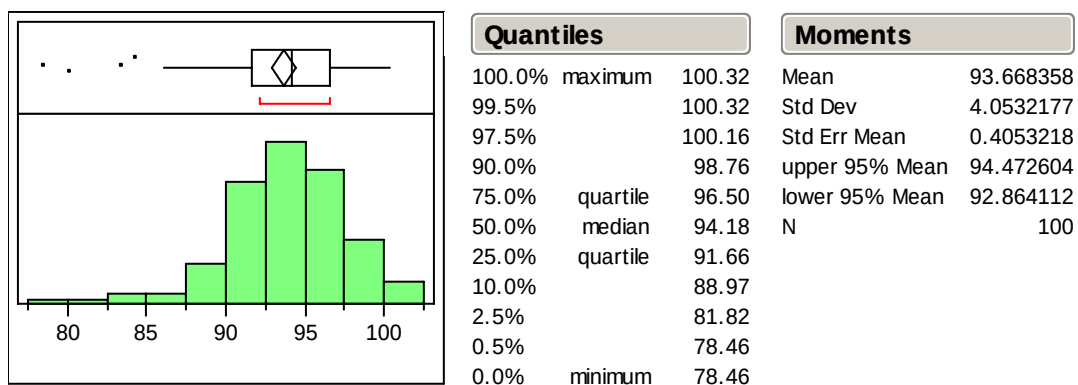


Figure 24 The histogram of simulated data from the mixed model with 100% contamination from the weak model.

Table 5 Comparison of the parametric 1st, 5th, 10th, 15th percentiles of the weak model and the strong model.

Parametric percentile	Weak model	Strong model
1 st percentile	80.1180	99.3450
5 th percentile	85.1036	110.7485
10 th percentile	87.4030	116.1927
15 th percentile	88.8174	119.5989

percentiles on simulated data, and median censored simulated data are summarized in tables with different mixture leverages (**Tables 6 to 9**). The simulated data from the mixed model is produced in this way, when there is zero percentage contamination from the weak model, all the simulated 1000 random numbers are produced by the strong model. When there is 10% contamination from the weak model, this means that 90% of the simulated data randomly comes from the strong model, while 10% of the simulated data is from the weak model. In our case, 1000 random numbers were produced separately for the strong and weak models. Next, 900 numbers out of 1000 numbers were randomly selected by the strong model while 100 numbers out of 1000 numbers were randomly selected from the weak model. These random selections were then combined to form the simulated data for the mixed model with a 10% contamination level from the weak model. The contamination level increases by 10% each time until the total contamination level reaches 100%, meaning that all of the data come from the weak Weibull distribution. We calculate the shape and scale parameters by assuming that the simulated data of the mixed model comes from one Weibull distribution. Then, we obtain the corresponding percentiles. Comparisons are made for the parametric 1st percentile with the 1st percentile from the whole simulated data and from the median censored simulated data. When the contamination level is between 0 to 10% ,both 1st percentiles from the completely simulated data and from the median censored simulated data are very close to parametric 1st percentiles of the mixed model. When the

Table 6. Comparison of 1st percentile from the whole simulated data and median censored simulated data with the parametric 1st percentile having different levels of contamination from the weak Weibull distribution.

Contamination percentage	Parametric 1 st percentile	1 st percentile from the whole simulated data	1 st percentile from the median censored simulated data
0	99.345	100.3636	102.1967
10%	86.959	89.4787	83.1645
20%	84.8505	79.2374	69.8904
30%	83.636	70.3377	63.118
40%	82.7831	63.0295	63.7547
50%	82.1267	57.2816	80.2634
60%	81.5938	52.9948	82.2902
70%	81.1458	50.1005	82.2915
80%	80.7595	48.8572	82.1498
90%	80.4203	50.2135	81.9573
100%	80.118	80.9368	81.7552

Table 7. Comparison of 5th percentile from the whole simulated data and median censored simulated data with the parametric 5th percentile having different levels of contamination from the weak Weibull distribution.

