United States Department of Agriculture

Forest Service

Forest Products Laboratory

Research Paper FPL–RP–576

Monitoring of Visually Graded Structural Lumber

David E. Kretschmann
James W. Evans
Linda Brown

Abstract

To satisfy the increased demand for forest products, much of the future timber supply is expected to be derived from improved trees grown on managed plantations. This fast-grown resource will tend to be harvested in short-age rotations and will contain higher proportions of juvenile wood compared with wood in current harvests. As a result, current allowable properties may need to be reduced in the future. This report explores four options for monitoring the properties of fast-grown wood and briefly discusses the advantages and disadvantages of these approaches. The recommended multiple-stage sampling approach is illustrated in detail using simulated results based on the North American In-Grade test results for Southern Pine. Finally, the report presents details of a "real world" example of monitoring lumber properties currently being conducted by the Southern Pine Inspection Bureau.

Keywords: monitoring, visually graded structural lumber, Southern Pine

July 1999

Kretschmann, David E.; Evans, James W.; Brown, Linda. 1999. Monitoring of visually graded structural lumber. Res. Pap. FPL–RP–576. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 18 p.

A limited number of free copies of this publication are available to the public from the Forest Products Laboratory, One Gifford Pinchot Drive, Madison, WI 53705–2398. Laboratory publications are sent to hundreds of libraries in the United States and elsewhere.

The Forest Products Laboratory is maintained in cooperation with the University of Wisconsin.

Research Highlights

Previous research has helped draw attention to a potential problem associated with fast-grown plantation material: the higher percentage of juvenile wood. The lower mechanical properties of juvenile wood, compared with mature wood, account for the inferior properties of aggressively grown plantation wood harvested on short rotations compared with the properties of longer rotation, naturally suppressed timber. Independent grading agencies have a strong interest in determining whether a significant change (particularly a decrease in material properties) in the lumber resource has occurred. The work reported here explores four options for monitoring lumber properties and briefly discusses the advantages and disadvantages of these options. The recommended multiple-stage approach is discussed in detail using simulated results based on the North American In-Grade test data for Southern Pine.

Some possible choices for monitoring the lumber resource are the use of control charts, Bayesian statistics, repetition of an In-Grade program, and multiple-stage sampling. Of these choices, a multiple-stage sampling approach has been deemed the most practical for visually graded lumber. Simulations demonstrate the relative sensitivity of this method to sample size, juvenile wood content, and selected “trigger levels.” The term trigger level is used to describe the property value level associated with a targeted shift in property. The Southern Pine Inspection Bureau has been monitoring modulus of elasticity since 1994. The results of this program are presented here.

Contents

Monitoring of Visually Graded Structural Lumber

David E. Kretschmann, Research General Engineer
James W. Evans, Supervisory Mathematical Statistician
Forest Products Laboratory, Madison, Wisconsin

Linda Brown, Engineer
Southern Pine Inspection Bureau, Pensacola, Florida

Introduction

Early research drew attention to the potential for lower mechanical properties in plantation material. The study involved an 8-year-old plantation of Caribbean pine (*Pinus caribaea* Morelet) from Puerto Rico (Boone and Chudnoff 1972). Specific gravity, bending stiffness, and strength of this plantation-grown wood were 50% lower than that of published values for virgin lumber of the same species. In actuality, Boone and Chudnoff compared juvenile wood to mature wood rather than plantation wood to virgin lumber. In this study, differences between juvenile and mature wood were probably accentuated because the trees were very young. Thus, the study evaluated early-formed juvenile wood, which has significantly lower properties than later-formed juvenile wood.

Research conducted at North Carolina State University and the Forest Products Laboratory (FPL) on clear wood of loblolly pine demonstrated that the problem of lower properties was the result of juvenile wood, not plantation wood per se (Pearson and Gilmore 1971, Bendtsen and Senft 1986). This research demonstrated that juvenile wood is substantially lower in mechanical properties than is mature wood and generally accounts for the inferior properties of plantation wood compared to that of virgin timber. Later studies further investigated the effect of juvenile wood on full-size structural lumber properties for various species. In New Zealand, in-grade testing was completed on radiata pine lumber cut from 40- to 60-year-old stands (Walford 1982) and 28-year-old stands (Bier and Collins 1984). In Canada, work by Barrett and Kellogg (1989) and Smith and others (1991) looked at plantation Douglas-fir and red pine. In the United States, several studies were conducted on the bending and tension parallel-to-grain properties of Douglas Fir and Southern Pine dimension lumber cut from plantations (Bendtsen and others 1988, Biblis 1990, Clark and Saucier 1989, Kretschmann and Bendtsen 1992, MacPeak and others 1990, Pearson 1984, Ying and others 1994).

Independent grading agencies have a strong interest in determining whether a significant change in the resource has occurred, resulting in a decrease in strength properties. Section 13, "Reassessment and Affirmation," was included in American Society for Testing and Materials (ASTM) Standard D1990 to ensure that a procedure was developed to detect significant changes in allowable properties as a result of change in the raw material resource or product mix (ASTM 1997). Satisfying Section 13 requires a method for detecting a significant change in the resource. In 1994, the Southern Pine Inspection Bureau (SPIB), with the assistance of the FPL, initiated a resource monitoring program to track possible changes that may occur in modulus of elasticity of No. 2 dimensional (2×4) lumber (SPIB 1994) with time. (Note: 2×4 denotes nominal 2- by 4-in. (standard 38- by 89-mm) lumber.)

As part of the background for this program, the FPL initiated a study to investigate options for monitoring properties of structural lumber. The objective of this study was to answer a series of questions, which will help grading agencies make informed decisions on a monitoring program. These questions included the following:

- Is one property, such as modulus of elasticity (MOE), a better predictor of true changes or shifts in the resource than are other properties, such as modulus of rupture (MOR), or the presence of pith?
- What are the relationships between shifts in MOE and MOR?
- How much of a shift in MOE can occur before there is a significant shift in MOR?
- How much of a shift in MOR can occur before there is a significant shift in MOE?
- If monitoring MOE, how does the probability of correctly detecting no shift in MOE increase with each additional sampling for MOE?
- What sample size is required: 50, 100, 200, 400, or 1,000?

The FPL conducted simulations that provide information to assist in understanding the meaning of shifts in MOE of structural lumber over time.

Background—Approaches to Monitoring

Of the many possible choices for monitoring structural lumber properties, this report discusses four approaches: control charts, Bayesian statistics, full In-Grade testing, and multiple sampling.

Control Charts

Control charts have the advantage of being a well-recognized method of quality control with a well-established procedure for tracking changes in properties. Control charts are graphical displays of summary statistics of a process plotted over time. Trends in the data are examined to see if the process is "in control" or "out of control." Through the use of control chart theory, it is possible to set up control limits (boundaries on performance properties) that the results of routine production must meet (Montgomery 1997). If the sampled properties are outside the control limits or there is a persistent pattern of the monitored properties to be consistently above or below the production mean, the operation is deemed out of control. To apply the method correctly requires very intensive repeated sampling at specific mills, which, depending on selection of mills, might ultimately be insensitive to changes in the overall resource. If the method is used incorrectly, there is a high risk that it will mistakenly indicate that allowable properties should be reassessed even when this is not needed.

Bayesian Statistical Approach

The Bayesian statistical method of sampling has been pursued by FPL researchers David W. Green and James W. Evans and University of Wisconsin–Madison statistician Richard Johnson (Johnson and others, 1995, 1999). In this approach, the distributional data of an initial sample is used to weight subsequent samples to obtain global properties. The benefit of the Bayesian method is that it uses existing information to provide guidance on sampling. This can potentially result in a smaller required sample size. The Bayesian approach could be used for a single sample or in a three-stage approach similar to that outlined for multiple-stage sampling. However, because the Bayesian approach is less traditional than other approaches, it would require additional sensitivity studies to confirm its applicability. In addition, adoption by a consensus organization might be difficult to obtain.

Full In-Grade Program

Another sampling option is to repeat the full In-Grade program periodically (Green and others 1989), an approach that has the advantage of being an accepted procedure. However, this program would merely indicate whether a change in design values had occurred. The long time-frame needed for in-grade testing also reduces the usefulness of the sequential nature of data collection. Decades would be needed to acquire an indication of trends. Furthermore, since the original In-Grade program was very expensive, any further full-scale testing program is likely to be even more costly. Finally, this approach does not indicate whether repeated testing is necessary until all testing has been completed. Thus, if no change in properties had indeed occurred, the expense of conducting a full In-Grade program would have been unnecessary.

Multiple-Stage Sampling

The method selected by SPIB is a multiple-stage sample across the growth region. With this method, the detection of a change in resource is approached in stages. The advantages of multiple-stage sampling are that it requires smaller sampling in the initial stages, it can be adjusted for cost and risk, and it focuses resources on the grade–size combination expected to be most sensitive to changes. The disadvantages of this method are that considerable time is required for sampling, consensus must be reached on an acceptable boundary for significant changes and number of stages, and multiple stages are needed to confirm significant changes in the resource.

