EFFECTS OF SAMPLE SIZE ON CHARACTERIZATION OF WOOD-PARTICLE LENGTH DISTRIBUTION

Quang V. Cao
Professor

Fei Yao
Research Associate

*Qinglin Wu**†
Professor
School of Renewable Natural Resources
Louisiana State University Agricultural Center
Baton Rouge, LA 70803

(Received October 2009)

Abstract. The effects of sample size on fitting length distribution of wood particles used for manufacturing wood-based composites were investigated. A simulation study was conducted to evaluate the variations of the first four sample moments and the ability of the sample distributions to characterize the population represented by the original data. Results showed that a sample size of 2000 was deemed necessary to adequately represent distributions of wood particle length.

Keywords: Composite, K-S statistic, log normal, oriented strandboard, segmented distribution, Weibull.

INTRODUCTION

The dimensions (eg length, width, and thickness) of particles used for manufacturing wood composites as influenced by the milling, drying, blending, and mat-forming processes are known to affect various physical and mechanical properties of the final products. These properties include horizontal density distribution, dimensional stability, elastic modulus, and strength (Steiner and Xu 1995; Suzuki and Takeda 2000; Nishimura and Ansell 2002; Nishimura et al 2002, 2004; Sumardi et al 2008). Therefore, accurate simulation of the dimension distribution for wood particles can lead to an improved quality control process during product manufacturing and better prediction of the final product properties.

The number of observations collected in a sample (ie sample size) is an important factor in the estimation of particle dimension distributions. A distribution derived from a small sample size may not be representative of the population. Statistical rules such as the law of large numbers and the central-limit theorem state that larger sample size leads to increased precision in estimating various properties of the population. However, larger sample size obviously demands more time and labor in collection and processing the material. It is therefore of practical importance to determine a minimum sample size that can adequately model wood-particle distribution.

There is no clear-cut answer to the question of determining a minimum or adequate sample size to characterize distributions of wood-particle length. It depends on various factors such as material type (fiber, particle, strand, flake, etc), statistical distribution, and precision tolerance. Different sample sizes have been chosen by researchers in previous studies. For example, 1850 samples were observed in the study carried out by Kruse et al (2000) on the relationship between sample dimensions and horizontal density distributions of oriented strandboard (OSB). Chen et al (2008) measured dimensions of almost 9000 strands from nine 1.22- × 2.44-m panels

* Corresponding author: QWu@agcenter.lsu.edu
† SWST member

in their study on strand characteristics and alignment of commercial OSB panels.

Different sample sizes may make it difficult to compare results from multiple sources. Furthermore, too large a sample produces little return for the extra time and effort invested, whereas too small a sample might result in a distribution different from that of the population. The objective of this study was to investigate the effect of sample size on fitting length distributions of wood particles and strands. The emphasis was on determining an estimate of adequate sample sizes by means of analyzing several statistical indices.

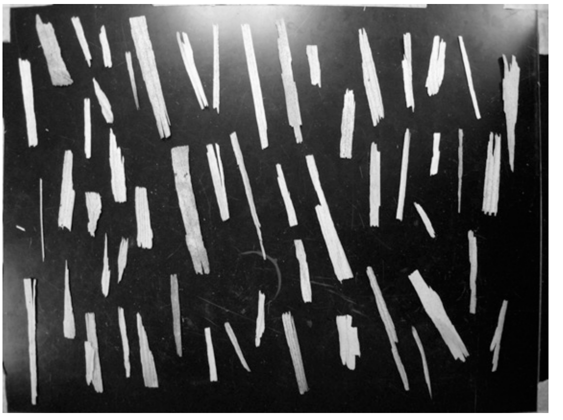

Figure 1. Typical wood particle and strand samples used in the study.

EXPERIMENTAL DATA

Length data for this study came from OSB particles and strands obtained from three different locations in an OSB plant. They were denoted as dry fine (DF), blended fine (BF), and mat reject (MR), respectively. Three packages of samples with similar weights were collected from each of the three sites, and they were sealed in plastic bags and stored at room temperature and moderate humidity. Two of the three packages were then randomly selected for measurements.

