Journal of Testing and Evaluation

Steve P. Verrill,[1] Frank C. Owens,[2] David E. Kretschmann,[3] Rubin Shmulsky,[2] and Linda S. Brown[4]

DOI: 10.1520/JTE20180472

Visual and MSR Grades of Lumber Are Not 2-parameter Weibulls and Why This May Matter

Journal of Testing and Evaluation

doi:10.1520/JTE20180472 available online at www.astm.org

Steve P. Verrill,[1] Frank C. Owens,[2] David E. Kretschmann,[3] Rubin Shmulsky,[2] and Linda S. Brown[4]

Visual and MSR Grades of Lumber Are Not 2-parameter Weibulls and Why This May Matter

Reference

S. P. Verrill, F. C. Owens, D. E. Kretschmann, R. Shmulsky, and L. S. Brown, "Visual and MSR Grades of Lumber Are Not 2-parameter Weibulls and Why This May Matter," *Journal of Testing and Evaluation* https://doi.org/10.1520/JTE20180472

Manuscript received July 9, 2018; accepted for publication January 22, 2019; published online April 24, 2019.

[1] United States Department of Agriculture Forest Service Forest Products Laboratory, 1 Gifford Pinchot Dr., Madison, WI 53726, USA (Corresponding author) e-mail: sverrill@fs.fed.us, http://orcid.org/0000-0001-7627-9712

[2] Department of Sustainable Bioproducts, Box 9820, Mississippi State, Starkville, MS 39762, USA

[3] American Lumber Standard Committee, 7470 New Technology Way, Suite F, Frederick, MD 21703, USA

[4] Southern Pine Inspection Bureau, PO Box 10915, Pensacola, FL 32524, USA

ABSTRACT

It has been common practice to assume that a 2-parameter Weibull probability distribution is suitable for modeling lumber strength properties. Previous work has demonstrated theoretically and empirically that the modulus of rupture (MOR) distribution of a visual grade of lumber or of lumber that has been binned by modulus of elasticity (MOE) is not a 2-parameter Weibull. Instead, the tails of the MOR distribution are thinned via pseudo-truncation. Simulations have established that fitting 2-parameter Weibulls to pseudo-truncated data via either full or censored data methods can yield poor estimates of probabilities of failure. In this article, we support the simulation results by analyzing large In-Grade type data sets and establishing that 2-parameter Weibull fits yield inflated estimates of the probability of lumber failure when specimens are subjected to loads near allowable properties. In this article, we also discuss the censored data or tail fitting methods permitted under ASTM D5457, *Standard Specification for Computing Reference Resistance of Wood-Based Materials and Structural Connections for Load and Resistance Factor Design*.

Keywords

2-parameter Weibull distribution, pseudo-truncated Weibull distribution, machine stress rated lumber, modulus of elasticity binned lumber, visually graded lumber, thinned tail, lumber property distribution, lumber reliability, censored data 2-parameter Weibull fits

Introduction

Verrill et al.[1–4] established empirically and theoretically that visual and machine stress rated (MSR) grades of lumber are not distributed as 2-parameter Weibulls. Instead, they have (at least to a first approximation) pseudo-truncated distributions. They also performed simulations that strongly suggested that both censored (see Section X2 of ASTM D5457-19, *Standard Specification for Computing Reference Resistance of Wood-Based Materials and Structural Connections for Load and Resistance Factor Design*[5]) and uncensored fits of 2-parameter Weibulls to pseudo-truncated data can lead to significant over- or underestimation of probabilities of failure when loads are near allowable properties. (Censored data techniques are also known among wood scientists as tail fitting and involve situations in which we have [or use] full information for only a subset of the data. Censored data techniques are discussed in many statistical textbooks that deal with reliability or lifetime estimation methods. See, for example, Lawless.[6])

Pseudo-truncation has a technical meaning. The concept, at least, of pseudo-truncation was recognized in an American Society of Civil Engineers pre-standard report.[7] Section B3 of that standard notes that "an improved strength distribution can be obtained by … thinning the lower tail by sorting on a correlated variable." For example, if the full "mill run" bivariate modulus of elasticity–modulus of rupture (MOE–MOR) distribution were a bivariate Gaussian (normal)–Weibull, then truncating or binning on the basis of MOE values (as in MSR lumber) would lead to a pseudo-truncated MOR distribution. That is, because MOE and MOR are not perfectly correlated, truncating on the basis of lower and upper MOE limits does not lead to perfect truncation of the MOR distribution, but it does, of course, lead to a MOR distribution whose tails are thinned. For the case in which the mill run joint MOE–MOR distribution is a bivariate Gaussian–Weibull, Verrill et al.[1,4] derived the exact form of this pseudo-truncated Weibull distribution. (They obtained its probability density function.) They also showed that it cannot have tail behavior that matches that of a Weibull distribution.

In this article, we analyze "In-Grade type" data sets to establish that modeling the MOR distributions of visual grades of lumber by 2-parameter Weibull distributions can lead to poor reliability estimates. In particular, when loads are close to the allowable properties calculated for those data sets, estimates of probabilities of failure will tend to be inflated. Our preceding simulation work—see Section 5.4 of Verrill et al.[2]—demonstrated that this positive bias in the mean tends to decrease with censoring, but also, as we would expect, censored estimates are more variable than correct pseudo-truncated estimates. Consequently, censored data techniques can, with appreciable probability, yield failure probability estimates that are considerably too high and, again with appreciable probability, yield estimates that are considerably too low.

The Data

The data come from 19 of the original In-Grade data cells (species-size-grade-property combinations), 6 data cells from a 2011 Southern Pine Inspection Bureau (SPIB) repeat of the In-Grade testing program, and 1 data cell from a 2014 SPIB resource monitoring program study. The data cells are identified in columns 1–4 of **Table 1**. The In-Grade program and some of its results are discussed in Green and Evans,[8] Green, Shelley, and Vokey,[9] and Evans and Green[10] Testing procedures for the In-Grade testing program are described in ASTM D4761, *Standard Test Methods for Mechanical Properties of Lumber and Wood-Based Structural Materials*,[11] and the data were adjusted in accordance with ASTM D1990, *Standard Practice for Establishing Allowable Properties for Visually-Graded Dimension Lumber from In-Grade Tests of Full-Size Specimens.*[12] The original SPIB resource monitoring program (1994–2010) is discussed in Kretschmann, Evans, and Brown.[13] In recent years, the program has been modified to ensure conformity with the requirements in the most recent version of ASTM D1990 and to add action points that depend upon both strength and stiffness measurements.

An Extension of An Earlier Analysis

Verrill et al.[3] presented a table that provided Cramér–von Mises (CVM) and Anderson–Darling (AD) goodness-of-fit test p values for tests of the null hypotheses that 19 In-Grade data cell MOR distributions were 2-parameter Weibulls. In **Table 1** of the current article, we extend this 2014 table to include the 2011 and 2014 SPIB data. To

TABLE 1

p values for Cramér–Von Mises and Anderson–Darling goodness-of-fit tests of 2-parameter Weibull fits to In-Grade, 2011 SPIB In-Grade, and 2014 SPIB resource monitoring program data