Figure 1 shows an example of a potential system for the multiple-stage approach. In Stage I, a nondestructive test program is conducted on what is anticipated to be the most sensitive grade and size combination. This stage may have multiple steps to further decrease sample size. Steps are defined as repeated sampling on a regular or periodic basis. Stage II is reached if the most sensitive grade–size combination indicated that a change in properties had occurred. Change is defined as a targeted shift of *x* amount in a property. The term trigger level describes the property value level associated with the targeted shift. The trigger level is defined as the original property value of the sample minus the targeted shift amount.

In Stage II, additional tests (destructive or nondestructive) are conducted on one or more sizes and grades to confirm that the trigger level has been reached. If the targeted shift is not confirmed, the original periodic testing of Stage I is reinitiated. If the targeted shift is confirmed in Stage II, a further stage of destructive testing in the full range of grade–size combinations may be initiated.

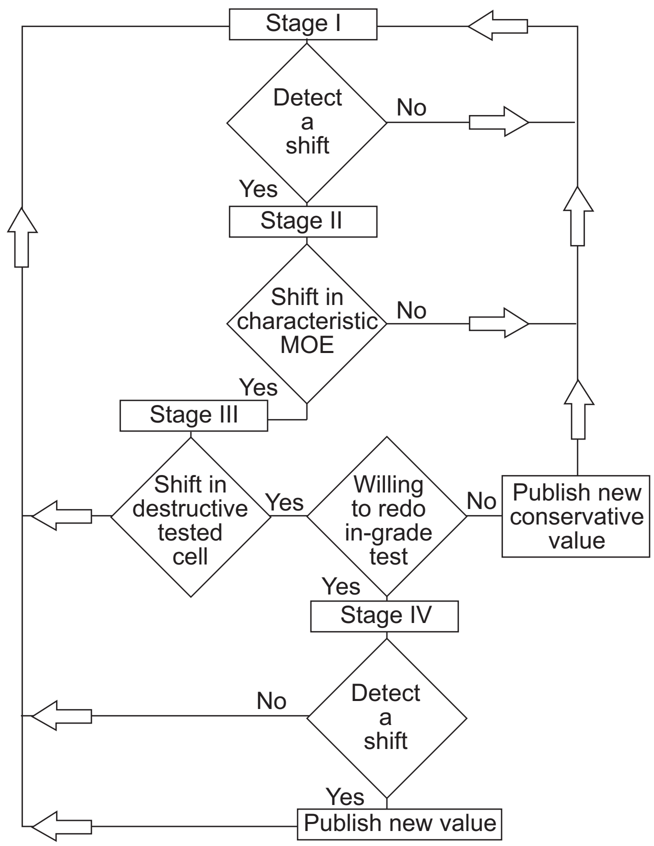

Figure 1—Potential flowchart for monitoring program.

Questions for Multiple-Stage Approach

In a multiple-stage approach, three initial questions must be answered:

1. Which material is most sensitive to changes in the resource; that is, what grade–size combination would be first to reflect a change in the resource?

2. How much of a change in that sensitive material is significant?

3. What are the economic impacts of reaching Stage II and beyond?

Choosing the most sensitive grade(s) and size(s) reduces the amount of material that must be tested. It is also assumed to be the most efficient way of detecting a change in the resource. Which grade(s) and size(s) are most sensitive to changes in the resource is judged by agency personnel on the basis of past experience and available test data.

The determination of a significant change is preferably made by consensus, but it can also be made by individual agencies. Pertinent questions to consider include the following:

- Is the agency most concerned with detecting a shift that will affect a rounding rule or with detecting a shift of a particular absolute value?

- How much confidence is needed that a change has really occurred and therefore how powerful does the test of change have to be? (Power is defined as the probability that if there is a difference between two distribution means, it will be detected.)

- How does the selection of power affect the sample size tested?

- How many samples is it practical to test?

The economic impact of progressing beyond Stage I is considerable. Progressing to Stage II and beyond will require testing a large number of pieces. Consequently, it is critical to determine that a change in the resource has occurred with a considerable degree of certainty.

Basic statistical procedures may be used to help answer these questions. Assuming normality, a quick estimate for the sample size required to detect differences in two means of two normal distribution means can be calculated. The agency must know its acceptable “comfort level” for detecting a difference in means when there is none (α) and for not detecting a difference in means when there is one (β). Alpha (α) is the probability that a difference will be detected between the two means when none exists. Ideally, α should be minimized. For the probability that a difference between the two means will not be detected even though one exists, the power is $1-\beta$. Ideally, this power should be maximized, so β should be small. Equation (1) rearranges the definition of standard variate at the chosen α level in terms of the mean of the sample tested. Equation (2) rearranges the definition of the standard variate at the chosen β level in terms of the mean for the sample tested. Equating Equations (1) and (2) and solving for *n* results in a quick estimate of sample size *n* (Eq. (3)) required to detect differences in two means of two populations, where the α and β levels and the standard deviation *s* for the population are known.

$$\frac{\bar{x}-\mu_0}{s/\sqrt{n}} = Z_\alpha \Rightarrow \bar{x} = \mu_0 + Z_\alpha\, s/\sqrt{n} \qquad (1)$$

$$\frac{\bar{x}-\mu_1}{s/\sqrt{n}} = -Z_\beta \Rightarrow \bar{x} = \mu_1 + Z_\beta\, s/\sqrt{n} \qquad (2)$$

where

Z_α is standard normal variate at chosen α level,

Z_β standard normal variate for chosen β level for second sample,

n	sample size,
s	common standard deviation of two populations,
μ_0	average value of original population,
μ_1	average value of new sample to be detected (trigger level),
$\bar{x}$	value of sample average above which no difference is declared, and
$\mu_0 - \mu_1$	magnitude of change to be observed (targeted shift).

Setting Equation (1) = Equation (2) gives sample size n:

$$n = \left(\frac{(-Z_\alpha + Z_\beta)s}{(\mu_0 - \mu_1)} \right)^2 \quad (3)$$

Figure 2 illustrates the effect of selecting different α and β levels. For example, if the agency is monitoring MOE, it may want to limit the chance of falsely detecting a change in MOE to 5% ($\alpha = 0.05$, $Z_\alpha = -1.645$) and to set the probability of proceeding to Stage II to 50% if there is a change in MOE ($\beta = 0.5$, $Z_\beta = 0$) (Fig. 2a). This is quite different from setting criteria so that there is 90% probability of detecting that MOE has shifted (Fig. 2b). Parts (c) and (d) of Figure 2 provide examples of 95% and 99% assurance, respectively, that MOE has shifted.

There are further complications, however, to choosing a particular α or β level on MOE for monitoring structural properties. The true concern is not only whether MOE has changed, but also how this change affects other mechanical properties. Existing information can be used to make assumptions about the relationship between other properties and MOE. However, it is preferable to run a number of simulations, making use of real strength and stiffness data, to provide a clearer understanding of the potential variability and sensitivity of a multiple-stage monitoring program. The rest of this report provides simulated results based on the North American In-Grade test data for Southern Pine and the specific example of the monitoring program initiated by SPIB. The simulated results are dependent on the MOE–MOR relationship observed during the In-Grade test program. If this relationship itself is altered, the simulation results may not be valid.

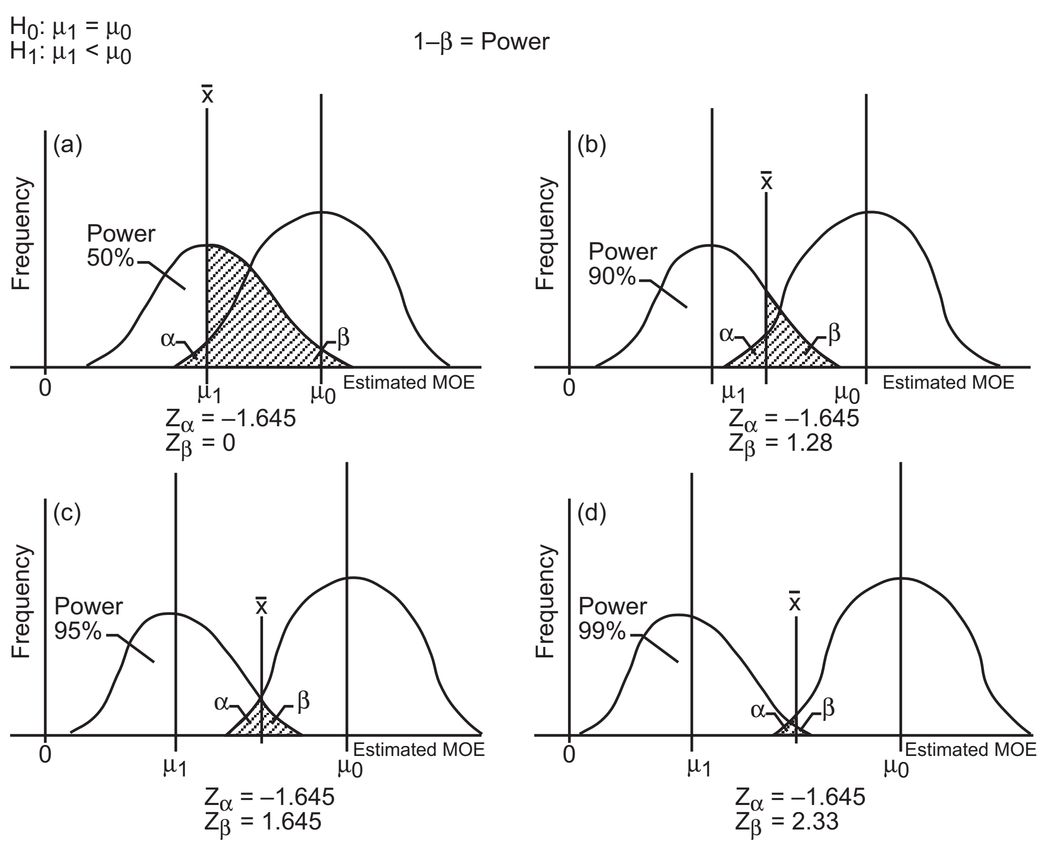

Figure 2—Differences in selection of α and β levels.

