

The Statistics of Wood Assays for Preservative Retention

Patricia K. Lebow

USDA Forest Products Laboratory
Madison, Wisconsin

Scott W. Conklin

Universal Forest Products, Inc.
Grand Rapids, Michigan

ABSTRACT

This paper covers general statistical concepts that apply to interpreting wood assay retention values. In particular, since wood assays are typically obtained from a single composited sample, the statistical aspects, including advantages and disadvantages, of simple compositing are covered.

Keywords: Wood assay, retention value, composite, central limit theorem

The Forest Products Laboratory is maintained in cooperation with the University of Wisconsin. This article was written and prepared in part by U.S. Government employees on official time, and it is therefore in the public domain and not subject to copyright. The use of trade or firm names in this publication is for reader information and does not imply endorsement by the U.S. Department of Agriculture of any product or service.

INTRODUCTION

The issues and/or questions addressed in this paper center around defining and understanding the natural and sampling variability that can arise in determining assay retention values for a batch of treated lumber as part of a treating plant's QA/QC program. The basic process involves putting trams (small rail cars) of wood product into a treating cylinder and then treating the product following a predefined schedule. After treatment, the product is removed from the cylinder and cores are taken (minimum of 20). These cores are then used to assess preservative penetration based on a measurement made on each core and the retention of preservative in a specified assay zone, commonly 0.0 to 0.6 in. (0-15 mm) for dimensional lumber and 0.0 to 1.0 in. (0-25mm) for 4x4 and larger lumber. For retention measurement the assay zone of all cores in a charge are composited for chemical analysis (an x-ray measurement, typically). Each charge is tested by the treater with approximately 5% of the charges also tested by a third-party approximately on a monthly basis (i.e., after the treated product has cured for a period of time).

The AWP standards provide guidance on many aspects of this process. Several of the standards related to assay retentions include A9 (standard method for analysis of treated wood and treating solutions by x-ray spectroscopy—calibration, sample preparation, calculations), A12 (density of wood species), T1 (use category system, general) and M2 (inspection of treated wood products (cores and borings)). Details can be found in the AWP Standards (2010) which is published annually.

Retention, addressed herein, is a measure of the concentration of preservative retained after treatment in a specified assay zone. Retention is expressed in kg/m^3 (lb/ft^3) and is calculated from the weight of preservative in a given volume of treated wood. For water-borne preservatives, the dry weight of the preservative active ingredients is used in the calculations.

Tsoumis (1991) lists factors that can affect treatment and, hence, retention levels, including method of treatment, type of preservative, target preservative retention level and wood characteristics. Particular product and dimension characteristics are also factors. Preparation (especially moisture content) and anatomical structure, sapwood/heartwood, density, species, grain orientation (especially with regard to direction of impregnation—ratio of axial-to-side permeability) can all contribute to treatment variability. In regards to species, significant factors that can contribute to treatment differences include softwood versus hardwood, especially the occurrence of pits in softwoods and tyloses in heartwood of hardwoods. Schultz, et al. (2004) discuss within board and between board retention variation in southern yellow pine treated with waterborne preservatives, as well as relationships between retention values and specific gravities and other factors that can influence retention levels.

BACKGROUND STATISTICS NECESSARY FOR UNDERSTANDING VARIABILITY

Knowledge of the previously listed factors and how they contribute to variability in an observed retention value can help with understanding both measurement and process variation. These factors can be a key component of observed variability when a single batch is assessed using different sample sets (Barrentine, 2003). Since typically only a single composite measurement is made for a single charge, it is difficult to interpret its value. However, assuming the sampling procedure and

AMERICAN WOOD PROTECTION ASSOCIATION

compositing follow certain guidelines along with other statistical assumptions, inferences about its value are possible (Patil et al. 2011).