					Goodness-of-fit *p* value	
Data Set	Species	Lumber Size	Grade	Sample Size	CVM	AD
In-Grade	DF	2 × 4	SS	414	.086	.033
	DF	2 × 8	SS	493	.472	.418
	DF	2 × 10	SS	414	.955	.771
	DF	2 × 4	2	386	.116	.066
	DF	2 × 8	2	1,964	.001	.001
	DF	2 × 10	2	388	.001	.001
	HF	2 × 4	SS	428	.026	.002
	HF	2 × 8	SS	375	.034	.033
	HF	2 × 10	SS	368	.062	.049
	HF	2 × 4	2	406	.004	.002
	HF	2 × 8	2	372	.009	.004
	HF	2 × 10	2	361	.010	.002
	SP	2 × 4	SS	413	.029	.010
	SP	2 × 8	SS	626	.028	.023
	SP	2 × 10	SS	413	.002	.001
	SP	2 × 4	2	413	.001	.001
	SP	2 × 6	2	413	.001	.001
	SP	2 × 8	2	1,367	.001	.001
	SP	2 × 10	2	412	.086	.113
2011	SP	2 × 4	SS	420	.254	.158
	SP	2 × 8	SS	409	.316	.226
	SP	2 × 10	SS	410	.043	.023
	SP	2 × 4	2	408	.001	.001
	SP	2 × 8	2	420	.196	.146
	SP	2 × 10	2	420	.091	.060
2014	SP	2 × 4	2	362	.807	.590

Note: *p* values listed as 0.001 might actually be lower.

perform the goodness-of-fit tests, we used an R[14] goodness-of-fit function, WEDF.test (Krit[15]), that provides more precise estimates of the *p* values than the estimates provided in Verrill et al.[3] This updated table continues to strongly suggest that visual grades of lumber are poorly fit by 2-parameter Weibulls.

Verrill et al.[3] also discussed Weibull probability plots of the data. In **figure 1**, we provide an example of such a plot. Verrill et al.[3] noted that 16 of the 19 data sets available at that time led to probability plots that had the short or thinned left tails that one would expect from pseudo-truncated data. (That is, the points in the left tails of the probability plots tended to lie above $y = x$ lines.) Twenty of the 26 data sets currently available to us display such a short left tail. To give readers an idea of how unlikely this would be if the data truly were 2-parameter Weibull, we performed 26 simulations. To do this, for each of the 26 data sets, we first obtained the maximum likelihood fit of a 2-parameter Weibull to the data. If the original data set contained n points, we then generated a data set of size n from the fitted 2-parameter Weibull distribution and plotted the corresponding Weibull probability plot. An example of such a plot is provided in **figure 2**. (Actual and generated probability plots for all 26 data sets can be viewed at http://www1.fpl.fs.fed.us/weib2.pp.html.) None of the 26 generated probability plots displayed shortened left tails. If the real MOR distributions are 2-parameter Weibulls, an approximate estimate of the probability of seeing all 20 of the observed short left tails among the original 26 probability plots and none among the probability plots associated with generated data or vice versa is

FIG. 1

Weibull probability plot for the 2011 SPIB In-Grade Southern Pine, 2 × 4, No. 2 data. Ordered empirical data versus expected ordered data under a 2-parameter Weibull model. The straight line is the $y = x$ line.

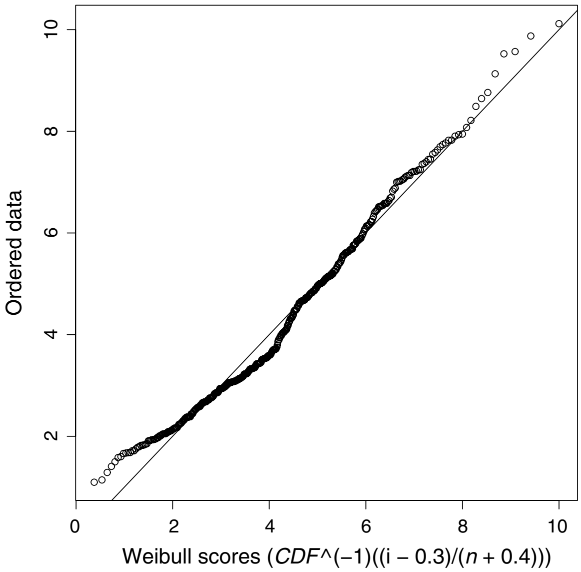
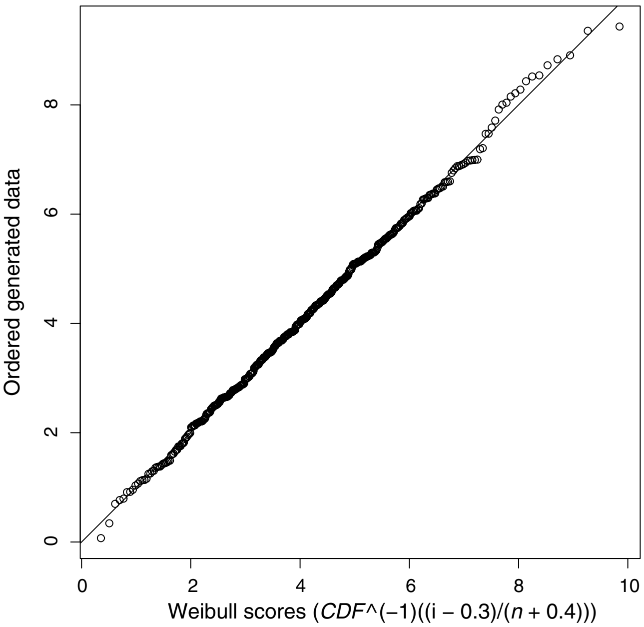


FIG. 2

Weibull probability plot of generated data. Ordered generated data versus expected ordered data under a 2-parameter Weibull model. The generated data was generated from a 2-parameter Weibull maximum likelihood fit of the 2011 SPIB In-Grade Southern Pine, 2 × 4, No. 2 data. The straight line is the $y = x$ line.



$$2 \times \genfrac{}{}{0pt}{}{26}{20} \qquad \genfrac{}{}{0pt}{}{52}{20} = 4 \times 10^{-9}$$

Note that this estimate ignores differences that may be due to species, lumber size, grade, and sample size differences, so it is merely suggestive rather than definitive. In fact, the data and intuition suggest that select structural (SS) and No. 2 fits may behave differently—13 of 14 No. 2 probability plots display overly heavy predicted (thin observed) left tails, whereas only 7 of 12 SS plots do. This might be associated with the fact that SS grades, with potential physical right tail limits, are all left-skewed, whereas No. 2 grades that might contain No. 1 and SS specimens are all right-skewed. (See the full, actual data, skewness (Skew) estimates in **Table 2**. The complete **Table 2** can be found in Verrill et al.[16] In this article, we only include that portion of **Table 2** that covers Douglas Fir specimens.)

Regardless, it is unlikely that the 26 MOR distributions are 2-parameter Weibulls. Of course, we essentially already knew this from the Cramér–von Mises and Anderson–Darling goodness-of-fit tests. However, this second analysis focuses our attention on the thinned observed left tails as one source of the lack-of-fit.

So far, we have been seeking to conclude that 2-parameter Weibull fits may be statistically rejected. In the next section, we identify a practical reason for rejecting 2-parameter Weibull lumber strength models when making reliability predictions.