Simulations

Data

The data used in the simulations were the SPIB In-Grade test results (Green and Evans 1987). The test data were adjusted using the procedures outlined in ASTM D1990 that are currently used to calculate the allowable properties listed in the American Forest & Paper Association (AF&PA) National Design Specification (NDS) for Wood Construction (AF&PA 1997).

Methods

Simulations were conducted in two phases. In Phase 1, we examined the impact of different sample sizes on the ability to predict changes in properties. In Phase 2, we investigated the sensitivity of MOE as a monitoring parameter by simulating potential scenarios of future resource changes.

Phase 1

Phase 1 investigated the likelihood of falsely predicting targeted shifts in properties when no change had occurred in the data set for a range of sample sizes. A program was developed to randomly select monitoring samples from the existing In-Grade data sets for bending with no changes. The properties considered were MOE and MOR. All trials looked at data randomly selected from either the entire data set or within each geographic region, weighting by production percentages. The use of closed-form solutions was also investigated.

Phase 2

Phase 2 addressed the sensitivity of the monitoring technique to detect predetermined changes in strength properties (trigger levels) with one to three steps. Targeted shifts in MOE were investigated. A targeted shift is a change of a specific amount *x* in MOE. The corresponding values of MOR were adjusted in accordance with MOE–MOR regression equations determined from the original In-Grade Southern Pine data.

In a number of simulations, regression equations were used to evaluate possible shifts of concern. Required sample sizes to detect a given change in MOE were investigated. Also, possible shifts in MOE caused by increases in juvenile wood were simulated.

Results

Regressions

Data from several sizes of No. 2 lumber and several grades of 2×4 lumber were regressed to establish the relationship between MOR and MOE. Regression relationships are shown in Table 1. These relationships were used to establish the anticipated required change in MOE (ΔMOE = $x_1 - x_2$) associated with change of a certain amount in MOR. For example, for a change of one rounding rule in MOR (50 lb/in^2 (344 kPa) as defined in ASTM D1990), the numeric value of the change in MOR (ΔMOR = $y_1 - y_2$) would equal the safety-load duration factor (2.1) multiplied by the rounding rule or 2.1·50 lb/in^2 =105 lb/in^2 (724 kPa).

Using $\quad \mathrm{MOR} = b \cdot \mathrm{MOE} + a$

and $\quad (y_1 - y_2) = b(x_1 - x_2)$

the corresponding $\Delta\mathrm{MOE} \approx 0.02 \times 10^6$ lb/in^2 (0.14 GPa).

Table 1—Regression relationships for various grade–size combinations[a]

Lumber size[b]	Grade	Samples with pith (no.)	Samples without pith (no.)	Intercept (*a*)	Slope (*b*)	One rounding rule in MOR is ΔMOE (×10^6 lb/in^2	(GPa))
2×4	No. 1	15	91	0.306	5.422	0.019	(0.13)
	No. 2	91	320	–0.864	5.767	0.018	(0.12)
	Select Structural	49	360	–1.329	5.294	0.020	(0.14)
	Stud	62	117	–0.521	5.229	0.020	(0.14)
2×8	No. 2	166	592	–0.151	4.016	0.026	(0.18)
2×4	No. 2 juvenile[c]	48	—	0.559	4.457	0.024	(0.17)
2×4	No. 2 juvenile[d]	106	—	0.711	4.706	0.022	(0.15)

[a]MOE and MOR values adjusted to 15% moisture content were used for calculations (Green and Evans 1987).
[b]Note: 2×8 denotes nominal 2- by 8-in. (standard 38- by 184-mm) lumber.
[c]Juvenile defined as presence of pith *and* less than 8 rings/inch (approximately 3 rings/cm).
[d]Juvenile defined as presence of pith *or* less than 4 rings/inch (approximately 1.5 rings/cm).

There is considerable variability about the MOE–MOR regression relationships for lumber but a well-established correlation (Green and Kretschmann 1991). For the sizes and grades considered, one rounding rule in MOR corresponds to, on average, a change in MOE of approximately 0.02×10^6 lb/in^2 (0.14 GPa). Using defined juvenile wood data sets, one rounding rule in MOR corresponds to a change in MOE of 0.025×10^6 lb/in^2 (0.17 GPa). These changes in MOE are much smaller than the rounding rule change of 0.1×10^6 lb/in^2 (0.69 GPa) in ASTM D1990.

Sample Size

The effect of sample size on the percentage of false positive readings, for one rounding rule in MOE and MOR (Table 2) and three targeted shifts in MOE (0.1, 0.025, and 0.02, Tables 3 to 5, respectively), was investigated using 100,000 simulated samples for each case. For each sample, a failure was considered to have occurred when the average MOE or the 5th percentile MOR for the sample was below the respective trigger level for the property. The MOE trigger level was the original mean MOE minus the targeted shift of one rounding rule of 0.1×10^6 lb/in^2 (0.69 GPa). For a sample size of 50 and a change in MOE of 0.1×10^6 lb/in^2 (0.69 GPa), when two successive sample failures are required there is a 0.033% chance of falsely declaring a shift in MOE (Table 2). The MOR trigger level was the original MOR 5th percentile minus 50 lb/in^2 (344 kPa). For the same example, there is a 23% chance of a false indication of a shift in MOR.

In Table 2, a step represents one complete sample of material. In the example just described, one step represents the survey of data from 1 year. Step 2 represents sampling from the next year and so on. For the MOE and MOR data in columns 3 and 4, "overall" designates samples selected randomly, with replacement, from all of the original 413 2×4 Select Structural (SS) pieces until the test sample was obtained. The number of No. 2 2×4 pieces tested in the In-Grade program was 413. The second run (columns 5 and 6) was a repeat of the original overall simulation. "Regional selection" designates material sampled randomly from within a given region, proportional to production in that region, until the 413-piece sample was reached.

The results shown in Table 2 suggested that the false positive rates for the two overall simulations were not substantially different than those for the regional selection. Therefore, the rest of the tests were run using random selection from the overall database.

Tables 3 to 5 compare the effect of sample size on the chances of falsely detecting a change in MOE for three levels of change in MOE (0.1, 0.025, and 0.020×10^6 lb/in^2 (0.69 GPa, 0.17 GPa, 0.14 GPa), respectively). If multiple steps are used, these results suggest little difference between sample sizes of 1000, 413, or 200. If the sample size is below 200, there is a significant increase in the chance of false positive readings.

The results of these simulations demonstrate the importance of sampling method in regard to the sample size required. If an agency were dependent on the information gathered over only 1 year (that is, one step), then the sample size required to reduce the chance of detecting false positive below 1% for a change in MOE of 0.025 would usually be well over 1,000 specimens (Table 4). This same level of confidence can be reached with slightly more than 200 specimens after three steps have been taken. For a random sample size, repeated sampling (multiple-stage approach) helps to ensure that a detected shift is not a result of the natural variability of the random sample by reducing the chances of falsely detecting a property shift. The original In-Grade grade–size sample of 413 pieces is very unlikely to give a false positive indication of a shift for MOE in a multiple-step process. The sample size simulation results in Tables 2 to 5 also suggest that detection with three steps and a sample size of 200 is roughly equivalent to a two-step procedure with a sample size of 413.