Consider the first hypothetical situation examining a single board from which an infinite number of possible cores may be taken and a measurement of the retention is made on the outer zone of each core. Plotting a histogram of those core measurements may resemble the typical bell-shaped normal curve; however, retention values must be nonnegative and it is likely the distribution of the individual core retention values may actually be skewed, such as occurs in the lognormal distribution. However, to understand retention values of the cores, for now assume that the underlying distribution has some finite mean, μ , and finite variance, σ^2 (note the population standard deviation is σ).

It is not realistic to measure every possible core, thus consider that a finite number of cores, n , are taken from random locations on the side of the board from the appropriate assay zone and each is measured for a retention value. Based on this simple random sample, simple statistics can be calculated, such as the sample mean:

$$\bar{X} = \sum_{i=1}^n X_i / n,$$

and sample standard deviation:

$$S = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)}$$

Statistical theory describes the behavior of the sample statistics, and allows inferences about the underlying distribution parameters μ and σ based on the realized estimates $\bar{X}$ and s (Schilling and Neubauer, 2009). In particular, the central limit theorem states:

If random samples of n observations are taken from a population with mean μ and standard deviation σ (where μ and σ are finite), then the distribution of the sample mean $\bar{X}$ is approximately normally distributed with mean μ and standard deviation $\sigma / \sqrt{n}$. The approximation generally increases in accuracy as n becomes larger (Brown and Hollander, 1977).

This, in part, allows for the construction of confidence intervals for the true population mean, such as the typical 95% confidence interval:

$$\bar{X} - 1.96 * \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 * \frac{S}{\sqrt{n}}$$

For example, assume we have an underlying normal distribution with population parameters, $\mu=0.15$ and standard deviation $\sigma=0.06$. Using the statistical computing package R, nine separate random samples of size $n=10$ were generated from this parent distribution, with histograms displayed in Figure 1.

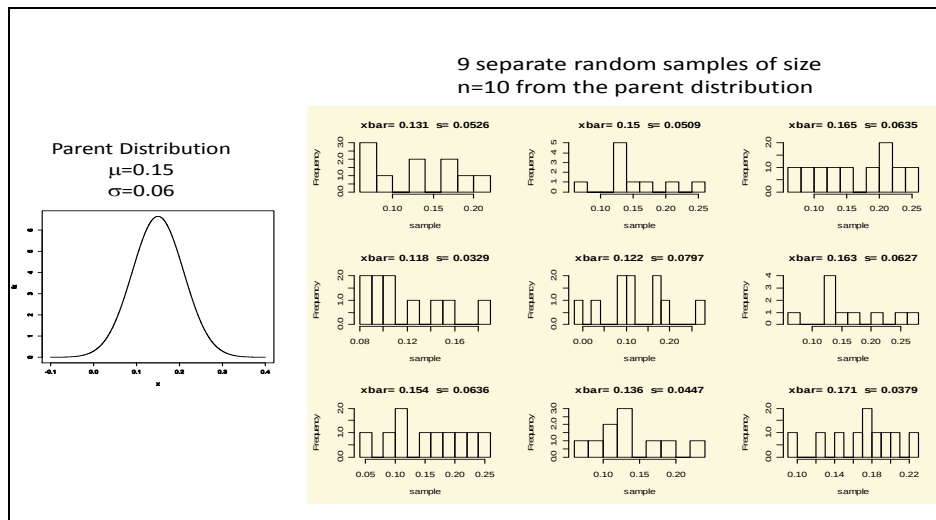


Figure 1. Nine separate simulations of a simple random sample (size $n=10$) from the parent distribution.

AMERICAN WOOD PROTECTION ASSOCIATION

For each simulation, sample statistics $\bar{x}$ and s resemble the parent population parameters. In this case, the average of the nine sample means is 0.145 and the standard deviation of those sample means is 0.0195. Table 1 summarizes what happens as the simulation process is repeated over and over again. The collection of sample statistics for $\bar{x}$ look more and more like what the central limit theorem claims, that the average of the sample means approaches a distribution with mean $\mu=0.15$ and standard deviation $\sigma/\sqrt{n}=0.06/\sqrt{10}=0.0190$ when sample size is 10 or standard deviation $\sigma/\sqrt{n}=0.06/\sqrt{20}=0.0134$ when sample size is 20. That is, as sample size increases from 10 to 20, the standard deviation of the sample mean decreases to $0.0190/\sqrt{2}=0.0134$.