The New Analysis

We used full and censored data maximum likelihood methods to fit 2-parameter Weibull distributions to each of the 26 data sets. A listing of the program that did the fitting (fit26.w2.cens.4.web.f) can be found at

TABLE 2

In-Grade Douglas Fir data

			Actual Data				Generated Data			
Data Set	Data Cell	Fraction of Data	Shape	Scale	Skew	Exc. Kurt	Shape	Scale	Skew	Exc. Kurt
In-Grade	DF2 × 4SS	full	4.62	11.07	−0.197	−0.176	4.65	11.15	−0.202	−0.172
		20 %	5.56	10.56	−0.326	−0.033	4.96	10.74	−0.248	−0.126
		10 %	6.07	10.10	−0.380	0.046	6.08	9.64	−0.382	0.048
		5 %	8.58	8.46	−0.568	0.403	6.33	9.42	−0.406	0.087
	DF2 × 8SS	full	3.75	8.45	−0.034	−0.274	3.88	8.53	−0.062	−0.264
		20 %	3.71	8.57	−0.024	−0.277	3.46	9.03	0.034	−0.288
		10 %	4.24	7.81	−0.133	−0.226	3.62	8.65	−0.003	−0.282
		5 %	4.41	7.50	−0.164	−0.204	3.98	7.83	−0.083	−0.254
	DF2 × 10SS	full	3.99	7.87	−0.084	−0.254	4.00	7.87	−0.088	−0.252
		20 %	4.76	7.28	−0.220	−0.155	4.22	7.64	−0.129	−0.228
		10 %	5.72	6.52	−0.343	−0.009	4.03	7.90	−0.094	−0.249
		5 %	8.01	5.48	−0.534	0.328	5.40	6.27	−0.306	−0.058
	DF2 × 4_2	full	3.02	8.73	0.163	−0.272	3.21	8.84	0.104	−0.287
		20 %	4.18	7.47	−0.121	−0.233	2.94	9.46	0.188	−0.263
		10 %	4.50	7.16	−0.178	−0.192	3.35	8.37	0.065	−0.289
		5 %	5.46	6.21	−0.314	−0.048	2.87	10.10	0.212	−0.252
	DF2 × 8_2	full	2.60	6.38	0.317	−0.181	2.62	6.28	0.309	−0.188
		20 %	3.52	5.44	0.021	−0.287	2.78	6.01	0.245	−0.233
		10 %	4.46	4.53	−0.172	−0.197	2.84	5.92	0.224	−0.245
		5 %	4.67	4.40	−0.205	−0.169	3.02	5.44	0.162	−0.273
	DF2 × 10_2	full	2.44	5.68	0.388	−0.113	2.37	5.57	0.421	−0.076
		20 %	4.96	3.80	−0.248	−0.126	2.07	6.06	0.583	0.160
		10 %	5.59	3.56	−0.329	−0.028	2.10	5.92	0.566	0.129
		5 %	5.68	3.50	−0.338	−0.015	2.24	5.34	0.489	0.012

Note: (1) Estimated shape and scale and corresponding skewness (Skew) and excess kurtosis (Exc. Kurt) values for 2-parameter Weibull fits (full, censored 20, censored 10, and censored 5) to six In-Grade Douglas Fir data sets, and 2) corresponding estimates for six data sets generated from the full sample 2-parameter Weibull fits to the six In-Grade Douglas Fir data sets. Corresponding results for Hem Fir and Southern Pine can be found in table 2 of Verrill et al.[16]

http://www1.fpl.fs.fed.us/weib2.prog.html. The program also obtained nonparametric estimates of the fifth percentiles of the 26 strength populations.

For In-Grade type data cell j, the 2-parameter Weibull estimate of the probability that the strength of a member of the cell would fall below x was calculated as

$$\text{Prob}_{\text{W},j} = 1 - \exp(-(x/\hat{\lambda}_j)^{\hat{\beta}_j})$$

where $\hat{\lambda}_j$ and $\hat{\beta}_j$ were the maximum likelihood estimates of the scale and shape parameters for cell j, respectively. These estimates are provided in **Table 2** for full, censored 20, censored 10, and censored 5 fits. A censored 20 fit, for example, is one in which in our maximum likelihood fit, we make use of the bottom 20 % of the data and the number, N_j, of specimens sampled for cell j. (N_j corresponds to the n in equation X2.1 of ASTM D5457.[5])

The empirical estimate of this probability was

$$\text{Prob}_{\text{emp},j} = n_j/N_j$$

where n_j was the number of specimens in cell j with MORs that fell below x, and N_j was the total number of specimens sampled for cell j. For the purposes of this article we will refer to these probabilities as "probabilities of failure." We considered cases in which x equaled the nonparametric estimate of the fifth percentile of the distribution divided by 1.9, 2.1, and 2.3 (that is, for cases in which x was in the neighborhood of the allowable property).

In **Table 3** (corresponding to the divisor 2.1), $\text{Prob}_{\text{W,j}} \times N_j$ values (the expected number of failures under full, censored 20, censored 10, and censored 5 2-parameter Weibull fits) are presented in columns 6–9, and n_j (the observed number of failures) is presented in column 10. (Corresponding tables for divisors 1.9 and 2.3 can be found as tables 3 and 5 of Verrill et al.[16]) A listing of the program that was used to produce **Table 3** can be found at http://www1.fpl.fs.fed.us/weib2.prog.html.

When $\text{Prob}_{\text{W},j} \times N_j \gg n_j$, the 2-parameter Weibull fit is likely to be overestimating the true probability of a failure. In some cases, the estimates based on a 2-parameter Weibull fit could be said to be highly inflated (consider the 20.8, 5.4, 11.0, and 5.9 predictions in column 6 of **Table 3**).

For the 2.1 divisor, the total number of cases in which specimen strengths actually lay below nonparametric fifth/2.1 values was 9. That is, the observed overall failure probability was 9/13,275 = 0.00068. For the full, censored 20, censored 10, and censored 5 2-parameter Weibull fits, the expected total numbers of failures were 83.8, 29.8, 15.9, and 12.4. The 83.8 value is more than nine times the observed number of failures and yields an overall probability of failure estimate of 83.8/13,275 = 0.0063.

This factor of nine suggests to us that a 2-parameter Weibull model for the MOR distribution of grades of lumber is a poor one. A 2-parameter Weibull model leads to inflated estimates of probabilities of failure when loads are near allowable properties. This result can be expected from the tail thinning due to pseudo-truncation that was explored in Verrill et al.[1–4] (Of course, the inflation factor might actually vary with the cell. For example, if we restrict ourselves to the SS cells and uncensored fits, the ratio is 10.8/4 = 2.7, whereas if we restrict ourselves to the No. 2 cells and uncensored fits, the ratio is 73.0/5 = 14.6.)

On the other hand, one could argue that despite the fact that Verrill et al.[1,4] have established theoretically that pseudo-truncated strength distributions are unlikely to be 2-parameter Weibulls, and despite the fact that formal goodness-of-fit tests reject 2-parameter Weibull models for the strength distributions of grades of lumber, and despite the fact that 2-parameter Weibull fits to In-Grade data sets lead to overestimates of probabilities of failure for loads near allowable properties, one could—on a purely ad hoc basis—use censored data fits to 2-parameter Weibulls to predict probabilities of failure for loads near allowable properties. Such an argument might be made by someone who noted from **Table 3** (and tables 3 and 5 of Verrill et al.[16]) that although censored data fits also lead to overestimates of the numbers of failures for loads near allowable properties, the overestimates appear to

TABLE 3

Evidence from In-Grade, 2011 SPIB In-Grade, and 2014 SPIB resource monitoring program data that 2-parameter Weibull fits (both uncensored and censored) lead, on average, to inflated estimates of failure probabilities when loads are at the nonparametric 5th/2.1