Table 2—Probability of falsely detecting change of one rounding rule[a] when no change had occurred

		Probability (%)					
		Overall		Second run of overall		Regional selection	
Sample size	Step	MOE	MOR	MOE	MOE	MOR	MOR
50	1	1.76	48.0	1.75	47.8	1.31	47.7
	2	0.033	23.0	0.41	22.9	0.016	22.8
200	1	0	34.4	0.001	34.3	0.001	31.9
	2	0	11.9	0	11.8	0	10.2
413	1	0	27.1	0	26.9	0	26.9
	2	0	7.3	0	7.2	0	7.3
1,000	1	0	17.5	0	17.6	0	14.5
	2	0	3.1	0	3.1	0	2.1

[a]Rounding rule defined as 0.1×10^6 lb/in^2 (0.69 GPa) in MOE or 50 lb/in^2 (344 kPa) in MOR.

Table 3—Probability of falsely detecting 0.1 $\times$ 10^6 lb/in^2 (0.69 GPa) change in MOE when no change had occurred

Sample size	Step	Probability (%) 2x4 No. 2	2x4 SS[a]	2x4 Stud	2x8 No. 2
50	1	2.964	1.665	3.223	4.377
	2	0.092	0.019	0.088	0.193
	3	0.005	0	0.004	0.009
100	1	0.381	0.120	0.411	0.746
	2	0.004	0	0.001	0.010
	3	0	0	0	0
200	1	0.009	0.001	0.015	0.036
	2	0	0	0	0
	3	0	0	0	0
413	1	0	0	0	0.001
	2	0	0	0	0
	3	0	0	0	0
1000	1	0	0	0	0
	2	0	0	0	0
	3	0	0	0	0

[a]SS is Select Structural.

Table 4— Probability of falsely detecting 0.025 $\times$ 10^6 lb/in^2 (0.17 GPa) change in MOE when no change had occurred for various sizes and grades (%)

Sample size	Step	Probability (%) 2x4 No. 2	2x4 SS	2x4 Stud	2x8 No. 2
50	1	33.571	29.481	32.690	34.155
	2	11.235	8.811	10.725	11.679
	3	3.754	2.722	3.517	4.032
100	1	27.369	22.267	26.122	27.877
	2	7.411	4.875	6.893	7.839
	3	2.028	1.095	1.797	2.182
200	1	19.788	13.972	18.173	20.234
	2	3.932	1.911	3.271	4.033
	3	0.776	0.286	0.604	0.785
413	1	11.087	5.990	9.420	11.359
	2	1.189	0.344	0.877	1.319
	3	0.142	0.016	0.086	0.154
1000	1	2.832	0.741	2.130	3.009
	2	0.051	0.007	0.054	0.088
	3	0	0	0.002	0.003

Table 5—Probability of falsely detecting 0.02 $\times$ 10^6 lb/in^2 (0.14 GPa) change in MOE when no change had occurred

Sample size	Step	Probability (%) 2x4 No. 2	2x4 SS	2x4 Stud	2x8 No. 2
50	1	37.137	33.283	36.123	37.329
	2	13.729	11.256	13.056	13.965
	3	5.028	3.832	4.727	5.291
100	1	32.108	27.065	30.663	32.014
	2	10.176	7.269	9.470	10.310
	3	3.286	1.988	2.928	3.304
200	1	25.476	19.257	23.335	25.458
	2	6.498	3.666	5.448	6.349
	3	1.641	0.704	1.286	1.576
413	1	17.204	10.545	14.689	16.819
	2	2.891	1.082	2.106	2.776
	3	0.480	0.111	0.332	0.435
1000	1	7.148	2.582	5.193	6.630
	2	0.458	0.060	0.293	0.425
	3	0.033	0	0.017	0.020

Therefore, a smaller sample size could be used with more steps to detect a targeted shift in MOE. A sample size of at least 360 should be considered to obtain good representation of all geographic regions. Also, if a shift in MOE of one rounding rule is observed, a shift of several rounding rules may have occurred for MOR.

Juvenile Wood

Increases in the percentage of juvenile wood in a piece were simulated in three ways. In each simulation, the criteria for defining the presence of juvenile wood were presence or absence of pith and number of rings per unit length, with 4 rings/inch equal to approximately 1.5 rings/cm.

Simulation 1	Presence of pith
Simulation 2	Presence of pith *and* <4 rings/inch Absence of pith *or* ≥4 rings/inch
Simulation 3	Presence of pith *or* <4 rings/inch Absence of pith *and* ≥4 rings/inch

The effects of these three sets of sorting criteria on MOE and MOR are illustrated in Tables 6, 7, and 8, respectively. Information on the presence or absence of pith was not available for all of the original In-Grade samples. Therefore, the total number of pieces listed by presence or absence of pith may not equal the actual number of pieces tested.

Table 6—Effect of Simulation 1 sorting criteria on MOE and MOR

Lumber size	Grade	Sorting criteria	Sample size	Mean MOE ($x10^6$ lb/in^2	(GPa))	5th percentile MOR ($x10^3$ lb/in^2	(MPa))
2×4	SS	All	409	1.823	(12.6)	6.987	(48.2)
		With pith	49	1.660	(11.4)	4.996	(34.4)
		Without pith	360	1.845	(12.7)	7.121	(49.1)
2×4	No. 2	All	411	1.534	(10.6)	3.842	(26.5)
		With pith	91	1.334	(9.2)	3.835	(26.4)
		Without pith	320	1.591	(11.0)	3.821	(26.3)
2×4	Stud	All	179	1.487	(10.3)	3.246	(22.4)
		With pith	62	1.356	(9.3)	2.908	(20.1)
		Without pith	117	1.556	(10.7)	3.798	(26.2)
2×8	No. 2	All	758	1.543	(10.6)	2.584	(17.8)
		With pith	166	1.464	(10.1)	2.473	(17.1)
		Without pith	592	1.566	(10.8)	2.588	(17.0)

Table 7—Effect of Simulation 2 sorting criteria on MOE and MOR

Lumber size	Grade	Sorting criteria[a]	Sample size	Mean MOE ($x10^6$ lb/in^2	(GPa))	5th percentile MOR ($x10^3$ lb/in^2	(MPa))
2×4	SS	All	409	1.823	(12.6)	6.987	(48.2)
		With pith *and* <4 rpi	1	1.490	(10.3)	—	(—)
		Without pith *or* ≥4 rpi	408	1.824	(12.6)	6.983	(48.1)
	No. 2	All	411	1.534	(10.6)	3.842	(26.5)
		With pith *and* <4 rpi	5	0.971	(6.7)	3.575	(24.6)
		Without pith *or* ≥4 rpi	406	1.541	(10.6)	3.863	(26.6)
2×4	Stud	All	179	1.487	(10.3)	3.246	(22.4)
		With pith *and* <4 rpi	9	0.967	(6.7)	2.901	(20.0)
		Without pith *or* ≥4 rpi	170	1.515	(10.4)	3.405	(23.5)
2×8	No. 2	All	758[b]	1.543	(10.6)	2.584	(17.8)
		With pith *and* <4 rpi	36	1.153	(7.9)	2.473	(17.1)
		Without pith *or* ≥4 rpi	721	2.584	(17.8)	2.588	(17.8)

[a]rpi is rings per inch; 4 rpi is approximately 1.5 rings/cm.
[b]One piece was eliminated from study as result of lack of information on both pith and rings per unit length.

For Simulation 1, MOR values for No. 2 lumber were not very sensitive to the presence of pith (Table 6). This was probably caused by the exclusion of some pieces with larger knots from the pith group when pieces were selected on the basis of pith alone. However, for all grades, MOE was sensitive to changes in the presence of pith. We concluded that defining juvenile wood on the basis of presence of pith and the number of rings per unit length resulted in too few pieces for a proper simulation. Therefore, for the remaining simulations, presence of pith was used as the criterion for juvenile wood.

Changes in properties were simulated using random sampling with discarding. Increases in the percentage of pieces with juvenile wood were simulated by discarding pieces that did not meet the criteria defining juvenile wood. Thus, for a 20% discard rate, if a piece were drawn that did not meet the juvenile wood definition (that is, a piece without pith), it would be discarded from the sample 20% of the time. Note that this is different from requiring certain percentages of pieces with juvenile wood in the sample. The 20%, 40%, 60%, and 80% discard rate sample sets were established by discarding, with that probability, pieces without pith and replacing those pieces by resampling the original 413-piece sample. The resampled pieces did or did not have pith.

Table 9 lists the random sampling results for SS 2×4 lumber for various properties when the discard rate was increased. An example is given for each property simulated to demonstrate how the boundaries were established for various comparisons of the 100,000 simulation samples.