Table 1. Simulations illustrating central limit theorem.

Number of simulations	Sample size = 10		Sample size=20	
	Average of sample means	Standard deviation of sample means	Average of sample means	Standard deviation of sample means
9	0.145	0.0195	0.151	0.0128
1000	0.151	0.0185	0.150	0.0130
10000	0.150	0.0189	0.150	0.0136

When applied to data with a normal distribution, the central limit theorem is exact; for other distributions it is approximate, although remarkably robust as sample sizes increase. Figure 2 illustrates the behavior of the sample mean when the underlying distribution is lognormal. As the sample size decreases from 20, the sampling distribution will become wider and less normal looking; as sample size increases beyond 20, the sampling distribution centered about the mean will become narrower.

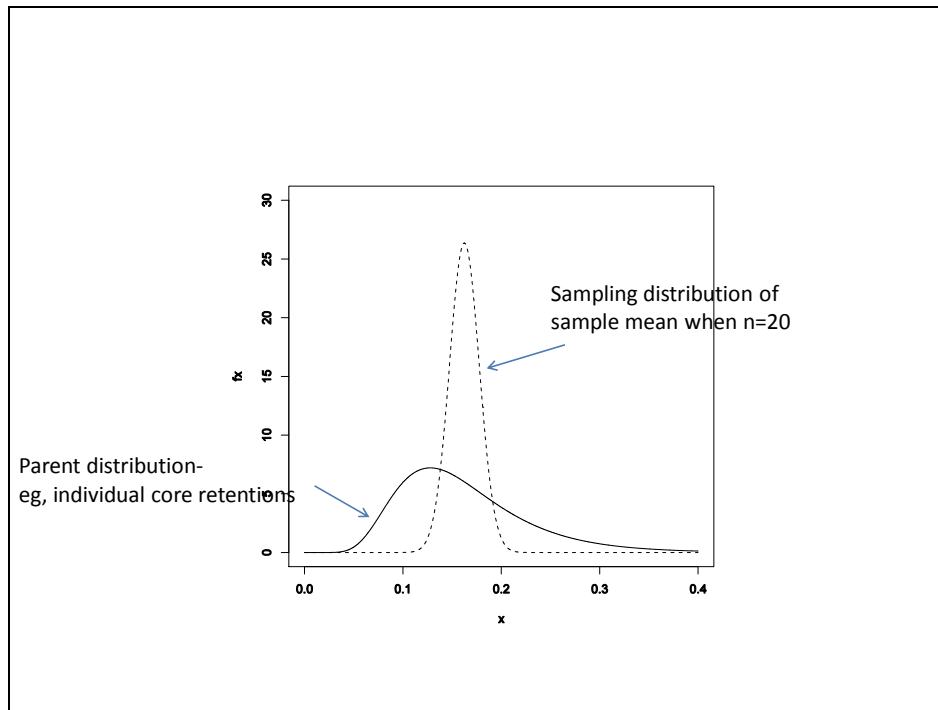


Figure 2. The sampling distribution of the sample mean when the underlying parent distribution is lognormal and when samples of size 20 are drawn randomly from the parent.

In the context of relating this back to obtaining a single retention value from a composite of assay cores, the compositing of the cores (into a homogeneous sample) under the appropriate conditions is the *physical* averaging of measurements. These assumptions include randomization, independence, equal weighting and perfect mixing for compositing and no measurement error. (See Patil, et al. (2011) and Schilling and Neubauer (2009) for specifics.)