					Predicted # Failures				
Data Set	Species	Lumber Size	Grade	Sample Size	No cens	Censored 20	Censored 10	Censored 5	Observed # Failures
In-Grade	DF	2 × 4	SS	414	0.77	0.28	0.19	0.04	0
	DF	2 × 8	SS	493	1.60	1.61	1.05	0.98	1
	DF	2 × 10	SS	414	1.17	0.54	0.27	0.06	0
	DF	2 × 4	2	386	2.84	0.83	0.63	0.34	1
	DF	2 × 8	2	1,964	20.84	7.34	3.71	3.18	0
	DF	2 × 10	2	388	5.41	0.48	0.30	0.29	0
	HF	2 × 4	SS	428	1.10	0.26	0.35	0.33	1
	HF	2 × 8	SS	375	0.82	0.45	0.20	0.06	0
	HF	2 × 10	SS	368	0.91	0.48	0.10	0.02	0
	HF	2 × 4	2	406	4.17	0.93	0.20	0.35	0
	HF	2 × 8	2	372	3.52	1.11	0.32	0.19	0
	HF	2 × 10	2	361	3.80	0.41	0.26	0.22	0
	SP	2 × 4	SS	413	0.98	0.18	0.31	0.26	0
	SP	2 × 8	SS	626	1.11	0.80	1.25	1.13	1
	SP	2 × 10	SS	413	0.22	0.10	0.18	0.06	0
	SP	2 × 4	2	413	4.44	0.62	0.25	0.05	1
	SP	2 × 6	2	413	3.87	0.72	0.17	0.02	0
	SP	2 × 8	2	1,367	11.00	4.92	2.30	1.76	2
	SP	2 × 10	2	412	1.43	0.71	0.35	0.12	0
2011	SP	2 × 4	SS	420	1.30	0.50	0.27	0.48	0
	SP	2 × 8	SS	409	0.38	0.89	0.45	0.45	0
	SP	2 × 10	SS	410	0.40	0.13	0.20	0.41	1
	SP	2 × 4	2	408	5.88	0.90	0.16	0.10	0
	SP	2 × 8	2	420	2.76	1.52	0.88	0.59	0
	SP	2 × 10	2	420	1.19	1.94	1.28	0.83	1
2014	SP	2 × 4	2	362	1.88	1.15	0.24	0.05	0
Total					83.80	29.79	15.85	12.39	9
Inflation factor					9.3	3.3	1.8	1.4	

decline as the censoring increases. That is, the upward bias in the estimate of the number of failures appears to decrease as the censoring increases.

Our short response is that the simulations of Verrill et al.[2,3] established that although the bias in the estimate of the probability of failure declines as censoring increases, the variance of the estimate significantly increases (as one would expect from censored data estimates) with the result that, for any given data set, censored data estimates of failure probabilities will have good chances of being seriously inflated or seriously deflated (whereas correct pseudo-truncated estimates will not).

In the next section, we provide a longer response to the suggestion that the strength distributions of visual or MSR grades of lumber can be well approximated by censored data fits of 2-parameter Weibull models.

Censored Data

ASTM D5457[5] permits 2-parameter Weibull distributions to be fit to data via censored data methods. Some scientists refer to this as tail fitting. Two permitted censored data fitting techniques (maximum likelihood and method of least squares) are described in Section X2 of ASTM D5457. To produce a tail fit, one might explicitly use only the bottom 20 % of the data (the bottom n_c data values in the notation of Section X2 of ASTM D5457) and the fact that 80 % of the data (n_s of the data values in the notation of Section X2 of ASTM D5457) exceed the maximum of the bottom 20 %.

In Section X1.1.3 of ASTM D5457, the authors of the standard write: "In addition, by permitting tail fitting of the data, it provides a way of fitting data in this important region that is superior to full-distribution types."

We would argue that this notion of the superiority of tail fitting is mistaken. This issue is relevant, important, and somewhat opaque. Thus, we feel that it is appropriate to address it in some detail here.

Statisticians refer to tail fitting methods as censored data methods and know that they are derived for the case in which we know the probability density function associated with the bottom (in the notation of Section X2 of ASTM D5457) n_c strength values in a sample, know those strength values, and further know that the remaining $n-n_c$ values in the sample are larger than the largest of the bottom n_c values. We might encounter such data if we applied a maximum load to all of the specimens in a data set (rather than loading all of the specimens to failure). We might also encounter such a data set if we chose to simply record strengths larger than the bottom n_c strengths as "larger than the bottom n_c strengths" (as is contemplated in Section X2 of ASTM D5457).

Among statisticians, it is a well-known fact that if our probability models are correct, we obtain better (lower mean squared error) parameter estimates and thus, in reliability situations, better estimates of the probability of failure (anywhere, including the left tail) by performing full data fits rather than censored data fits.

That is, if we have, for example, 400 random draws from a distribution, we obtain better estimates of the parameters of this distribution (and thus, estimates of the percentiles of this distribution) by using explicit knowledge of all 400 data values than by using just the explicit bottom 80 values and knowledge that the remaining 320 values exceed the maximum of the 80 values (the censored data approach covered in Section X2 of D5457). For samples of size 400, we performed simulations that confirmed that when the full underlying distribution is a 2-parameter Weibull, we do better with full data analyses than with censored data analyses. Results from these simulations are reported in table 6 of Verrill et al.[16] A listing of the Fortran computer program that was used to perform the simulations can be found at http://www1.fpl.fs.fed.us/weib2.prog.html.

Thus, there is no theoretical justification for performing a censored data fit if we believe that our model (e.g., a 2-parameter Weibull) holds for the whole population. So, we must assume that an advocate of censored data fits believes that a 2-parameter Weibull does not fit the whole population. What do they believe? We assume that they are thinking one of three things:

1. The population is a mixture—for example, 45 % of MOR values are drawn from one normal (or Weibull or lognormal or …) population, and 55 % of the MOR values are drawn from a different normal (or Weibull or lognormal or …) population. (A pseudo-truncated version of such a model is considered in Verrill et al.[17])
2. The population is a "chimera"—for example, for data below x_c, the population's probability density function at x is the 2-parameter Weibull density $\gamma^{\beta}\beta x^{\beta-1}\exp(-(\gamma x)^{\beta})$, and for $x > x_c$, it is something else. Actually, we assume that there are few wood scientists who would truly believe in such a creature. (What would be the mechanism that yielded the chimera?) But we mention it for completeness.
3. The advocates of censored data fits have no proposed theoretical model for the distribution of the MOR population. They simply believe that, for practical purposes, they can perform a censored data 2-parameter Weibull fit on some portion of the left tail of the data and get fairly decent predictions of, say, the bottom 10–20 % of the data.

The censored data methods described in Section X2 of ASTM D5457[5] are not designed to handle Case 1. In the notation of Section X2 of ASTM D5457, the censored data techniques make use of the lowest n_c values in the data set and the number n in the full random draw from the population. In the mixture case, you don't know the n associated with the leftmost subpopulation. You know N, the total number of observations drawn from the various populations in the mixture. If you had done a mixture analysis, you could estimate n as $\hat{p} \times N$ where $\hat{p}$ was your estimate of the proportion of the leftmost subpopulation in the mixture (even then, you would have to assume that the bottom n_c observations came solely from this leftmost population in the mixture). (Of course, if you did a complete pseudo-truncated mixed analysis as was done in Verrill et al.,[17] you could calculate the complete resulting probability density function and predict probabilities of failure at various loads.)

In the chimera case, if you knew that the underlying probability density function was a 2-parameter Weibull for x less than some x_c, then you could, indeed, perform a censored data fit based on the x_i's that lay below x_c and the total number of observations in the sample. However, as we note previously, we have seen no mechanism advanced for a chimera, and we assume that wood scientists have never actually estimated cutoffs between the various portions of a chimera or actually believe in their existence.

Instead, we assume that proponents of modeling MOR distributions with censored data estimates of 2-parameter Weibulls know that poor fits and probability plots are obtained when full data 2-parameter Weibull fits are made, and they see that censored data fits yield left tail probability plots in which observed and predicted order statistics more closely align. They then argue that for reliability purposes, we are concerned about the left tail (leaving this loosely defined), not the whole distribution, so we need only get a good fit in some portion of the left tail. That is, they only hope to use a censored data 2-parameter Weibull fit as a good interpolator for some portion (bottom 10 %? bottom 20 %?) of the left tail data.