Contamination percentage	Parametric 5 th percentile	5 th percentile from the whole simulated data	5 th percentile from the median censored simulated data
0	110.7485	111.8651	110.2903
10%	93.3578	103.3043	95.3427
20%	90.5575	94.8389	83.7622
30%	89.0957	87.1038	76.7788
40%	88.1033	80.4072	75.429
50%	87.3529	74.7759	85.5027
60%	86.7503	70.2952	86.4465
70%	86.2473	66.9403	86.1863
80%	85.816	64.8987	85.8741
90%	85.4387	64.9382	85.5639
100%	85.1036	85.9616	85.2887

Table 8. Comparison of 10th percentile from the whole simulated data and median censored simulated data with the parametric 10th percentile having different levels of contamination from the weak Weibull distribution

Contamination percentage	Parametric 10 th percentile	10 th percentile from the whole simulated data	10 th percentile from the median censored simulated data
0	116.1927	117.3599	114.0746
10%	98.3608	110.0425	101.2434
20%	93.5548	102.6428	90.6864
30%	91.7857	95.7116	83.7198
40%	90.6561	89.5131	81.2544
50%	89.8249	84.1419	87.9246
60%	89.1675	79.6542	88.3528
70%	88.6243	76.0698	87.9584
80%	88.1617	73.5659	87.5759
90%	87.7592	72.7336	87.2229
100%	87.403	88.2794	86.8979

Table 9. Comparison 15th percentile from the whole simulated data and median censored simulated data with the parametric 15th percentile with different levels of contamination from units from weak Weibull distribution

Contamination percentage	Parametric 15 th percentile	15 th percentile from the whole simulated data	15 th percentile from the median censored simulated data
0	119.5989	120.797	116.4052
10%	111.5507	114.3329	104.9957
20%	95.9163	107.6791	95.167
30%	93.6225	101.3191	88.2016
40%	92.3217	95.5011	84.9714
50%	91.4019	90.3036	89.4101
60%	90.6893	85.8553	89.514
70%	90.1076	82.1579	89.0468
80%	89.6164	79.336	88.6163
90%	89.1916	77.8839	88.2228
100%	88.8174	89.7103	87.8852

contamination level is from 20% to 50%, both the 1st percentile from the completely simulated data and from the median censored simulated data are a little smaller than the parametric 1st percentile of the mixed model. Between 50% and 100%, the percentile from the median censored simulated data are very close to the parametric 1st percentile, while the 1st percentile from the whole simulated data is much smaller than the parametric 1st percentile. When the contamination level reaches 100%, the mixed model data is assumed to come from one Weibull distribution only, with the three percentiles approaching similar values. The same pattern is observed on the 5th, 10th 15th percentiles.

Figures 25 and 26 are the overlays of the parametric percentiles of the mixed model, the parametric percentiles of the weak model, percentiles from the whole simulated data and from the median censored simulated data. By comparing these figures, we find that the median censored method performs better for estimating the higher percentiles. For example, when the contamination level is from 10% to 50%, the median censored method does not perform as well as the whole simulated data for estimating the 1st percentile. However, this case is only true when the contamination level is from 20% to 50% for estimating the 15th percentile. When we compare the three lines of percentile from the whole simulated data and median censored simulated data with the parametric percentile