AMERICAN WOOD PROTECTION ASSOCIATION

COMPOSITING

Compositing, the combining of cores into a homogeneous mixture from which a single retention value is obtained, is a tool that can reduce within sample variation, saving time and money by not needing to measure each core separately, and providing more material from which to make a measurement. There are, however, disadvantages associated with sampling in this manner, including increased complexity (i.e., estimated variances, when available, may correspond to something different than what we assume), increased reliance on assumptions, and possible measurement errors resulting in increasing variances.

Under the appropriate conditions compositing leads to similar results as the CLT for estimating a mean, however, knowledge about variation may be lost if only a few composite samples are used and knowledge is lost if only a single composite sample is obtained. To further interpret and understand the variation associated with composite measurements when only a single value is obtained, it is necessary to use other experiments or data to understand *possible* variation. This approach carries the risk that the variation used is not comparable, and inappropriate inferences may be made.

In reality, there are many levels of possible grouping in a treatment charge that could lead to different sources of variation which could be components of individual core retentions. For example, there may be retention variation between cores within a board, between boards in a bundle, between bundles in a lot, and between sampling time periods for a lot. An experiment to study assay zone retention variation within boards versus between boards for a set of homogenous southern yellow pine (SYP) 2x4's (actual 38 by 89 mm) treated to a retention of 0.15 lb/ft³ (2.4 kg/m³) was performed by several industrial engineers. Further details can be found in Conklin (2009). A subset of the experiment, in which the outer 0.5 in. (13 mm) assay zone from ten boards treated in one charge, was further analyzed for this paper. The data is plotted in Figure 3.

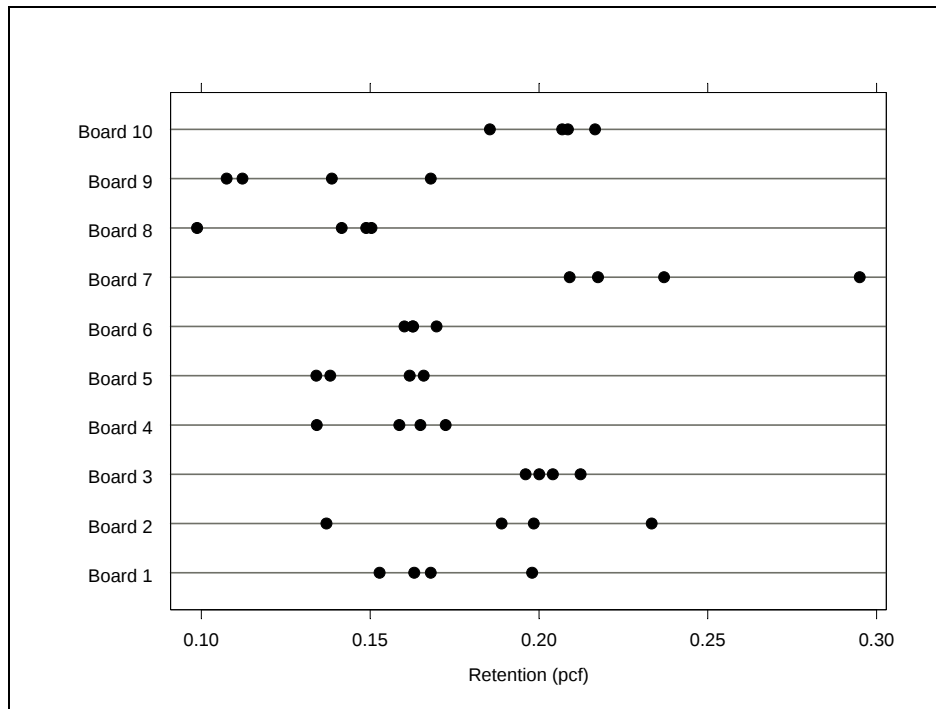


Figure 3. Data from an experiment to study the outer 0.5 in. (13 mm) assay zone retention.