Table 8—Effect of Simulation 3 sorting criteria on MOE and MOR

Lumber size	Grade	Sorting criteria	Sample size	Mean MOE ($\times 10^6$ lb/in^2	(GPa))	5^{th} percentile MOR ($\times 10^3$ lb/in^2	(MPa))
2×4	SS	All	409	1.823	(12.6)	6.987	(48.2)
		With pith *or* <4 rpi	53	1.654	(11.4)	4.662	(32.1)
		Without pith *and* ≥4 rpi	356	1.848	(12.7)	7.230	(49.9)
2×4	No. 2	All	411	1.514	(10.4)	3.842	(26.5)
		With pith *or* <4 rpi	106	1.314	(9.1)	3.879	(26.7)
		Without pith *and* ≥4 rpi	305	1.541	(10.6)	3.835	(26.4)
2×4	Stud	All	179	1.487	(10.3)	3.246	(22.4)
		With pith *or* <4 rpi	78	1.303	(9.0)	2.893	(19.9)
		Without pith *and* ≥4 rpi	101	1.629	(11.2)	3.835	(26.4)
2×8	No. 2	All	758	1.543	(10.6)	2.584	(17.8)
		With pith *or* <4 rpi	274	1.416	(9.8)	2.391	(16.5)
		Without pith *and* ≥4 rpi	484	1.616	(11.1)	2.630	(18.1)

Table 9—Effect of discard rate on properties and pith[a]

Property	No discard	20% discard	40% discard	60% discard	80% discard
Pith range	25–78	32–95	42–113	68–147	126–218
Pith mean	49	59	76	104	166
MOE mean ($\times 10^6$ lb/in^2 (GPa))	1.824 (12.6)	1.819 (12.5)	1.812 (12.5)	1.799 (12.4)	1.771 (12.2)
MOR 5% ($\times 10^3$ lb/in^2 (MPa))	6.939 (47.8)	6.873 (47.4)	6.769 (46.7)	6.589 (45.4)	6.290 (45.4)

[a]Data from 2×4 SS sample. Sample size = 413. Target MOE = 0.025×10^6 lb/in^2 (0.17 GPa).

Example Simulations

MOE Simulation

In Table 10, the boundary for MOE is the average MOE (1.823×10^6 lb/in^2 (12.6 GPa), Table 6) of the original In-Grade SS 2×4 sample minus 0.025×10^6 lb/in^2 (0.17 GPa) or 1.798×10^6 lb/in^2 (12.4 GPa). For step 1 (the first 413-piece sample), the average MOE was determined and compared against 1.798×10^6 lb/in^2 (12.4 GPa). A failure was considered to have occurred if this average value was less than the boundary MOE; 5,990 of 100,000 samples had averages below the boundary (Table 10). Step 2 represents the number of times that two failures occurred in a row. Thus, if a failure had already occurred and the average MOE of the next sample of 413 pieces fell below the boundary, one piece was added to the total of failures for step 2. If the average MOE of the subsequent sample was above 1.798×10^6 lb/in^2 (12.4 GPa), no step 2 failure had occurred and the comparison was conducted the next time when the average MOE of a 413-piece sample fell below the boundary.

For the 99,999 samples checked, there were 344 occurrences when two samples in a row had an average MOE below the boundary by chance with no discard. The same process was used for three samples in a row. For SS 2×4, there were only 16 occurrences when three samples in a row had an average MOE < 0.025×10^6 lb/in^2 (<0.17 GPa) less than the average MOE of the original sample when no change in population had occurred (that is, no pieces had been discarded). As the discard rate increased (that is, increase in percentage of pieces with juvenile wood), the percentage of failures detected with repeated samplings (steps 1 to 3) increased dramatically.

MOR Simulation

In Table 10, the boundary used for MOR was the safety-load duration factor of 2.1 times one rounding rule ($2.1 \cdot 50 = 105$ lb/in^2 (723 kPa)) subtracted from the 5^{th} percentile (6.987×10^3 lb/in^2 (48.2 MPa)). The actual value of the SS 2×4 MOR boundary was 6.882×10^3 lb/in^2 (47.5 MPa). For each 413-piece sample, the 5^{th} percentile MOR was determined and compared to this boundary. The regression relationship in Table 1 was used to determine what an equivalent change in MOE would be with a rounding rule change in MOR.

Table 10—Percentage of failures in average MOE, MOR, pith, or proof load observed at different steps for various discard rates[a]

		Failures (%)				
Property	Step	No discard	20% discard	40% discard	60% discard	80% discard
MOE[b]	1	5.990	10.337	20.788	48.501	94.892
	2	0.344	1.060	4.290	23.451	90.029
	3	0.016	0.110	0.899	11.385	85.449
MOR[c]	1	35.830	46.116	51.586	82.899	98.664
	2	12.847	21.188	37.864	68.706	97.347
	3	4.564	9.645	23.259	56.932	96.047
Pith[d]	1	5.837	49.221	98.278	100	100
	2	0.371	24.115	96.583	100	100
	3	0.019	11.833	94.915	100	100
Proof load[e]	1	4.800	8.146	15.931	36.270	82.205
	2	0.232	0.676	2.533	13.088	67.489
	3	0.012	0.060	0.434	4.688	55.372

[a]Simulation based on 2×4 SS sample. Sample size = 413.
[b]Target MOE = 1.798×10^6 lb/in^2 (12.4 GPa).
[c]Target MOR = 6.882 lb/in^2 $\times 10^3$ lb/in^2 (47.4 MPa).
[d]Target number of pieces with pith ≥60.
[e]Target order statistic >27th order.

From Table 10, of 100,000 samples, there were 35,830 occurrences when the calculated MOR 5th percentile was below the boundary. For the 99,999 samples checked in step 2, there were 12,847 occurrences when two samples in a row had a calculated MOR 5th percentile below the boundary. Finally, for the 99,998 samples checked in step 3, there were 4,564 occurrences when three samples in a row had a calculated MOR 5th percentile below the boundary.

Pith Simulation

A binomial distribution was used to determine a one-sided upper 95% confidence interval for the expected number of specimens discarded, based on the lack of pith in a 413-piece sample. Equation (4) was used to calculate the bounds of the 95% confidence interval,

$$n\left[p \pm \left(1.645\sqrt{\frac{pq}{n}} + \frac{1}{2n}\right)\right] \tag{4}$$

where p is the probability a specimen has pith, q is $1 - p$, and n is sample size.

The continuity correction factor ($1/2n$) was ignored because the large sample sizes made it negligible. The upper boundary for the confidence interval was used for both pith and proof-load simulations. Using the upper boundary ensured a high level of confidence that the shift had in fact occurred and was not merely observed as a result of chance. We were interested in the side of the interval where there would be too many, rather than too few, pieces with pith or proof-load failures.

In the original In-Grade SS 2×4 sample, 49 of 413 pieces contained pith:

$$p = \frac{49}{413} = 0.119 \qquad q = (1 - p) = 0.881 \qquad n = 413 \tag{5}$$

Equation (5) can be written in terms of pieces containing pith:

$$413\frac{49}{413} + 1.645\sqrt{413 \cdot \frac{49}{413} \cdot \frac{364}{413}} = 49 + 10.81 = 59.81 \tag{6}$$

Therefore, the boundary for pith that should be used is $49 + (0.026 \cdot 413) \cong 60$. In this case, a "failure" was said to occur if ≥60 pieces in the 413-piece sample contained pith. From Table 10, 5,837 of 100,000 samples had ≥60 pieces with pith. However, in only 371 of 99,999 occurrences did two sample sets in a row have ≥60 pieces with pith, and in only 19 of 99,998 occurrences did three sample sets in a row have ≥60 pieces with pith.

Table 11—Effect of discard rate on properties and presence of pith and percentage of failures in No. 2 2×4 sample for various discard rates[a]

		Effect of discard rate				
Pith or property	Step	No discard	20% discard	40% discard	60% discard	80% discard
Pith range	—	56–125	72–150	95–174	131–220	200–286
Pith mean	—	91	108	132	171	242
MOE mean (x10^6 lb/in^2 (GPa))	—	1.530 (10.5)	1.520 (10.5)	1.505 (10.4)	1.482 (10.2)	1.438 (9.9)
MOR 5% (x10^3 lb/in^2 (MPa))	—	3.845 (26.5)	3.845 (26.5)	3.846 (26.5)	3.848 (26.5)	3.851 (26.6)
		Failures (%)				
MOE	1	11.087	25.775	57.339	93.604	99.997
	2	1.189	6.662	32.978	87.616	99.993
	3	0.142	1.674	18.857	82.024	99.989
MOR	1	12.730	12.180	11.100	9.743	7.597
	2	1.640	1.541	1.265	0.926	0.564
	3	0.191	0.207	0.151	0.079	0.038
Pith	1	5.759	64.222	99.859	100	100
	2	0.341	41.371	99.717	100	100
	3	0.015	26.652	99.576	100	100
Proof load	1	4.850	4.682	4.203	3.772	3.142
	2	0.233	0.231	0.193	0.118	0.093
	3	0.011	0.008	0.008	0.003	0.003

[a]Target MOE = 0.025×10^6 lb/in^2 (0.17 GPa).