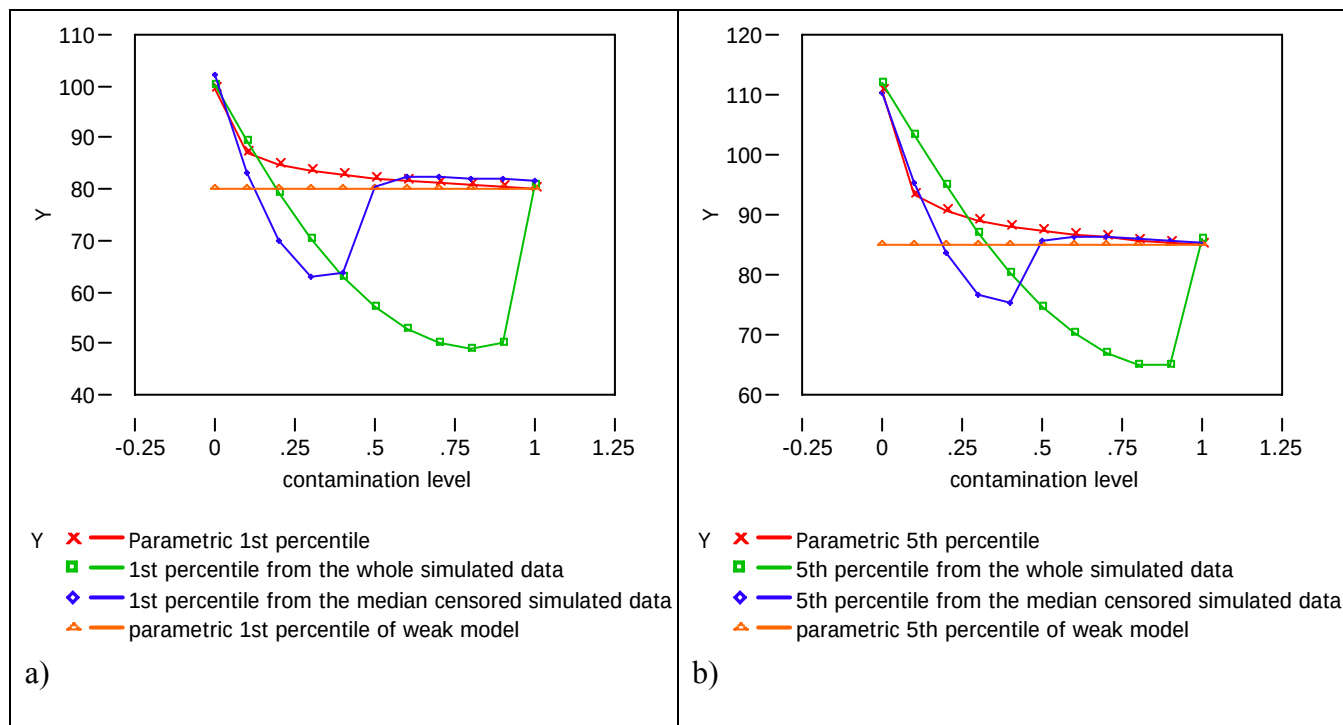


Figure 25. a) Comparison of 1st percentile from the whole simulated data and median censored simulated data with the parametric 1st percentile having different levels of contamination from the weak Weibull distribution b) Comparison of 5th percentile from the whole simulated data and median censored simulated data with the parametric 5th percentile having different levels of contamination from the weak Weibull distribution