The problem with this approach is that if we do not have a good mechanistic model for the generation of the data (for example, pseudo-truncation of a bivariate mill run distribution if we focus on MOE–MOR data or visual grade–MOR data, or pseudo-truncation of a trivariate mill run distribution if we focus on MOE–visual grade–MOR data), but instead, simply fit an interpolator to the left tail of a visual grade data set, we run into the weakness associated with all empirical models—they tend to perform poorly when we attempt to apply them beyond the data used to fit them. Thus, if we are interested in estimates of probabilities of failure on the order of 0.001 or even 0.0001, we might not do well if we base our predictions on interpolative fits to the bottom 10 or 20 % of a sample of size 400, even if the bottom 10 or 20 % of the observed and predicted data align well along the $y = x$ line in a probability plot. (Of course, we also won't do well if the load can fall above the 5 or 10 or 20 % of the data fully included in the MOR censored data fit, and the form of the appropriate probability density function in this area differs significantly from the 2-parameter Weibull estimated from the censored data fit.)

In support of this claim, we first note that, as stated in Section 4, Verrill et al.[2,3] performed simulations that established that using censored data 2-parameter Weibull methods on pseudo-truncated Weibull data can often lead to probability of failure estimates that are either well above or well below true values, and that this problem is much reduced when estimates are based on the correct pseudo-truncated model. A censored data 2-parameter Weibull approach might yield probability plots that look better than those produced from a full data 2-parameter Weibull approach, but the censored data 2-parameter Weibull approach still performs much more poorly than a correct pseudo-truncated Weibull approach. Censored data 2-parameter Weibull fits to pseudo-truncated Weibull data appear to lead to decreases in the biases of probability of failure estimates (compared with non-censored data 2-parameter Weibull fits) but appear to lead to increases in their variances. In short, increases in censoring lead to interpolation overfitting caused not by an increase in the number of parameters in the model, but by decreases in the range and effective number of data points. Again, for simulation details, see Section 5.4 of Verrill et al.[2]

Second, we note that if there really were an unvarying 2-parameter Weibull that fit low-tail data, then (depending upon where the low-tail is supposed to begin) censored data maximum likelihood techniques would obtain similar (because maximum likelihood methods are asymptotically unbiased regardless of the degree of censoring) estimates whether we fit the bottom 20, bottom 10, or bottom 5 % of the data (unless, of course, the low-tail doesn't begin until we are at or below some or all of these percentiles). However, as displayed in columns 4 and 5 of **Table 2** (the full **Table 2** can be found in Verrill et al.[16]), for 12 of the 14 No. 2 In-Grade type data sets (and 4 of the 12 SS data sets), as censoring goes from none, to 20th percentile, to 10th percentile, to 5th percentile, the shape parameter obtained from a 2-parameter Weibull censored data fit monotonically increases, and the scale parameter monotonically decreases. (A listing of the Fortran program that produced the fits can be obtained at http://www1.fpl.fs.fed.us/weib2.prog.html.) This is not what should occur theoretically or what does occur empirically when we use censored data 2-parameter Weibull methods to fit generated 2-parameter Weibulls—see columns 8 and 9 of **Table 2** and see the discussion of table 3 in Section 5.2 of Verrill et al.[2] However, it is what we expect when we incorrectly use censored data 2-parameter Weibull methods to fit pseudo-truncated Weibull data. See the discussion of table 5 in Section 5.3 of Verrill et al.[2] All of this suggests that low-tail In-Grade type strength data is not well modeled by the left tail of a 2-parameter Weibull, despite, for example, 2-parameter Weibull probability plots that might appear more nearly linear when they are based on censored data fits and only include low-tail data.

The Fortran program that produced **Table 2** also produced plots that make our point visually. The full set of plots is available at http://www1.fpl.fs.fed.us/weib2.156plots.pdf. We attach six of these plots as **figures 3**–**8**.

In **figure 3**, we plot a histogram of the 413 southern pine, 2 × 6, No. 2 In-Grade MOR values. We overlay this histogram with the estimated probability density function from a 2-parameter Weibull maximum likelihood uncensored fit to the In-Grade MOR values and with the estimated probability density functions from censored 20, censored 10, and censored 5 maximum likelihood fits to the In-Grade MOR values. This plot makes clear the systematic change in the fits as the censoring increases. In contrast, in **figure 4**, we plot a histogram of 413 2-parameter Weibull values generated from the full data 2-parameter Weibull fit of the 413 southern pine, 2 × 6, No. 2 In-Grade MOR values. We overlay this histogram with the estimated probability density function from a 2-parameter Weibull maximum likelihood uncensored fit to the generated data, and with the estimated probability density functions from censored 20, censored 10, and censored 5 maximum likelihood fits to the generated data. These fits (to generated data that we know to be 2-parameter Weibull) do not display the systematic changes with increased censoring displayed by the corresponding fits to In-Grade MOR values.

The dependency of left tail predictions on the degree of censoring in our fitting procedures is further illustrated by **figures 5**–**8**. In these figures we plot ordered data versus predicted ordered data. That is, these figures display 2-parameter Weibull probability plots. They correspond, respectively, to full, censored 20, censored 10, and censored 5 2-parameter Weibull fits of the southern pine, 2x6, No. 2 In-Grade data. The horizontal lines in these plots mark the 20th, 10th, and 5th percentiles of the ordered data. In **figures 6**–**8** (corresponding to

FIG. 3

Histogram of In-Grade Southern Pine, 2 × 6, No. 2 data. It is overlaid with a 2-parameter Weibull full data fit (solid line), a 2-parameter Weibull censored 20 fit (dashed line), a 2-parameter Weibull censored 10 fit (dot-dashed line), and a 2-parameter Weibull censored 5 fit (dotted line). Solid vertical lines are plotted at the nonparametric 5th/2.1, and at the nonparametric 5th, 10th, and 20th percentiles of the data.

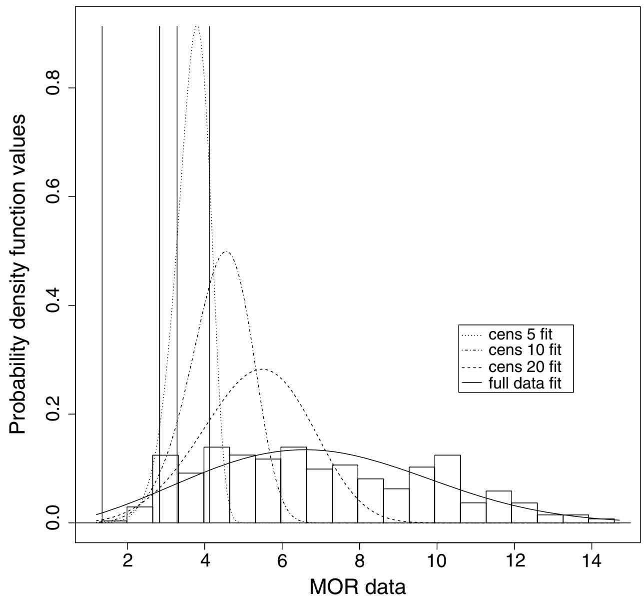


FIG. 4

Histogram of generated In-Grade Southern Pine, 2×6, No. 2 data. It is overlaid with a 2-parameter Weibull full data fit (solid line), a 2-parameter Weibull censored 20 fit (dashed line), a 2-parameter Weibull censored 10 fit (dot-dashed line), and a 2-parameter Weibull censored 5 fit (dotted line). Solid vertical lines are plotted at the nonparametric 5th/2.1, and at the nonparametric 5th, 10th, and 20th percentiles of the data.