Proof-Loading Simulation

Proof loading is a process of loading a member to a selected level to obtain "proof" that the member will perform at that load level. For wood, proof loading is usually used to determine which samples pass a 5th percentile criterion for a strength property. To simulate the effect of proof loading, the comparison boundary was determined again by using a binomial function. Using the 5th percentile definition and Equation (4), the upper 95% confidence limit boundary was determined to be anything greater than the 27th order statistic of the MOR of the original 413-piece SS 2×4 sample. Therefore, a "failure" was considered to have occurred if the MOR of ≥28 of 413 pieces was less than that of the targeted proof-load:

$$413 \cdot 0.05 + 1.645\sqrt{413 \cdot 0.05 \cdot 0.95} = 20.65 + 7.286 = 27.9 \tag{7}$$

Tables 11 to 13 present results similar to those presented in Tables 8 to 10 for different grades and sizes and a sample size of 413. The results suggest that proof-load testing for MOR is not a meaningful monitoring scheme when a shift in MOR is linked to a shift in MOE because monitoring for MOE is more powerful. Also, using the presence of pith to detect a change in resource may be overly sensitive to shifts in MOE but may have potential as an early detection mechanism for property shifts. However, pith sensitivity varies from cell to cell.

The results of these simulations show that several steps are needed to limit the probability of detecting a shift in MOE when one does not truly exist. In Tables 6 and 9 and the upper part of Table 11, results on the effect of discard rate on properties and presence of pith show that for a given change in MOE, there may be several rounding rule changes in MOR for higher grade lumber and few changes for low-grade material. If MOR is chosen as the monitoring property, the simulations suggest that shifts in MOR can be detected easily. Monitoring of MOR may also be grade dependent. MOE, on the other hand, seems to be quite grade-independent and does not give as high a rate of false declarations of a change. Therefore, MOE is a better monitoring property than is MOR.

MOE Trigger Levels

The MOE trigger is the amount of a shift in MOE that is acceptable prior to initiating Stage II testing. To determine how the discard rate affected the percentage of samples resulting in a MOE shift of a particular magnitude, we considered MOE shifts of 0.02, 0.04, 0.05, and 0.1×10^6 lb/in^2 (0.14, 0.28, 0.34, and 0.69 GPa, respectively). The results are listed in Tables 14 and 15.

Table 12—Effect of discard rate on properties and presence of pith and percentage of failures in 2×4 Stud sample for various discard rates[a]

		Effect of discard rate				
Pith or property	Step	No discard	20% discard	40% discard	60% discard	80% discard
Pith range		99–182	125–213	149–243	193–279	258–338
Pith mean		143	165	194	235	300
MOE mean (x10^6 lb/in^2 (GPa))		1.487 (10.3)	1.477 (10.2)	1.463 (10.1)	1.442 (9.9)	1.411 (9.7)
MOR 5% (x10^3 lb/in^2 (MPa))		3.245 (22.4)	3.205 (22.1)	3.161 (21.8)	3.106 (21.4)	3.036 (20.9)
		Failures (%)				
MOE	1	9.420	22.259	49.941	86.200	99.790
	2	0.877	4.930	24.994	74.345	99.579
	3	0.086	1.080	12.501	64.124	99.368
MOR	1	30.704	36.401	44.172	55.774	71.189
	2	9.455	13.308	19.672	31.279	50.718
	3	2.871	4.810	8.746	17.551	36.063
Pith	1	5.565	72.856	99.977	100	100
	2	0.323	53.027	99.953	100	100
	3	0.016	38.605	99.929	100	100
Proof load	1	2.078	2.999	4.872	8.735	18.033
	2	0.049	0.088	0.215	0.760	3.254
	3	0.002	0.004	0.008	0.075	0.577

[a]Target MOE = 0.025 × 10^6 lb/in^2 (0.17 GPa).

Table 13—Effect of discard rate on sample properties and presence of pith and percentage of failures in No. 8 2×4 sample for various discard rates[a]

		Effect of discard rate				
Pith or property	Step	No discard	20% discard	40% discard	60% discard	80% discard
Pith range		56–131	70–149	93–171	123–218	201–284
Pith mean		91	107	132	170	241
MOE mean (x10^6 lb/in^2 (GPa)		1.543 (10.6)	1.539 (10.6)	1.533 (10.6)	1.524 (10.5)	1.506 (10.4)
MOR 5% (x10^3 lb/in^2 (MPa))		2.580 (17.8)	2.577 (17.8)	2.573 (17.7)	2.567 (17.7)	2.553 (17.6)
		Failures (%)				
MOE	1	11.359	15.635	23.490	39.599	72.754
	2	1.319	2.403	5.491	15.604	52.912
	3	0.154	0.359	1.268	6.126	38.407
MOR	1	7.772	8.870	10.895	14.250	21.724
	2	0.602	0.760	1.187	2.003	4.676
	3	0.040	0.066	0.127	0.286	1.030
Pith	1	4.950	61.591	99.844	100	100
	2	0.247	37.948	99.689	100	100
	3	0.011	23.499	99.534	100	100
Proof load	1	5.186	5.490	6.021	6.850	8.531
	2	0.272	0.309	0.365	0.446	0.722
	3	0.008	0.017	0.021	0.037	0.068

[a]Target MOE = 0.025 × 10^6 lb/in^2 (0.17 GPa).

Table 14—Effect of discard rate on likelihood of detecting given targeted shift in MOE for SS 2x4 sample[a]

MOE shift ($\times 10^6$ lb/in^2 (GPa))	Step	Effect of discard rate for various MOE means: No discard 1.824×10^6 lb/in^2 (12.6 GPa)	20% discard 1.819×10^6 lb/in^2 (12.5 GPa)	40% discard 1.812×10^6 lb/in^2 (12.5 GPa)	60% discard 1.799×10^6 lb/in^2 (12.4 GPa)	80% discard 1.771×10^6 lb/in^2 (12.2 GPa)
0.02 (0.14)	1	10.545	16.9	30.338	60.375	97.357
	2	1.082	2.846	9.235	36.389	94.776
	3	0.111	0.469	2.821	21.878	92.273
0.04 (0.28)	1	0.635	1.445	4.102	17.020	76.828
	2	0.006	0.021	0.172	2.947	58.983
	3	0	0.002	0.005	0.496	45.318
0.05 (0.34)	1	0.083	0.257	0.985	6.025	55.308
	2	0	0	0.007	0.336	30.640
	3	0	0	0	0.024	16.902
0.10 (0.69)	1	0	0	0	0	0.194
	2	0	0	0	0	0
	3	0	0	0	0	0

[a]Sample size = 413.

Table 15—Effect of discard rate on likelihood of detecting given targeted shift in MOE for No. 2 2x4 sample[a]

MOE shift ($\times 10^6$ lb/in^2 (GPa))	Step	Effect of discard rate for various MOE means: No discard 1.530×10^6 lb/in^2 (10.5 GPa)	20% discard 1.520×10^6 lb/in^2 (10.5 GPa)	40% discard 1.505×10^6 lb/in^2 (10.4 GPa)	60% discard 1.482×10^6 lb/in^2 (10.2 GPa)	80% discard 1.438×10^6 lb/in^2 (9.9 GPa)
0.02 (0.14)	1	17.204	35.523	67.899	96.445	100
	2	2.891	12.601	46.120	93.011	100
	3	0.480	4.431	31.283	89.705	100
0.04 (0.28)	1	1.990	6.899	25.467	75.225	99.933
	2	0.029	0.479	6.506	56.600	99.865
	3	0.001	0.026	1.628	42.576	99.797
0.05 (0.34)	1	0.443	2.043	11.149	54.673	99.622
	2	0	0.038	1.246	30.017	99.245
	3	0	0.001	0.125	16.417	98.871
0.10 (0.69)	1	0	0	0.007	0.302	39.460
	2	0	0	0	0.001	15.634
	3	0	0	0	0	6.161

[a]Sample size = 413.

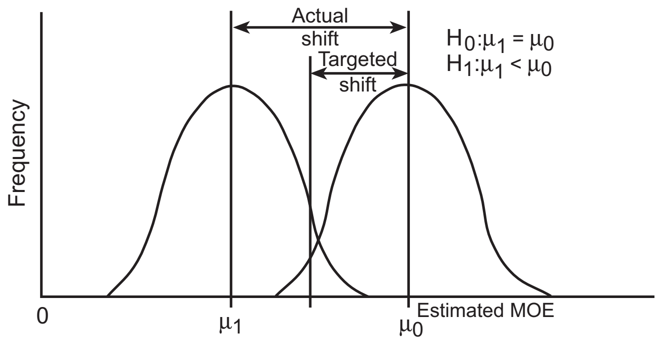

Figure 3—Illustration of shifts in MOE.

Implications of Simulations

The simulations indicate that "what if" scenarios can be considered using normal statistics and existing data instead of further simulations. If a constant is subtracted from the MOE of every piece (Fig. 3), the probability that a test will pick up that difference can be calculated. Equation (8) gives the standard normal variate $Z_{(1-\beta)}$ for detecting an actual shift of a given amount based on various targeted shifts.

$$Z_{(1-\beta)} = \frac{\text{targeted} - \text{actual shift}}{\sigma/\sqrt{n}} \tag{8}$$

Table 16 shows the likelihood of detecting a given actual shift in MOE if the targeted shift is set at different levels. If the actual shift equals the targeted shift, there is a 50% chance of detecting the change. As the actual shift increases or decreases with respect to the targeted shift, there is an increase or decrease in the likelihood of detecting it. This clearly indicates that there is very little chance that a minor shift in MOE would be detected if the trigger level were set to detect a targeted shift of 0.1×10^6 lb/in^2 (0.69 GPa).