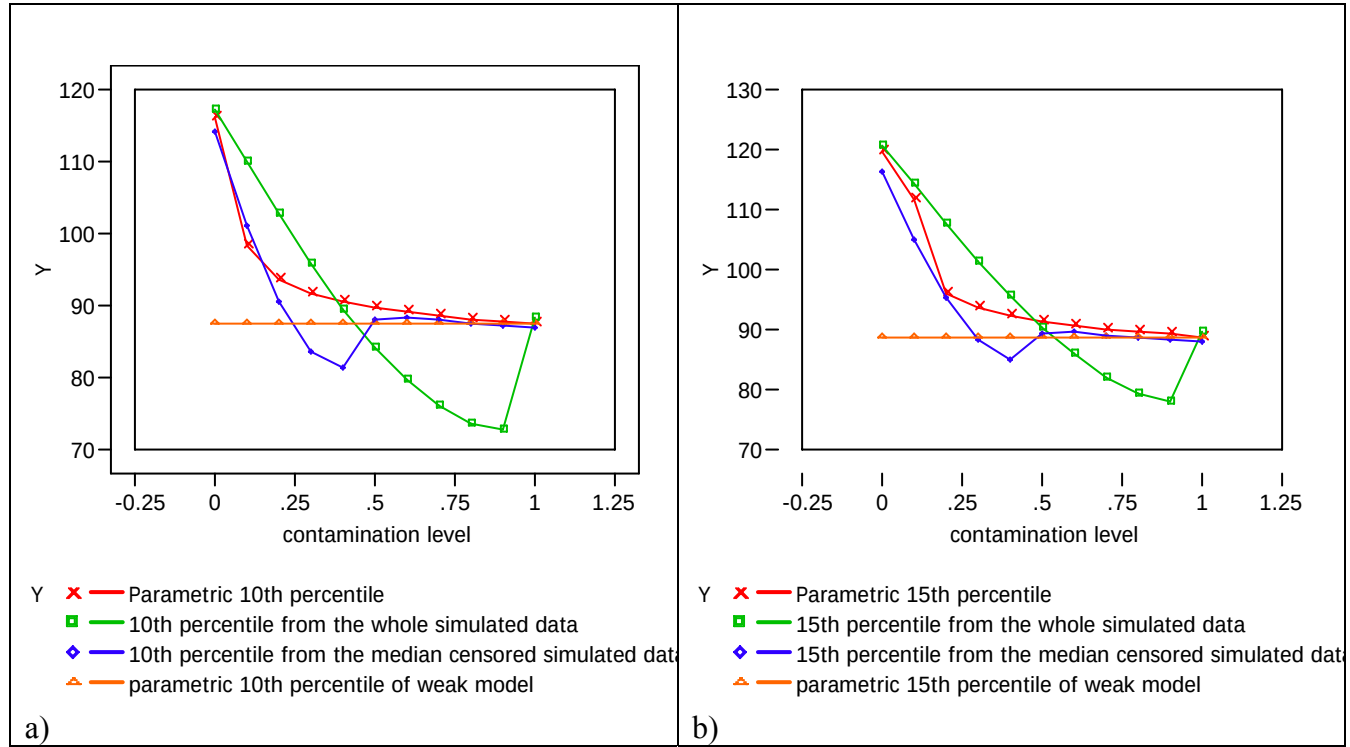


Figure 26. a) Comparison of 10th percentile from the whole simulated data and median censored simulated data with the parametric 10th percentile having different levels of contamination from the weak Weibull distribution b) Comparison of 15th percentile from the whole simulated data and median censored simulated data with the parametric 15th percentile having different levels of contamination from the weak Weibull distribution

of the mixed model, we can see the line of the median censored data is much closer to the parametric percentile of the mixed model than for the whole simulated data.

We study the forced median censored method in estimating small percentiles from simulated data of the mixture of two Weibull distributions. These simulations show the median censored method estimates lower percentiles very close to the true parametric percentiles. The median censored method provides strong improvements and protection in many places, while in other cases caution is needed. The simulations establish the median censoring method as a useful tool for improving estimation of the lower percentiles by the method fitting the lower distributional side better.

CHAPTER VII. CONCLUSIONS

Because of the wide application of engineered wood products in building and building-related products, it is important to focus on product quality. As previously discussed, significant research is devoted to improving the product quality of engineered wood. Statistical reliability and quality control methods are important tools for monitoring the quality of forest products and specifically engineered wood.

In the thesis data set of 529 OSB panels, the average value of IB is 49.7 p.s.i. with standard deviation equaling 11.3 p.s.i. One outlier in the IB data is identified from a boxplot that had an extremely small IB of 15.3 p.s.i. This outlier might be the result of a simple typo during data entry or may indicate a more serious problem, such as an infant mortality mode of failing strength of manufactured OSB, etc. According to probability plots and information criteria, the lognormal distribution is the best fit for IB when compared to the other frequently used reliability distributions, including the SEV, normal distribution, Weibull distribution, loglogistic distribution etc.