All particles and strands in the two selected packages were measured using a digital camera placed in a stationary stand. The particles were manually separated and placed under the view field of the camera. The distance between the lens and wood particle surface was kept unchanged so that particle lengths in all pictures were directly comparable. Image-Pro Plus software (Media Cybernetics Inc, Silver Spring, MD) was used to measure the particle length from the digital photographs. Calibration was done with a ruler to a precision of ± 0.1 mm. Figure 1 shows a sample digital photo taken from an MR sample. A total of 24,357 DF observations were collected at the exit of dryers in the plant, and the samples were adhesive-free particles used as the core layer of an OSB mat. BF data (29,782 observations) were collected at the exit of the blender, and the sample was composed of adhesive-coated DF. MR (17,612 observations) was collected at the forming line, and consequently the sample included all adhesive-coated particles and strands from both face and core layers.

METHODS

A preliminary investigation showed that the segmented distribution (Cao and Wu 2007) was more appropriate for modeling the distribution of wood-particle length than a Weibull or log normal distribution, or a mixture of both. The segmented distribution fit the data better than either the Weibull or log normal distribution and has one fewer parameter than the mixture of the Weibull and log normal. The segmented distribution as defined by Cao and Wu (2007) has the following probability density function (pdf):

$$f(x) = \begin{cases} f_1(x)/\beta, & 0 \leq x \leq t \\ \alpha f_2(x)/\beta, & x > t \end{cases} \tag{1}$$

where

$$f_1(x) = \left(\frac{1}{x\sigma\sqrt{2\pi}}\right) \exp\left[-\frac{1}{2}\left(\frac{\ln(x) - \mu}{\sigma}\right)^2\right]$$

= log normal pdf with parameters μ and σ,

$$f_2(x) = \left(\frac{cx^{c-1}}{b^c}\right) \exp\left[-\left(\frac{x}{b}\right)^c\right]$$

= Weibull pdf with scale parameter b and shape parameter c,

t = the join point, fixed as the median of x

$\alpha = f_1(t)/f_2(t)$, so that $f(x)$ is continuous at the join point t

$\beta = \int_0^t f_1(x)dx + \int_t^\infty \alpha f_2(x)dx = F_1(t) + \alpha[1 - F_2(t)]$; β is used to scale $f(x)$ such that $f(x)$ integrates to one.

Figure 2 shows the frequencies of wood-particle length for DF, BF, and MR and the segmented distributions that approximated these frequencies.

A simulation study was then conducted to examine the effects of sample size on the ability of samples to represent the wood-particle length distribution in the population. Simulations were carried out by increasing sample size from 500 to the population sizes (29,782, 24,357, and 17,612 for BF, DF, and MR, respectively). For each sample size, 1000 samples were generated by sampling with replacement from the population. The simulation was repeated for the three types of wood particles.

Evaluation of Sample Moments

For each sample, the first four moments were computed: mean, variance, skewness, and kurtosis. The variations of the sample moments were evaluated by use of the coefficient of variation (CV), which is defined here as:

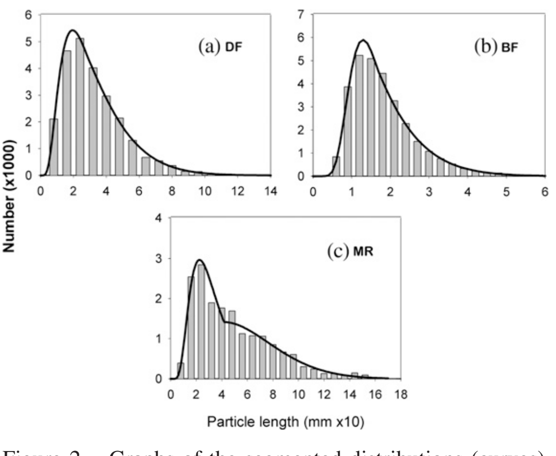

Figure 2. Graphs of the segmented distributions (curves) that fit the observed data (bars) for the three wood particle types.

$$\mathrm{CV} = \frac{\sum_{i=1}^{1000} (\hat{y}_i - y)^2}{1000y} \tag{2}$$

where y = one of the four true moments (calculated from the original data),

$\hat{y}_i$ = estimate of y from the i^{th} sample.