Using Equation (9), the random chance of detecting a shift when none had occurred for a given targeted shift is given in Table 17.

$$Z_{(\alpha)} = \frac{\text{targeted} - \text{actual shift}}{\sigma/\sqrt{n}} \tag{9}$$

These values, based on normal distribution statistics, are comparable to those presented in Tables 2 to 5, which were based on simulations. This information has been used as a guide in developing a resource monitoring program at the Southern Pine Inspection Bureau (SPIB).

SPIB Resource Monitoring Program

The SPIB has conducted a resource monitoring program annually since 1994. This program was developed as a means of recognizing whether a significant change has occurred in lumber strength properties. Both the SPIB Technical Committee and Board of Governors have shown solid support for this program since its inception.

Method of Sampling

Mills are selected in a similar fashion to that used in the original North American In-Grade program (Green and others 1989). Each SPIB mill is assigned to a geographic region and the production from each region is totaled. Based on the proportion of production from each region, the number of pieces to be selected from that region is calculated. The target minimum sample size is 360 pieces. Regardless of the actual number of pieces required for a given region, a minimum sample size of 10 pieces/mill and a maximum of 15 pieces/mill were established. Because of the minimum sample size per mill, the actual sample size has ranged from year to year from 380 to 405 pieces. Each year, approximately 30 mills are randomly selected from all of the SPIB subscriber mills that produce No. 2 2×4 lumber (SPIB 1994). Because of timing issues related to reporting production figures and scheduling data collection for the monitoring program, the regional production percentages are based on the annual production figures 2 years prior to the testing. The selection of mills to be sampled is conducted by the Forest Products Laboratory. The mills in each region are randomly ranked. The appropriate number of mills is then selected to obtain the required sample size for each region.

The data collected for each piece include width, thickness, length, moisture content, rings/inch, estimated percentage of latewood, grade-controlling characteristic, maximum strength-reducing characteristic, and two MOE values determined using a Metriguard E-Computer (Metriguard, Pullman, WA) . The E-Computer is calibrated using an 8-lb (3.6 kg) brass weight, and the theoretical constant (79.4) to maintain the "true, dynamic *E*" is used. All data are collected at the individual mill sites. Ambient and wood temperature data are also collected, and an effort is made to test the wood at temperatures above 32°F (0°C).

After the data are collected for each annual program, several data adjustments are made. First, the two E-Computer *E* values are averaged to obtain a single value for each piece. Because the E-Computer assumes nominal dimensions (that is, 2×4 is 1.5 by 3.5 in.), the average *E* value is adjusted for the actual dimensions of each piece. The *E* value is further adjusted for moisture content to 15% using the moisture content model in ASTM D1990.

In a separate, unpublished study conducted by SPIB, the issue of adjusting a flatwise, E-Computer *E* value to be equivalent to an edgewise, static bending *E* value was considered. Data were from a 1994 machine-stress-rated (MSR) tension test program. The sample consisted of 2×4 to 2×10 lumber with a total sample size of 140 pieces. The

Table 16—Likelihood of detecting actual shift for various targeted shifts for 2×4 samples

Grade			Actual shift in MOE ($x10^6$ lb/in^2 (GPa))	Likelihood (%) of detecting actual MOE shift for targeted shifts				
				–0.1 x 10^6 lb/in^2 (–0.69 GPa)	–0.05 x 10^6 lb/in^2 (–0.34 GPa)	–0.04 x 10^6 lb/in^2 (–0.28 GPa)	–0.025 x 10^6 lb/in^2 (–0.17 GPa)	–0.02 x 10^6 lb/in^2 (–0.14 GPa)
SS	Mean μ_0	1.824 (12.6)						
	σ	0.332 (2.3)						
	n	413	–0.15 (–1.0)	99.9	100.0	100.0	100.0	100.0
			–0.125 (–0.86)	93.7	100.0	100.0	100.0	100.0
			–0.1 (–0.69)	50.0	99.9	99.9	100.0	100.0
			–0.05 (–0.34)	0.1	50.0	72.9	93.7	96.7
			–0.04 (–0.28)	0	27.1	50.0	82.1	88.9
			–0.025 (–0.17)	0	6.3	17.9	50.0	62.2
			–0.02 (–0.14)	0	3.3	11.1	37.8	50.0
			–0.01 (–0.07)	0	0.7	3.3	17.9	27.1
No. 2	Mean μ_0	1.531 (10.6)						
	σ	0.366 (2.5)						
	n	413	–0.15 (–1.0)	99.7	100.0	100.0	100.0	100.0
			–0.125 (–0.86)	91.8	100.0	100.0	100.0	100.0
			–0.1 (–0.69)	50.0	99.7	99.7	100.0	100.0
			–0.05 (–0.34)	0.3	50.0	71.2	91.8	95.3
			–0.04 (–0.28)	0	28.8	50.0	79.7	86.7
			–0.025 (–0.17)	0	8.2	20.3	50.0	61.0
			–0.02 (–0.14)	0	4.8	13.4	39.0	50.0
			–0.01 (–0.07)	0	1.3	4.8	20.3	28.8

Table 17—Likelihood of detecting targeted shift when no shift had occurred for 2×4 samples

Grade		Actual shift in MOE ($x10^6$ lb/in^2 (GPa))	Likelihood (%) of detecting actual MOE shift for targeted shifts				
			–0.1 x 10^6 lb/in^2 (–0.69 GPa)	–0.05 x 10^6 lb/in^2 (–0.34 GPa)	–0.04 x 10^6 lb/in^2 (–0.28 GPa)	–0.025 x 10^6 lb/in^2 (–0.17 GPa)	–0.02 x 10^6 lb/in^2 (–0.14 GPa)
SS	Mean μ_0	1.824 (12.6)					
	σ	0.332 (2.3)	0	0.1	0.7	6.3	11.1
	n	413					
No. 2	Mean μ_0	1.531 (10.6)					
	σ	0.366 (2.5)	0	0.3	1.3	8.2	13.4
	n	413					

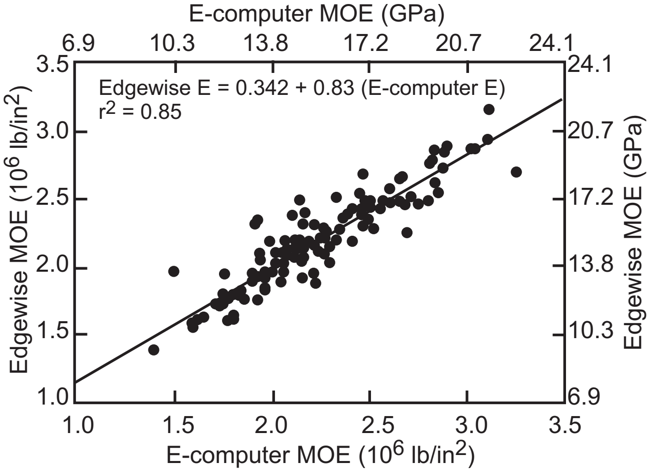

Figure 4—Dynamic flatwise MOE (E-computer) as opposed to edgewise static MOE.

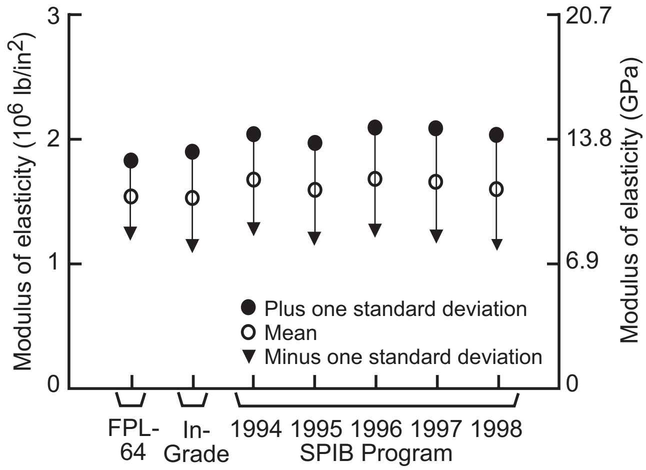

Figure 5—Results of SPIB resource monitoring program, 1994–1998.

E-Computer *E* value was the average of two lengthwise orientations. Information on edgewise static *E* was collected from a Metriguard 312 proof loader using a 21:1 span to depth ratio. The results are plotted in Figure 4.