The mean value of EI for the same data set of 529 OSB panels is 58212 p.s.i with a standard deviation of 4205.7 p.s.i. Six potential outliers of EI data are identified from a boxplot with three outliers of extremely small values and the other three of extremely

large values. This may indicate that further investigation is required to identify the cause of the six outliers. One unusual feature of EI data is its kurtosis of 2.69 which indicates a very “peaked” distribution. The lognormal distribution fits EI data best when compared to other most frequently used probability distributions for reliability data.

A nonparametric plot such as the Kaplan-Meier plot may be used to capture the reliability function of both IB and EI data. By comparing the Kaplan-Meier curves from different groups of OSB, manufacturers may improve product quality, while minimizing raw material inputs and reducing sources of variation.

Sometimes, strong units of OSB may be contaminated with weak OSB units, e.g., undiscovered introduction of “fines” or suboptimal small flakes in the process. Both strong units and weak units may come from a Weibull distribution. In order to find a better way to estimate small percentiles more accurately, a novel method called median censoring is used on simulated data sets. Simulated data of a mixed model with two differing two-parameter Weibull distributions are produced using MATLAB code. The Weibull model with a shape parameter equal to 27 and a scale parameter of 95 is referred to as the weak model. The Weibull model with a shape parameter of 15 and a scale parameter of 135 is referred to as the strong model. Histograms of the simulated data from the mixed model are used as powerful visualization tools to explore changes of the

simulated data with different mixture leverages. MLE estimated percentiles are calculated from the simulated data. Parametric percentiles of the mixed model are calculated from the cumulative probability function equation of the mixed model by the Newton-Raphson method. The 1st, 5th, 10th, and 15th parametric percentiles, the MLE estimated percentiles on simulated data and median censored simulated data of different mixture leverages are summarized. It is concluded from the comparison of these percentiles, that when the mixture leverage is equal to or bigger than 50%, MLE estimated percentiles on median censored simulated data are very close to the parametric percentiles of the mixed model which is in contrast to the MLE estimated percentiles on the completely simulated data. When the mixture leverage is between 0 and 50%, the MLE estimated percentiles on simulated data with median censoring do not lose much information compared with MLE estimated percentiles on simulated data without median censoring.

In conclusion, we find that graphically and statistically exploring OSB reliability, as measured by IB and Parallel EI, provides valuable information about OSB. The median censored method is a good tool to estimate lower percentiles. This method estimates lower percentiles more accurately than the non- median censored method in the simulation case. More research could be done to explore the application and effective estimation of median censored methods for the Weibull distribution with different scale

and shape parameters. Besides the mixed model with two Weibull distributions, other mixed models with different distributions could be examined, such as the mixed model with two lognormal distributions. Mixed models of two different distributions should be considered for future research, i.e., a mixed model with one lognormal distribution and one Weibull distribution. In addition to median censored methods, other censored methods can be applied, such as a quartile censored method.

REFERENCES

LIST OF REFERENCES

- Adair, C. (2005). Market outlook for the OSB and structural panel market. APA – Engineered Wood Association. Tacoma, WA.
- 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, Acaderniai Kiado, Budapest, 267-281
- American Society of Testing and Materials. (2003). Standard test method for tensile properties of plastics. ASTM D 638-03. ASTM. West Conshohocken, PA.
- Bozdogan, H. (2000). Akaike's information criterion and recent developments in information complexity. *Journal of Mathematical Psychology*, 44(1), 62-91.
- Bozdogan, H. (2004). Lecture notes for Statistics 564, the University of Tennessee, Knoxville, TN.
- Chen, W. 2005. A reliability case on estimating extremely small percentiles of strength data for continuous improvement of medium density fiberboard product quality. M.S. Thesis. University of Tennessee. Knoxville, TN.
- Chen, W., R.V. León, T.M. Young, and F.M. Guess, F.M. 2006. Applying a forced censoring technique with accelerated modeling for improving estimation of extremely small percentiles of strengths. *International Journal of Reliability and Application*. 7(1):27-39.
- Deming, W.E. (1986). Out of the Crisis. Massachusetts Institute of Technology's Center for Advanced Engineering Design, Cambridge, MA.
- Deming, W.E. (1993). The New Economics. Massachusetts Institute of Technology's Center for Advanced Engineering Design, Cambridge, MA
- English, L. P. (1999), Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits. New York: John Wiley & Sons.