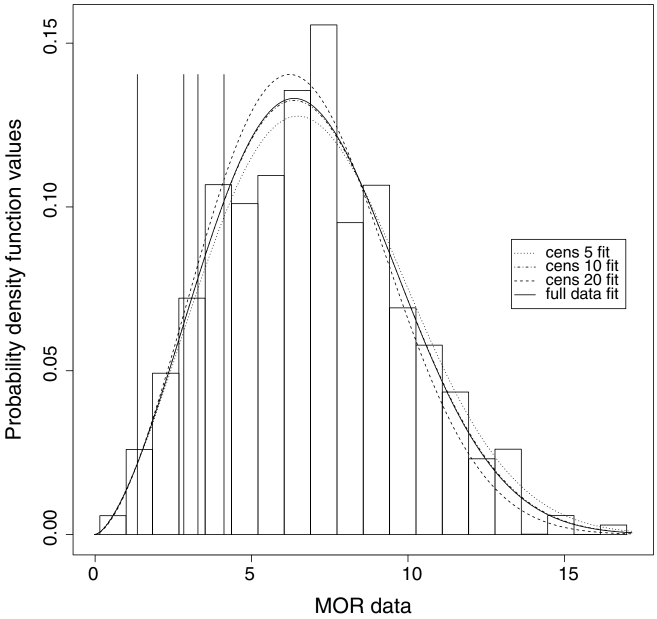
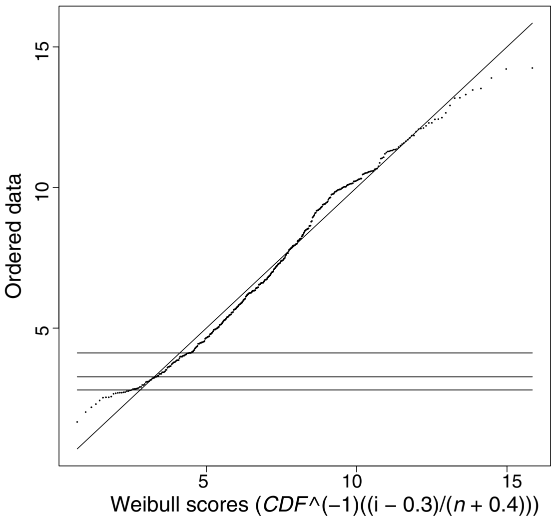


FIG. 5

Weibull probability plot of In-Grade Southern Pine, 2×6, No. 2 data. Ordered observed data versus expected ordered data under a full data 2-parameter Weibull model fit. The straight nonhorizontal line is the $y = x$ line. The horizontal lines are at the 0.20, 0.10, and 0.05 quantiles of the data.



censored 20, censored 10, and censored 5 fits), the fits sharply deviate from the $y = x$ lines at or shortly above the data that are explicitly used in the fit. That is, 5 % fits don't do a good job of predicting 10 % data, and 10 % fits don't do a good job of predicting 20 % data. Locally good interpolants don't continue to perform well beyond the data used to produce the interpolants.

We admit that we are presenting plots from a case that does an especially good job of making our point visually. However, the plots in the 25 other cases make similar points. (All 156 plots can be viewed at

FIG. 6

Weibull probability plot of In-Grade Southern Pine, 2 × 6, No. 2 data. Ordered observed data versus expected ordered data under a censored 20 fit. The straight nonhorizontal line is the *y* = *x* line. The horizontal lines are at the 0.20, 0.10, and 0.05 quantiles of the data.

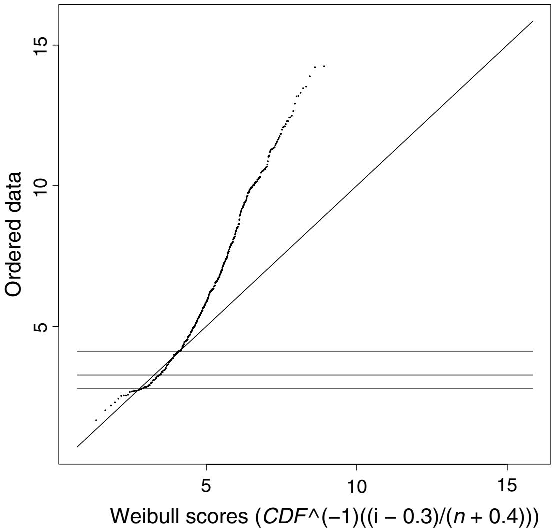
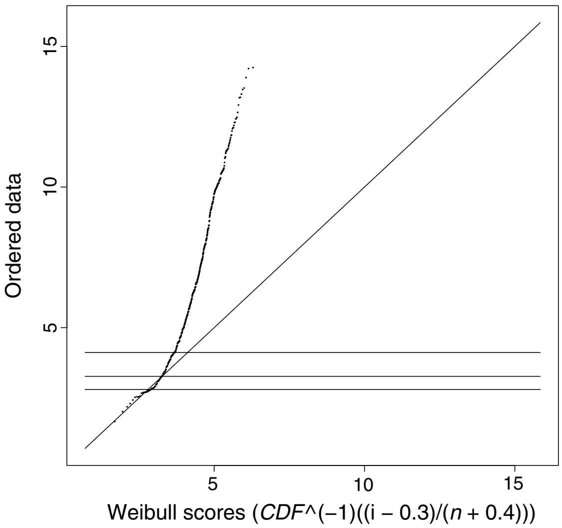


FIG. 7

Weibull probability plot of In-Grade Southern Pine, 2 × 6, No. 2 data. Ordered observed data versus expected ordered data under a censored 10 fit. The straight nonhorizontal line is the *y* = *x* line. The horizontal lines are at the 0.20, 0.10, and 0.05 quantiles of the data.

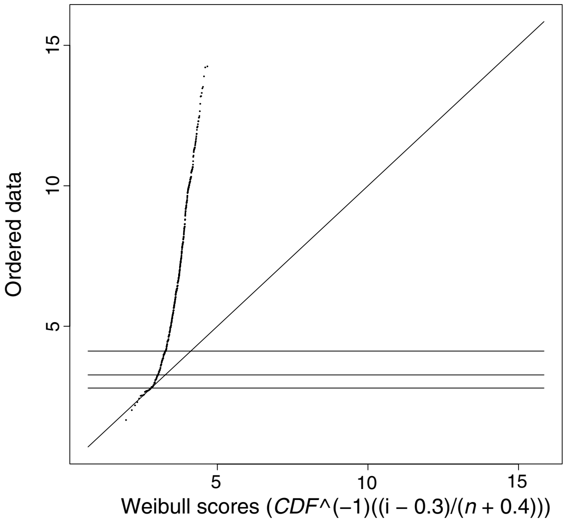


http://www1.fpl.fs.fed.us/weib2.156plots.pdf.) In fact, for 7 of the 14 No. 2–censored 5 probability plots and 4 of the 12 SS–censored 5 probability plots, there is a sharp bend at the 5th percentile of the data. For an additional 4 of the No. 2–censored 5 probability plots and an additional 4 of the SS–censored 5 probability plots, this sharp bend is slightly higher (between the 5th and 10th percentiles of the data).

Given the observed poor predictions above the interpolated data, we see no reason to trust predictions for the important region below the interpolated data. As stated previously, this empirical result is strongly bolstered by the simulation results reported in Section 5.4 of Verrill et al.[2] that establish the superiority of theoretically correct pseudo-truncated fits to theoretically incorrect censored data 2-parameter Weibull fits. (Of course, if we do fits to

FIG. 8

Weibull probability plot of In-Grade Southern Pine, 2 × 6, No. 2 data. Ordered observed data versus expected ordered data under a censored 5 fit. The straight nonhorizontal line is the $y = x$ line. The horizontal lines are at the 0.20, 0.10, and 0.05 quantiles of the data.



sufficiently large data sets, relevant percentiles will lie within the interpolation region, and this will improve the performance of interpolator-based percentile predictions.)