Evaluation of Goodness-of-Fit

For each type of wood particle and for each of the 1000 generated samples, the maximum likelihood technique (Cao and Wu 2007) was used to obtain estimates of parameters of the segmented distribution that characterized this sample. The resulting distribution for each sample was then evaluated against the original data (representing the population) by use of the Kolgomorov-Smirnov (KS) goodness-of-fit statistic. The K-S statistic has been widely used in testing that a theoretical distribution fits an empirical data set. A small KS value indicates that the segmented distribution from the sample did a good job in fitting the population.

RESULTS AND DISCUSSION

Evaluation of Sample Moments

Figure 3 shows CVs for the first four sample moments as related to sample size for the three types of wood particles. Curves from the three wood-particle types were indistinguishable from one another for all four sample moments. These curves seemed to follow a negative exponential function, in which the variation of the four sample moments decreased rapidly as sample size increased from 500 to 2000 and then decreased at a lower rate for sample size over 2000.

Evaluation of Goodness-of-Fit

The mean of the K-S goodness-of-fit statistics also decreased as sample size increased (Fig 4). The MR data included all kinds of particles and strands from both face and core layers, resulting in a more irregular distribution (Fig 2c), which

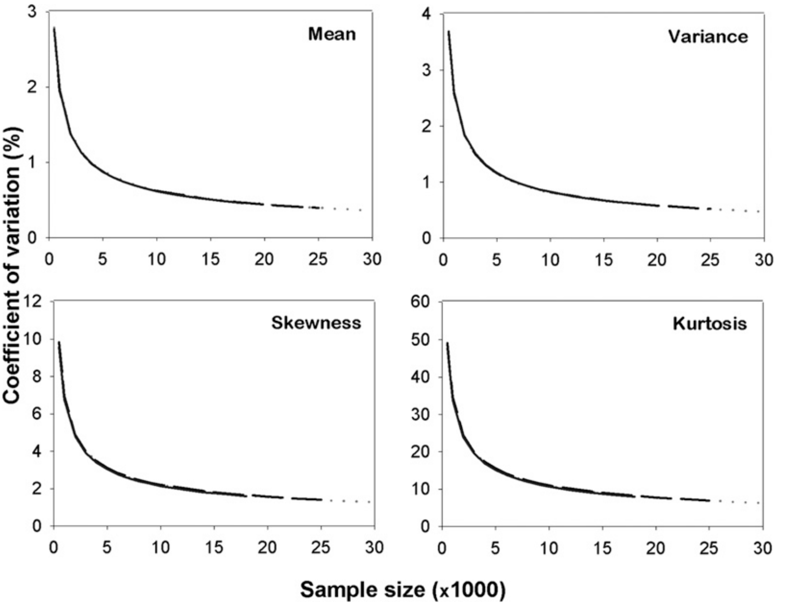

Figure 3. Coefficients of variation for the first four sample moments as related to sample size for blended fine (BF) (dotted), dry fine (DF) (dashed), and mat reject (MR) (solid). Curves from the three wood-particle types were indistinguishable from one another for all four sample moments.

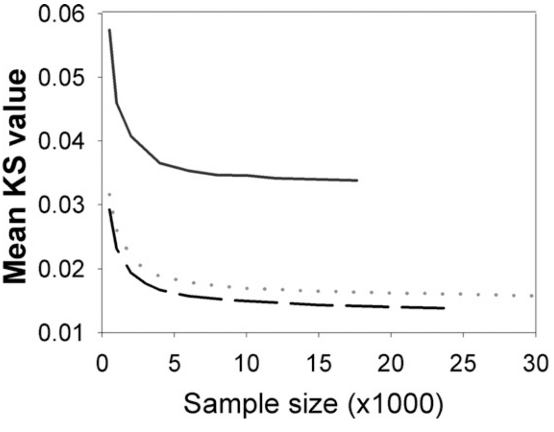

Figure 4. Mean Kolgomorov-Smirnov values as related to sample size for blended fine (BF) (dotted), dry fine (DF) (dashed), and mat reject (MR) (solid).

in turn produced higher K-S values than those for BF and DF. Nevertheless, the curve shapes remain similar for all three wood-particle types. Similar to Fig 3, the curve shapes were of a negative exponential type, which exhibited a sharp drop in mean K-S value as sample size increased from 500 to 2000. The rate of decrease in mean K-S value was more gradual for sample size over 2000.