- Edwards, D.J. 2004. An applied statistical reliability analysis of the internal bond of medium density fiberboard. M.S. Thesis. University of Tennessee. Knoxville, TN.
- 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.M., 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.M., León, R., Chen, W. and Young, T.M. (2004). Forcing a closer fit in the lower tails of a distribution for better estimating extremely small percentiles of strengths. *International Journal of Reliability and Application*, **5**(4), 129-143.
- Guess, F.M., and Proschan, F. (1988). Mean residual life: theory and applications. *Handbook of Statistics: Quality Control and Reliability*, **7**, 215-224.
- Guess, F.M., Walker, E., and Gallant, D. (1992). Burn-in to improve which measure of reliability. *Microelectronics and Reliability*, **32**, 759-762.
- Guess, F.M., Zhang, X., Young, T.M. and León, R. (2005). Using mean residual life functions for unique insights into strengths of materials data. *International Journal of Reliability and Application*. **6**(2), 79 – 85.
- Huang, K-T., Lee, Y. L. and Wang, R. Y. (1999). *Quality Information and Knowledge*. Prentice Hall, New York, NY.
- Jiang, S. (1992). Maximum likelihood estimates, from censored data, for mixed-Weibull distributions. *IEEE Transactions on reliability*. Vol 41, NO. 2, 1992.
- Kececioglu, D. B. (1998). Parameter estimation for mixed-Weibull distribution. *Proceedings annual reliability and maintainability symposium*.
- Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. Wiley & Sons, New York, NY.
- Neimsuwan, T. 2004. Effect of resin and wax ratio on OSB properties. M. S. Thesis. University of Tennessee. Knoxville, TN.

- Perhac, D.G., Young, T.M., Guess, F.M., and Leon, R. V. (2007). Exploring reliability of wood-plastic composites: stiffness and flexural strengths. Submitted.
- Redman, T. C. (1996). Data Quality for the Information Age. Artech House Computer Science Library, Boston/Norwood, MA.
- Redman, T. C. (2001). Data Quality: The Field Guide. Butterworth-Heinemann Digital Press, Boston, MA.
- Sharp, J.G. 2004. Formaldehyde – the big issue. Proc. 8th European Panel Products Symposium. Llandudno, UK. p. 2-1 to 2-25.
- Spake, A. 2007. Toxic trailers. Sun Herald Feb. 22, 2007 issue. Biloxi, MS.
- Steele, J.C. 2006. “Function domain sets” confidence intervals for the mean residual life Functions with applications in production of medium density fiberboard. M.S. Thesis. University of Tennessee. Knoxville, TN.
- Walker, E. and Guess, F.M. (2003). Comparing reliabilities of the strength of two container designs: A case study. Journal of Data Science, **1**, 185-197.
- Weibull, W. (1951). A statistical distribution function of wide applicability, Journal of Applied Mechanics, **18** (1), 293-297
- Winandy, J. E. (2004). Fundamentals of composite processing. Proceedings of a workshop; November 5-6, 2003.
- Young, T.M. and Guess, F.M. (1994). Reliability processes and structures. Microelectronics and Reliability **34**, 1107-1119.
- 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 Application **3**(4), 155-164.
- Youngquist, J. A. (1987). Resin distribution in hardboard: evaluated by internal bond strength and fluorescence microscopy. Wood and Fiber Science. 19(2), 215-224.