Load Distributions Rather than Fixed Loads

In "The New Analysis," we evaluated estimated and empirical probabilities of failure at fixed loads (at approximate allowable properties). In this section, we consider variable loads. We discuss calculations in which peak load distributions are modeled as lognormals. Such a model was suggested to us by a reviewer of Verrill et al.[2] The reviewer of Verrill et al.[2] suggested that we model the load distribution as a lognormal with coefficient of variation 0.3 that exceeds the allowable property with probability 0.02.

In our calculations, the load was modeled as a lognormal with coefficient of variation 0.3, and we considered the cases in which the load exceeded the approximate allowable property (nonparametric estimate of the 5th percentile divided by 2.1) with probabilities 0.01, 0.02, 0.05, 0.10, and 0.2. The mathematical details of the calculations are provided in the Appendix. The results appear in Table 4. (A listing of the program that was used to produce Table 4 can be found at http://www1.fpl.fs.fed.us/weib2.prog.html.)

The results presented in Table 4 suggest that 2-parameter Weibull fits to strength distributions can also yield inadequate estimates of failure probabilities when we incorporate load distributions into our calculations. For example, for load exceedance probabilities of 0.01 (the probability that the lognormal load exceeds the [approximate] allowable property is 0.01), and censored 20, censored 10, and censored 5 fits, the ratios of 2-parameter Weibull–based estimates of the number of expected failures to data-based estimates of the number of expected failures are, respectively, 4.1/0.54 = 7.6, 1.7/0.54 = 2.2, and 1.2/0.54 = 2.2. (Admittedly, these values do decline as the exceedance probability goes up. For example, for an exceedance probability of 0.02 rather than 0.01, the corresponding values are 5.8, 2.6, and 1.9.)

We note that the results suggest that failure rate biases decline as censoring increases, but we continue to emphasize that this is an improvement in interpolation rather than modeling and that simulations (again, see Section 5.4 of Verrill et al.[2]) suggest that the apparent reduction in the bias of the estimate of failure rates given censored data fits is accompanied by an increase in the variance of the estimate of failure rates and thus an increased chance of serious underestimates of failure rates.

TABLE 4

Evidence that Weibull fits yield inadequate estimates of failure probabilities even when we incorporate load distributions and censoring into the calculations

Data Used in Weibull Fit	Probability that the Load is Above the Allowable Property	Data-based Expected Failures	Weibull Fit–based Expected Failures	Column 4 Divided by Column 3
All	0.01	0.54	17.1	31.7
	0.02	0.96	21.4	22.3
	0.05	2.16	30.1	13.9
	0.10	4.27	41.0	9.6
	0.20	9.39	59.8	6.4
Bottom 20 %	0.01	0.54	4.1	7.6
	0.02	0.96	5.6	5.8
	0.05	2.16	8.9	4.1
	0.10	4.27	13.5	3.2
	0.20	9.39	22.4	2.4
Bottom 10 %	0.01	0.54	1.7	3.1
	0.02	0.96	2.5	2.6
	0.05	2.16	4.3	2.0
	0.10	4.27	7.2	1.7
	0.20	9.39	13.4	1.4
Bottom 5 %	0.01	0.54	1.2	2.2
	0.02	0.96	1.8	1.9
	0.05	2.16	3.3	1.5
	0.10	4.27	5.7	1.3
	0.20	9.39	11.2	1.2

Summary

Past theoretical work (Verrill et al.[1,4]) established that if a mill run MOE–MOR population has a bivariate normal-Weibull distribution, then the MOR distributions of visual grade or MSR subpopulations will be pseudo-truncated Weibulls (with thinned tails). Past empirical work (Verrill et al.[2,3]) and work reported in "An Extension of An Earlier Analysis" of the current article confirm that MOR distributions of visual grades of lumber are not 2-parameter Weibulls and do have thinned tails. Past simulation work (Verrill et al.[2,3]) suggested that 2-parameter Weibull fits to pseudo-truncated Weibull data led to inflated estimates of failure when loads are near allowable properties. Empirical work reported in "The New Analysis" of the current article suggests that modeling visual grade–MOR distributions with 2-parameter Weibulls can lead to estimates of failure probabilities when loads are near allowable properties that are inflated by a factor of 9 (at least for No. 2 lumber and uncensored fits).

Past simulation work (Verrill et al.[2,3]) and work reported in "Censored Data" and "Load Distributions Rather than Fixed Loads" of the current article suggest that, as one would expect, censored data 2-parameter Weibull fits can also perform poorly when applied to pseudo-truncated data. In particular, they can lead to highly variable estimates of probabilities of failure and thus (for different data sets) to both serious overestimates and serious underestimates of probabilities of failure.

Given these theoretical, empirical, and simulation results, we believe that additional full mill run data sets (see, for example, Owens et al.[18,19]) need to be obtained, and additional pseudo-truncated distributions (see, for example, Verrill et al.[17]) need to be developed in an attempt to identify alternatives to the 2-parameter Weibull as a model for visual and MSR strength distributions. We are engaged in such work.

However, given the fact that actual distributions may be complicated mixtures of base distributions that vary from mill to mill, region to region, time to time, size to size, and species to species, it may be that no satisfactory theoretical form(s) can be identified to form the basis of sophisticated reliability models that could yield improved design values.

We suspect that ultimately, if reliability engineers want to obtain accurate reliability estimates, they will need to develop detailed computer models that yield real-time, in-line estimates of lumber strength based on measurements of stiffness, specific gravity, knot size, and location, slope of grain, and other strength predictors.

Appendix—Calculations for Table 4

The calculations that yielded Table 4 were performed via a Fortran program that can be found at http://www1.fpl.fs.fed.us/weib2.prog.html. In this Appendix, we describe these calculations. None of the mathematics is novel. We describe it in some detail simply because this permits easy checking of our work. Some of the description replicates material found in the Appendix to Verrill et al.[2]

OBTAINING THE PARAMETERS OF THE LOGNORMAL LOAD DISTRIBUTION

By definition, a random variable X is distributed as a lognormal(μ,σ^2) if ln(X) is distributed as a normal(μ,σ^2). Thus, characterizing the lognormal is equivalent to determining the two parameters μ and σ. For a lognormal distribution, it can be shown that the coefficient of variation, CV, is given by

$$\mathrm{CV} = \sqrt{\exp(\sigma^2) - 1}$$

or

$$\ln(1 + \mathrm{CV}^2) = \sigma^2 \tag{A.1}$$

Thus, for lognormals, the parameter σ can be determined from a knowledge of the (alternative parameter) CV. In the calculations that produced Table 4, CV = 0.3 so

$$\sigma^2 = \ln(1.09) \approx 0.0862 \tag{A.2}$$

Next, we show that we can obtain μ_i for the ith of the 26 In-Grade type data sets from σ and the probability, q, that the lognormal lies above the approximate allowable property, a_i (the nonparametric estimate of the 5th percentile divided by 2.1 of the ith In-Grade type data set). In Table 4, the q values appear in column 2.

We have

$$\mathrm{Prob}(\mathrm{LN}(\mu_i, \sigma^2) \le a_i) = 1 - q$$

or

$$\mathrm{Prob}(\mathrm{N}(\mu_i, \sigma^2) \le \ln(a_i)) = 1 - q$$

or

$$\mathrm{Prob}\left(\mathrm{N}(0,1) \le \frac{\ln(a_i) - \mu_i}{\sigma}\right) = 1 - q$$

or

$$\frac{\ln(a_i) - \mu_i}{\sigma} = \Phi^{-1}(1 - q)$$

where Φ denotes the N(0,1) cumulative distribution function. Thus,

$$\mu_i = \ln(a_i) - \sigma \times \Phi^{-1}(1-q) \tag{A.3}$$

So, given the CV of the lognormal load distribution (we assume it to be 0.3) and the probability, q (0.01, 0.02, 0.05, 0.1, or 0.2 in our calculations), that the lognormal load distribution exceeds the approximate allowable property, a_i, we can use equations (A.1) and (A.3) to calculate the mean and variance, μ_i and σ^2, needed to characterize the lognormal load distribution appropriate for data set, i, and exceedance probability, q.