Adequacy of Sample Size

These results indicated that there was no clear-cut answer to the question of ascertaining a minimum sample size to characterize distributions of wood-particle length. We used the following graphical technique in an attempt to determine an adequate sample size.

The method was based on two tangents to the curve connecting data points from each statistic (Figs 3 and 4). The tangent starting from the smallest sample size was a straight line with intercept a_1 and slope b_1. On the other hand, the tangent starting from the largest sample size was a line of intercept a_2 and zero slope, because all the statistics approached a lower asymptote as sample size became very large. The bisector of the angle formed by these two tangent lines divided the points into two halves, the left side of steeper slopes and the right side of less steep slopes. The equation for the angle bisector is:

$$y = y_0 + b_3(x - x_0) \quad (3)$$

where y = value of an evaluation statistic

x = sample size

x_0 and y_0 = coordinates of the intersection of the two tangents

$b_3 = -1/[\tan\{0.5*\tan^{-1}(b_1)\}]$ = slope of the angle bisector

The intersection of the angle bisector and the curve connecting the data points indicate an adequate sample size that provides a reasonable tradeoff between accuracy/precision and effort. This graphical technique yielded a sample size of approximately 2000 based on the statistics from sample moments (Fig 3) and about 1000 based on the K-S statistics (Fig 4). A minimum sample size of 2000 is therefore consistent with what we observed: all of the evaluation statistics exhibited a sharp drop in value as sample size increased 500 to 2000, and the rate of decrease was more gradual for sample size over 2000.

CONCLUSIONS

The effects of sample size in fitting distributions of wood-strand length followed a negative exponential function, in which CV for four sample moments and the mean K-S statistics dropped sharply before sample size reached 2000 and afterward decreased gradually as sample size exceeded 2000. This observation was confirmed by a graphical technique based on the bisector of the angle formed by the two outermost tangents to the curve. These results from the simulation study revealed that a sample size of 2000 was deemed necessary to adequately represent distributions of wood-particle length.

REFERENCES

Cao QV, Wu Q (2007) Characterizing wood fiber and particle length with a mixture distribution and a segmented distribution. Holzforschung 61(2):124 – 130.

Chen SG, Du CG, Wellwood R (2008) Analysis of strand characteristics and alignment of commercial OSB panels. For Prod J 58(6):94 – 98.

Kruse K, Dai C, Pielasch A (2000) An analysis of strand and horizontal density distributions in oriented strand board (OSB). Holz Roh Werkst 58(4):270 – 277.

Nishimura T, Amin J, Ansell MP (2004) Image analysis and bending properties of model OSB panels as a function of strand distribution, shape and size. Wood Sci Technol 38(4):297 – 309.

Nishimura T, Ansell MP (2002) Monitoring fiber orientation in OSB during production using filtered image analysis. Wood Sci Technol 36(3):229 – 239.

Nishimura T, Ansell MP, Ando N (2002) Evaluation of the arrangement of wood strands at the surface of OSB by image analysis. Wood Sci Technol 36(1):93 – 99.

Steiner, PR, Xu W (1995). Influence of flake characteristics on horizontal density distribution of flakeboard. For Prod J 45(4):61 – 66.

Sumardi I, Kojima Y, Suzuki S (2008) Effects of strand length and layer structure on some properties of strandboard made from bamboo. J Wood Sci 54(2):128 – 133.

Suzuki S, Takeda K (2000) Production and properties of Japanese oriented strand board I: effect of strand length and orientation on strength properties of sugi oriented strand board. J Wood Sci 46(4):289 – 295.