APPENDIX

APPENDIX

Matlab code:

```
% produces the simulated data of mixed model with two Weibull distribution

% alpha is the mixed percentage from two different Weibull models;

% n equals how many random numbers are generated from two Weibull and uniform
%distribution

% b is shape parameter c is scale parameter;

%W1=alpha percentage of the model;

%W2=(1-alpha) percentage of the mixed model;

function data=mixed_data1(b1,c1,b2,c2,alpha,n)

    W1=wblrnd(b1,c1,n,1);

    W2=wblrnd(b2,c2,n,1);

    U=random('uniform',0,1,n,1);

    data=[];

    for i=1:1:n

        if U(i)>1-alpha

            data(i)=W1(i);

        else

            data(i)=W2(i);

        end

    end
```

```

end

end

data=data';

%Newton-Raphson Algorithm for Weibull MLE's

function first_esti=NRWeibull_revise(x)

n=length(x);

chatinit=100;

bhatinit=10;

chat(1)=chatinit;

bhat(1)=bhatinit;

ck=1;

bk=1;

end_b=0;

while end_b<1

    bk=bk+1;

    end_c=0;

    while end_c<1

        ck=ck+1;

        S(ck-1)=(n/chat(ck-1))-n*log(bhat(bk-1))+sum(log(x))-sum(((x./bhat(bk-1)).^chat(ck-1)).*log(x./bhat(bk-1))));

```

```

H(ck-1)=n/chat(ck-1)^2+sum(((x./bhat(bk-1)).^chat(ck-1)).*(log(x./bhat(bk-1))).^2);

chat(ck)=chat(ck-1)+(S(ck-1)/H(ck-1));

if abs(chat(ck)-chat(ck-1))<0.001

    end_c=1;

end

end

bhat(bk)=((1/n)*sum(x.^chat(ck)))^(1/chat(ck));

if abs(bhat(bk)-bhat(bk-1))<0.001

    end_b=1;

end

end

chat=chat(ck);

bhat=bhat(bk);

% Calculate Weibull Quantiles

p=0.01;

first_esti=bhat*log(1/(1-p))^(1/chat);

% produces the median censored data;

function data_cenc=median_data(x)

```

```

n=length(x);

x_median=median(x);

diff=x-x_median;

for i=1:1:n

    if diff(i)>0

        diff(i)=0;

    else

        diff(i)=1;

    end

end

combo=[diff,x];

nozeroindex=find(combo(:,1));

data_cenc=combo(nozeroindex,2);

% summarize the corresponding percentiles of the simulated and censored simulated data

function [uncencor_t, uncencor_e, cencor_t, cencor_e]=thesis(b1,c1,b2,c2,alpha,n)

    for i=1:1:100

        data=mixed_data1(b1,c1,b2,c2,alpha,n);

        uncencor_t(i)=prctile(data,0.01);

```

```

uncencor_e(i)=NRWeibull_revise(data);

cencor_data=median_data(data);

cencor_t(i)=prctile(cencor_data,0.01);

cencor_e(i)=NRWeibull_revise(cencor_data);

end

uncencor_t=mean(uncencor_t);

uncencor_e=mean(uncencor_e);

cencor_t=mean(cencor_t);

cencor_e=mean(cencor_e);

%calculate the parametric 1st percentile
t=0.99;
x=[];
sampletest(t);
x=x'

% the below is sample test function
function y = sample(t)
a = 0:0.1:1;
x=[];
for i = 1:length(a)
    x(i) = fzero(@(x)a(i)*exp(-(x/95)^27)+(1-a(i))*exp(-(x/135)^15)-t,-1)
end
x=x'

```

VITA

Yang Wang, originally from Wuhan, China, earned a MD degree from Tongji Medical College, Huazhong University of Science and Technology in May 2001. She completed a Master of Science degree in the Department of Microbiology at the University of Tennessee in August 2005. Since then, she has been pursuing a second Master's degree in Statistics at UT. She has served as a Graduate Teaching Assistant, Graduate Research Assistant, and a student member of American Statistics Association (ASA).