The data showed that for *E* values less than about 2.0×10^6 lb/in^2 (13.8 GPa), a slight upward adjustment to the E-computer *E* value would be appropriate; for *E* values greater than 2.0×10^6 lb/in^2 (13.8 GPa), a slight downward adjustment would be appropriate. No adjustment was made to convert E-Computer *E* values to edgewise *E* values because average *E* of the SPIB resource monitoring program data was less than 2.0×10^6 lb/in^2 (13.8 GPa) and it would be conservative to not make an adjustment.

Results to Date

The results of the SPIB resource monitoring program from 1994 to 1998 are presented in Table 18 and Figure 5. As a basis for comparison, Table 18 and Figure 5 include data from the original In-Grade test program (Green and Evans 1987) and from No. 2 2×4 tests in the FPL–64 Southern Pine In-Grade program conducted in the 1960s (Doyle and Markwardt 1966). The data for the In-Grade No. 2 2×4 sample, as well as the combined data for No. 2 of all sizes converted to a 2×4 basis, are presented.

In general, the results are favorable and do not indicate a significant departure from the results of the original In-Grade testing program or the FPL–64 test program. In fact, the observed MOE tends to be higher than that obtained in either of these programs. This may be due in part to the fact that fewer mills separated dense lumber after the In-Grade design values were adopted in 1991.

Discussion

The potential for the resource to change continuously over time indicates a critical need for a monitoring system to identify shifts in mechanical properties. The monitoring of In-Grade test results is a very complicated issue that lends itself to many potential methods. For the methods examined, the multiple-stage approach is apparently the most workable system, considering current industry practice and procedures.

Table 18—Summary of MOE test results for Southern Pine resource monitoring program

Variable	FPL–64 No. 2 2x4	In-Grade testing program		Resource monitoring program				
		No. 2 2x4	No. 2 all sizes	1994	1995	1996	1997	1998
Sample size	100	413	2,605	405	405	405	390	380
Average *E* (x10^6 lb/in^2 (GPa))	1.540 (10.6)	1.531 (10.6)	1.563 (10.8)	1.660 (11.4)	1.585 (10.9)	1.678 (11.6)	1.655 (11.4)	1.593 (11.0)
COV (%)	19.0	23.9	—	22.6	24.0	24.1	25.1	27.4

The multiple-stage approach, with multiple steps at Stage I, allows for small sample sizes with a high probability that the selected trigger level is truly a change and did not occur by random chance. The program initiated by the Southern Pine Inspection Bureau is a fine initial attempt at monitoring structural lumber. Ultimately, the decision of the best method for determining changes in the resource lies in the hands of the various grading agencies responsible for supervising the implementation of the current visual grading rules.

There is a definite need for more study of monitoring possibilities. Additional work is needed to assess more fully the economic impact of not recognizing a true shift in properties. For example, if there is a concern that a change in the resource will primarily increase the coefficient of variability (COV) of material properties, then a lower tail (that is, 5^{th} percentile value MOE) might be more sensitive to that change. Other simulations could be conducted with existing data to examine the impact of changes in production, such as increases in MSR production at dimension mills, changes in markets, or changing variability on reliability-based design calculations. Finally, a large representative study of plantation material could be collected across the growth range of the Southern Pine sample to obtain an accurate picture of the impact of rapid growth on properties.

We hope that the most important impact of this publication will be to provide a framework for discussing various options of monitoring properties. We envision that this document will serve as the starting point for determining appropriate methods for monitoring structural lumber.

Conclusions

The simulations demonstrated the relative sensitivity of a multiple-stage monitoring approach to sample size, juvenile wood content, and selected trigger levels:

Differences in a uniform shift in MOE detected by monitoring procedures may be explained as clearly in closed-form solutions based on normal distributions as by simulations.

The targeted shift of MOE should be less than 0.1×10^6 lb/in^2 (0.69 GPa).

For smaller target shifts, it is better to use a multiple-step process to confirm any observed shifts.

A sample size of approximately 400 is very unlikely to give a false positive indication of a shift for MOE after three steps.

Depending on the current relationship between MOE and MOR, if a shift of one rounding rule in MOE is detected, it is possible that many shifts in a rounding rule for MOR may have occurred, particularly in the higher lumber grades.

A multiple-stage approach can be smoothly implemented by a grading agency as part of its regular quality control program.

References

AF&PA. 1997. Design values for wood construction—a supplement to the 1997 edition of the National Design Specification. Washington, DC: American Forest & Paper Association.

ASTM. 1997. Annual book of standards. Vol. 04.10. Wood. ASTM D1990–97. Standard practice for establishing allowable properties for visually-graded dimension lumber from in-grade tests of full-size specimens. West Conshohocken, PA: American Society for Testing and Materials: p. 287.

Barrett, J.D.; Kellogg, R.M.. 1989. Strength and stiffness of dimension lumber. Second growth Douglas-fir: its management and conversion for value. Forintek Canada Corporation. SP–32 ISSN 0824–2119. April: 50–58. Chapter 5.

Bendtsen, B.A.; Senft, J. 1986. Mechanical and anatomical properties in individual growth rings of plantation-grown eastern cottonwood and loblolly pine. Wood and Fiber Science.18(1): 23–38.

Bendtsen, B. Alan; Plantinga, Pamela L.;. Snellgrove, Thomas A. 1988. The influence of juvenile wood on the mechanical properties of 2x4's cut from Douglas-fir plantations. In: Proceedings of the 1988 International Conference on Timber Engineering. Vol. 1.

Biblis, E.J. 1990. Properties and grade yield of lumber from a 27-year-old slash pine plantation. Forest Products Journal. 40(3): 21–24.

Bier, H.; Collins, M.J. 1984. Bending properties of 100 x 50 mm structural timber from a 28-year-old stand of New Zealand radiata pine. IUFRO Group S5.02. Xalapa, Mexico. New Zealand Forest Service Reprint 1774. December.

Boone, R.S.; Chudnoff, M. 1972. Compression wood formation and other characteristics of plantation-grown *Pinus caribaea*. Res. Pap. ITF–13. Rio Piedros, Puerto Rico: U.S. Department of Agriculture, Forest Service, Institute of Tropical Forestry.

Clarke III, A;. Saucier, J.R. 1989. Influence of initial planting density, geographic location and species on juvenile wood formation in southern pine. Forest Products Journal. 39(7/8): 42–48.

Doyle, D.V; Markwardt, L.J. 1966. Properties of Southern Pine in relation to strength grading of dimension lumber. Res. Pap. FPL–64. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

Green, D.W; Evans J.W. 1987. Mechanical properties of visually graded lumber: Vol. 1–8. Springfield VA: National Technical Information Service. U.S. Department of Commerce. PB–88–159–371.

Green, D.W.; Kretschmann, D.E. 1991. Lumber property relationships for engineering design standards. Wood and Fiber Science. 23(3): 436–456.

Green, D.W.; Shelley, B E.; Vokey H.P. 1989. In-grade testing of structural lumber. In: Proceedings of workshop sponsored by In-grade Testing Committee and Forest Products Society. FPS Proceedings 47363. Madison, WI: Forest Products Society.

Johnson, R A.; Evans J.W.; Green, D.W. 1995. Predictive distribution of order statistics. In: Balakrishnan, N., ed. Recent advances in life-testing and reliability. Boca Raton: CRC Press: 505–521.

Johnson, R.A.; Evans J.W.; Green, D.W. 1999. Nonparametric Bayesian predictive distributions for future order statistics. Statistics and Probability Letters. 41: 247–254.

Kretschmann, D.E.; Bendtsen, B.A. 1992. Ultimate tensile stress and modulus of elasticity of fast-grown plantation loblolly pine lumber. Wood and Fiber Science. 24(2): 189–203.

MacPeak, Malcolm D.; Burkart, L.F.; Weldon, D. 1990. Comparison of grade, yield, and mechanical properties of lumber produced from young fast-grown and older slow-grown planted slash pine. Forest Products Journal. 40(1): 11–314.

Montgomery D.C. 1997. Introduction to statistical quality control. 3d ed. New York, NY: John Wiley and Sons, Inc.

Pearson, R.G. 1984. Characteristics of structural importance of clear wood and lumber from fast-grown loblolly pine stands. In: Proceedings, on utilization of the changing wood resource in the Southern United States;1984 June 12–13; North Carolina State University, Raleigh, NC

Pearson, R.G.; Gilmore, R.C. 1971. Characterization of the strength of juvenile wood of loblolly pine. Forest Products Journal. 21(1): 23–30.

Smith, I.; Alkan, S.; Chui, Y.H. 1991. Variation of dynamic properties and static bending strength of a plantation grown red pine. Journal of Institute of Wood Science. 12(4): 221–224.

Southern Pine Inspection Bureau. 1994. Standard grading rules. Pensacola, FL: Southern Pine Inspection Bureau.

Walford, G.B. 1982. Current knowledge of the in-grade bending strength of New Zealand radiata pine. FRI Bull. 15. New Zealand.

Ying, L.; Kretschmann, D.E.; Bendtsen, B.A. 1994. Longitudinal shrinkage in fast-grown loblolly pine plantation wood. Forest Products Journal. 44(1): 58–62.