AMERICAN WOOD PROTECTION ASSOCIATION

This subset of data was analyzed as a nested variance component model allowing the breakdown of total variation into within board and between board variation (time was a factor in the study, but our analysis is restricted to a subset of the experiment where no time effect or board by time interaction was indicated). Each of the measurements for a board is a composite of 3 cores. Assuming normality in the underlying populations, we can calculate the expected variation associated with a single core measurement as

$$\sigma^2_{\text{between boards}} + \sigma^2_{\text{between cores within board}} = \sigma^2_{\text{between boards}} + 3 * \sigma^2_{\text{between composites within board}},$$

which is estimated as $0.001040 + 3 * 0.000561 = 0.002723$. At the mean for this subset of data, the relative standard deviation is calculated as $(0.002723)^{1/2} / 0.174 = 0.30$ or about 30%. The expected variation associated with the overall mean retention in this part of the experiment is

$$\sigma^2_{\bar{X}} = \frac{\sigma^2_{\text{between boards}}}{b} + \frac{\sigma^2_{\text{between composites within boards}}}{b * c}$$

where b = number of boards sampled and c = number of composites per board. Once variance estimates are obtained, different sampling scenarios can be investigated to see their impact on estimating the mean retention. For example, with only a single core from each board, the expected variation in the overall mean retention becomes

$$\sigma^2_{\bar{X}} = \frac{\sigma^2_{\text{between boards}}}{b} + \frac{\sigma^2_{\text{between cores within boards}}}{b}$$

If there were 20 boards treated with a mean retention of 0.15 lb/ft³ (2.4 kg/m³) and a single core extracted from each, the relative standard deviation for the mean is then approximated as $((0.001040 + 3 * 0.000561) / 20)^{1/2} / 0.15 = 0.078$, or about 8%. Keep in mind these estimates are from a designed experiment, controlling other factors that could contribute to variation in a processing environment, and, hence, likely provide an optimal scenario. In fact, complete analysis of this full experiment, shows further variation can be attributed to the other assay zones (more interior), as well as a possible initial time component. A companion experiment with a different charge schedule also experienced increased relative standard deviation estimates (data are not shown). Further experiments with multiple similar charges over multiple times and measuring assay retention at the core level could provide more robust estimates.

In the treating plant, if we take a random sample of boards from a charge, extract a core randomly from each board, and then combine the assay zones of the cores into a single homogeneous composite for measurement, we should obtain an estimate of a mean retention value for the appropriate assay zone. There will not be a measurement of variation associated with that value (see Chapter 9 in Schilling and Neubauer, 2009, for specific details). In general, the variation of the lot mean will be similar to what we observed for the experiment discussed above, except we would expect separate variance components for each level of sampling (and possibly a finite population correction factor) as described earlier.

Kleinknecht (1999) discussed how long term treating plant and agency testing data can be broken down to study variation in retention assays. He uses paired treating plant data, with estimated mean retention values for charges obtained by both the treating plants and a third-party agency, to separate the variation into that attributed to the charges, plants and agencies. The variance estimates include sampling plus lab variation for the plants, sampling plus lab variation for the agencies, and charge-to-charge variation. We have computed variance estimates for two sets of long-term data (SYP) and presented the data in terms of relative standard deviations (see Table 2, charge-to-charge variation is not listed). Variance estimates for the two data sets are similar to what Kleinknecht calculated. The relative standard deviation values for the 'combined' column of data are based on three separate estimators where variation is assumed to be common to the plant and the agency. Two estimators from Bland and Altman (1996) and Bland (2006) measure within charge coefficient of variations, and the third is computed from variance estimates from a nested random effect model and using a within-charge relative deviation measurement. Table 2 shows that generally relative standard deviations of the mean retention values are approximately 8-20% based on composites of 20 cores. If we attribute this variation solely to inherent sampling variability (which would not be accurate), this would imply individual core retentions could have possible relative standard deviations in the range 35-90%. Further study of the relationships between treater and agency statistics are possible, especially following some of the work of Bland and Altman for within subject variation measures.

AMERICAN WOOD PROTECTION ASSOCIATION

Table 2. Relative standard deviation (RSD) estimates based on long-term plant/agency assay retention values.