OBTAINING COLUMNS THREE AND FOUR OF TABLE 4

Let

$$f_{\mathrm{LN},i}(y) = \frac{1}{\sqrt{2\pi}} \times \exp(-((\ln(y)-\mu_i)/\sigma)^2/2) \times \frac{1}{\sigma} \times \frac{1}{y} \tag{A.4}$$

denote the lognormal probability density function appropriate for data set i, a CV equal to 0.3, and exceedance probability q ($q \in \{0.01, 0.02, 0.05, 0.10, 0.20\}$). Then, column 3 in Table 4 contains values of the form

$$\sum_{i=1}^{26} N_i \times \mathrm{Prob}(\mathrm{Load}_i > \mathrm{Strength}_i) \approx \sum_{i=1}^{26} N_i \times \quad \sum_{j=1}^{N_i} \int_{w_{ij}}^{\infty} f_{\mathrm{LN},i}(y)\mathrm{d}y \qquad N_i \tag{A.5}$$

where w_{ij} is the jth observed strength value for In-Grade type data set i ($i \in \{1,\ldots,26\}$) and N_i is the number of lumber specimens in data set i.

Here

$$\sum_{j=1}^{N_i} \int_{w_{ij}}^{\infty} f_{\mathrm{LN},i}(y)dy \qquad N_i$$

is a data-based estimate of the probability that the lognormal load would be greater than the strength in the ith of the 26 cases.

Column 4 in Table 4 contains

$$\sum_{i=1}^{26} N_i \int_0^{\infty} f_{\mathrm{LN},i}(x) \int_0^{x} \hat{\gamma}_i^{\hat{\beta}_i} \hat{\beta}_i w^{\hat{\beta}_i - 1} \exp(-(\hat{\gamma}_i w)^{\hat{\beta}_i}) dw\, dx \tag{A.6}$$

where $\hat{\beta}_i$ and $\hat{\gamma}_i$ are the maximum likelihood estimates of the shape parameter and the inverse of the scale parameter for the ith of the 26 In-Grade type data sets and the corresponding censoring level. (The censoring level is indicated in column 1 of the table).

Here

$$\int_0^{\infty} f_{\mathrm{LN},i}(x) \int_0^{x} \hat{\gamma}_i^{\hat{\beta}_i} \hat{\beta}_i w^{\hat{\beta}_i - 1} \exp(-(\hat{\gamma}_i w)^{\hat{\beta}_i}) dw\, dx$$

is a theoretical estimate of the probability that the lognormal load would be greater than the strength in the ith of the 26 cases under the assumption (essentially) of a 2-parameter Weibull left tail of the strength distribution (or at least of a 2-parameter Weibull strength distribution in the region of practical overlap of the load and strength distributions).

References

1. S. P. Verrill, J. W. Evans, D. E. Kretschmann, and C. A. Hatfield, *Asymptotically Efficient Estimation of a Bivariate Gaussian–Weibull Distribution and an Introduction to the Associated Pseudo-Truncated Weibull, Research Paper FPL-RP-666* (Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, 2012).
2. S. P. Verrill, J. W. Evans, D. E. Kretschmann, and C. A. Hatfield, *An Evaluation of a Proposed Revision of the ASTM D 1990 Grouping Procedure, Research Paper FPL-RP-671* (Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, 2013).
3. S. P. Verrill, J. W. Evans, D. E. Kretschmann, and C. A. Hatfield, "Reliability Implications in Wood Systems of a Bivariate Gaussian–Weibull Distribution and the Associated Univariate Pseudo-Truncated Weibull," *Journal of Testing and Evaluation* 42, no. 2 (March 2014): 412–419, https://doi.org/10.1520/JTE20130019
4. S. P. Verrill, J. W. Evans, D. E. Kretschmann, and C. A. Hatfield, "Asymptotically Efficient Estimation of a Bivariate Gaussian-Weibull Distribution and an Introduction to the Associated Pseudo-Truncated Weibull," *Communications in Statistics– Theory and Methods* 44, no. 14 (2015): 2957–2975, https://doi.org/10.1080/03610926.2013.805626
5. *Standard Specification for Computing Reference Resistence of Wood-Based Materials and Structural Connections for Load and Resistance Factor Design*, ASTM D5457-19 (West Conshohocken, PA: ASTM International, 2019). https://doi.org/10.1520/D5457-19
6. J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, 2nd ed. (Hoboken, NJ: John Wiley and Sons, 2003).
7. Task Committee on Load and Resistance Factor Design for Engineered Wood Construction, *Load and Resistance Factor Design for Engineered Wood Construction: A Pre-Standard Report* (New York, NY: American Society of Civil Engineers, 1988).
8. D. W. Green and J. W. Evans, *Mechanical Properties of Visually Graded Dimension Lumber: Vol. 4. Southern Pine, Publication PB-88-159-413* (Springfield, VA: National Technical Information Service, 1988).
9. D. W. Green, B. E. Shelley, and H. P. Vokey, eds., *In-Grade Testing of Structural Lumber, Conference Proceedings 47363* (Madison, WI: Forest Products Research Society, 1989).
10. J. W. Evans and D. W. Green, *Mechanical Properties of Visually Graded Dimension Lumber: Vol. 2. Douglas Fir-Larch, Publication PB-88-159-397* (Springfield, VA: National Technical Information Service, 1988).
11. *Standard Test Methods for Mechanical Properties of Lumber and Wood-Base Structural Material*, ASTM D4761-18 (West Conshohocken, PA: ASTM International, 2018). https://doi.org/10.1520/D4761-18
12. *Standard Practice for Establishing Allowable Properties for Visually-Graded Dimension Lumber from In-Grade Tests of Full-Size Specimens*, ASTM D1990-16 (West Conshohocken, PA: ASTM International, 2016). https://doi.org/10.1520/D1990-16
13. D. E. Kretschmann, J. W. Evans, and L. Brown, *Monitoring of Visually Graded Structural Lumber, Research Paper FPL-RP-576* (Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, 1999).
14. R Core Team, *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing), 2013, http://web.archive.org/web/20190328220111/https://www.r-project.org/
15. M. Krit, EWGoF: Goodness-of-Fit Tests for the Exponential and Two-Parameter Weibull Distributions (R package version 2.0), 2014, http://web.archive.org/web/20190328220559/https://cran.r-project.org/web/packages/EWGoF/index.html
16. S. P. Verrill, F. C. Owens, D. E. Kretschmann, R. Shmulsky, and L. S. Brown, "Visual and MSR Grades of Lumber Are Not 2-Parameter Weibulls and Why It Matters (With a Discussion of Censored Data Fitting)," USDA Forest Products Laboratory draft research paper, http://web.archive.org/web/20190328222315/https://www1.fpl.fs.fed.us/weib2.new.pdf.
17. S. P. Verrill, F. C. Owens, D. E. Kretschmann, and R. Shmulsky, *A Fit of a Mixture of Bivariate Normals to Lumber Stiffness–Strength Data, Research Paper FPL-RP-696* (Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, 2018).
18. F. C. Owens, S. P. Verrill, R. Shmulsky, and D. E. Kretschmann, "Distributions of MOE and MOR in a Full Lumber Population," *Wood and Fiber Science* 50, no. 3 (2018): 265–279, https://doi.org/10.22382/wfs-2018-027
19. F. C. Owens, S. P. Verrill, R. Shmulsky, and R. J. Ross, "Distributions of Modulus of Elasticity and Modulus of Rupture in Four Mill Run Lumber Populations," *Wood and Fiber Science* (in press).