	Data Set A RSD (%)			Data Set B RSD (%)		
Target Retention (pcf)	Plant	Agency	Combined	Plant	Agency	Combined
0.15	16%	14%	13-16%	14%	16%	14-15%
0.25	13%	18%	14-16%	17%	18%	15-18%
0.34	9%	15%	12-13%	12%	15%	14-16%
0.40	8%	18%	14-15%	12%	18%	15-16%
0.60	6%	19%	14-22%	17%	18%	19-22%

CONCLUSIONS

Compositing can be a time and cost saving strategy for sampling in an industrial environment, however, it should be clear what is being gained (better mean estimate with less time and cost) and what is being lost (no variance estimate if only one composite is made). Compositing can occur at many different levels and it is important to understand where it is occurring and make sure that proper compositing procedures are followed to ensure proper inferences. See Schilling and Neubauer (2009) for details on breaking down variance components in bulk sampling in an industrial setting, as well as cost considerations and development of sampling plans.

LITERATURE CITED

1. American Wood Protection Association. 2010. 2010 AWPB Book of Standards. AWPB. Birmingham, AL.
2. Barrentine, L.B. 2003. *Concepts for R&R Studies*. ASQ Quality Press. Milwaukee, WI. 76 pp.
3. Bland, J.M., Altman, D.G. 1996. Measurement error proportional to the mean. *BMJ*. 313:106.
4. Bland, J.M. 2006. How should I calculate a within-subject coefficient of variation? <http://www-users.york.ac.uk/~mb55/meas/cv.htm> (last accessed 5/20/11).
5. Brown, B.W., Hollander, M. 1977. *Statistics: A Biomedical Introduction*. John Wiley & Sons. New York. 456 pp.
6. Conklin, S.W. 2009. Research Needs in Wood Preservation—Thoughts on In-plant Quality Control. CWPA.
7. Kleinknecht, D. 1999. The Influence of Variability in Treated Wood Samples on “Accept/Reject” Decisions. Proceedings of the 95th Annual Meeting of the American Wood-Preservers’ Association. May 16-19, 1999. Fort Lauderdale, FL. 95:105-115.
8. Patil, G.P., Gore, S.D., Taillie, C. 2011. *Composite Sampling: A Novel Method to Accomplish Observational Economy in Environmental Studies*. Springer. New York, NY. 275 pp.
9. Schilling, E.G., Neubauer, D.V. 2009. *Acceptance Sampling in Quality Control. Second Edition*. CRC Press. Boca Raton, FL. 683 pp.
10. Tsoumis, G. 1991. *Science and Technology of Wood: Structure, Properties, Utilization*. Van Nostrand Reinhold. New York. 494 pp.
11. Schultz, T.P., Nicholas, D.D., Dalton, T.J., Keefe, D. 2004. Biocide retention variation of southern yellow pine products treated with waterborne preservatives in commercial, pilot plant, or laboratory cylinders. *Forest Products Journal*. 54(3):85-90.

PROCEEDINGS

One Hundred Seventh Annual Meeting

of the

**AMERICAN
WOOD PROTECTION
ASSOCIATION**

Marriott Harbor Beach Hotel
Fort Lauderdale, Florida
May 15-17, 2011

VOLUME 107

AMERICAN WOOD PROTECTION ASSOCIATION
P O BOX 361784 • BIRMINGHAM, ALABAMA 35236-1784 • USA

ISSN 0066-1198

COPYRIGHT, 2011

BY

AMERICAN WOOD PROTECTION ASSOCIATION
P.O. BOX 361784
BIRMINGHAM, ALABAMA 35236-1784
USA

Rights to republish papers and reports published herein, in whole or in part, or by reference are granted to all persons, provided that reference to the authors and to the AWPA Proceedings are made.

Statements made or opinions expressed in this publication shall not be the responsibility of the American Wood Protection Association.

Colin McCown, *Editor*
Beth Williams, *Assistant Editor*