Statistical Modeling for Process Control in the Sawmill Industry

Rudolf Beran*
Department of Statistics, University of California, Davis
Davis CA 95616, USA
Email: beran@wald.ucdavis.edu, Tel: (530) 754-7765, Fax: (530) 752-7099

SUMMARY

Softwood logs are processed into green boards through a series of horizontal or vertical sawing operations that reduce lumber thickness. This paper uses physical understanding to model how systematic and random errors in board thickness accumulate during sequential resawing. The error model is validated on board thickness measurements gathered at a northern California sawmill. The analysis

- explains previously puzzling patterns in the spatially averaged sample variances of board thickness measurements;
- enables estimating, from measured board thicknesses, the means and variances of the thickness errors introduced by each sawing operation in the observed sequence;
- generates stronger methods for sawmill quality control.

Through submodeling of the mean vector and the covariance matrix of measured thicknesses, the paper finds slight "wedging" in the mean thickness of certain boards and distance-based correlations among the random sawing errors on these boards. The one-way random effects model, used uncritically in earlier analyses of lumber thickness measurements, does not fit the study data.

KEYWORDS: Headrig error, vertical resaw error, horizontal resaw error, error accumulation, covariance structure, submodel fit

* This research was supported in part by the U.S. Forest Service, United States Department of Agriculture and in part by National Science Foundation Grant DMS 0404547.

1. INTRODUCTION

The increasing scarcity of top quality lumber in the western United States provides an economic incentive for strengthening process control in the sawmill industry. In western U.S. soft-wood mills, the green lumber end-product is the result of several distinct sawing operations that reduce board thickness. Vibration of the saws contributes to irregularities in thickness of the final green lumber. Misalignment of the saws produces green boards that are systematically wedge-shaped or tapered or otherwise deformed. The green lumber is dried, either naturally or in a kiln, and is then planed to standard dimensions for the market. The green lumber must be sawn thick enough to offset random and systematic irregularities in shape and to allow for shrinkage when it dries. Boards too thin to meet market standards for thickness must be resawn wastefully.

Because sawing errors accumulate during sequential sawing operations, it has not been clear how to monitor, from board thickness measurements, the performance of secondary sawing machines. Resolution of this puzzle was the primary motivation for the modeling and analysis reported in this paper. This account is a statistical detective story directed at professional statisticians who are interested in sawmill quality control issues. Some carpentry experience is advantageous in understanding sawmill operations.

The U.S. Forest Service made available lumber thickness measurements gathered in a study at a redwood mill in northern California. In this mill, tree trunks of large diameter are sectioned into long logs, which are then broken down into boards as follows:

- Each log, fastened to a wheeled steel carrier, is pulled repeatedly through a large vertical bandsaw called a *headrig*, which slices a large vertical slab from the log. The nominal thickness of the slab is either 4 inches or 2 inches.
- Turned onto its broad face, each slab is carried lengthwise on rollers through multiple, equally spaced rotary saws that divide it into boards of nominal 12 inch width.

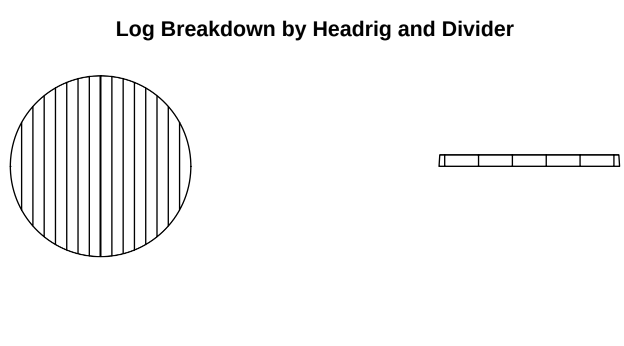

Figure 1. Breakdown of a softwood log into slabs by the headrig bandsaw (left) is followed by division of each slab into boards (right).

Some of the boards obtained by the foregoing breakdown of the log are resawn to obtain boards of half the thickness. Resawing of boards may be repeated. At the mill studied, resawing is done in two alternate ways:

- In a *horizontal resaw*, the board is pressed flat against fixed horizontal rollers which carry it through a saw that cuts down the length of the board, parallel to the rollers. This operation produces an upper offspring board and a lower offspring board of nominally equal thickness.
- In a *vertical splitter resaw*, the board is stood vertically on one edge, between two sets of spring-loaded rollers which guide it through a saw that cuts vertically down the length of the board. This operation produces a left offspring board and a right offspring board of nominally equal thickness. Ideally, the spring tensions on both sets of rollers in the vertical splitter are equal.

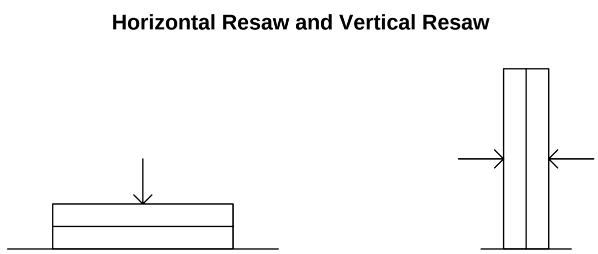

Figure 2. Division of a board into two thinner boards is accomplished either by a horizontal resaw (left) or by a vertical splitter resaw (right).

Principal goals of the study were:

- To develop a physically based statistical model that describes how systematic and random errors in lumber thickness accumulate through a sequence of sawing operations;
- To compare patterns of variability in the the model with patterns observed in the data;
- To use the fitted model to quantify the variability introduced by each resaw in a sequence of sawing operations;
- To point out the implications of the foregoing work for process control in sequential resawing;
- To relate the model and data analysis to assumptions made in earlier studies of lumber thickness data, such as the one-way random effects model with normally distributed errors (cf. Warren [8]).

Section 2.1 describes the experiment that produced the board thickness measurements used in this study. Section 2.3 presents spatially averaged sample means and sample variances of these measurements, computed for each set of offspring boards produced during the observed sawing sequence. It is seen that the sample variances of offspring boards depend strongly on the sawing sequence that produces them. The consistent patterns of magnitude found in these sample variances are puzzling at first sight. The statistical thickness error model developed in Section 3 explains the patterns quantitatively by considering the physical characteristics of each sawing operation. Section 4 develops implications of the statistical model for process control. Through submodeling of the mean vector and the covariance matrix of measured thicknesses, Section 5 finds slight "wedging" in the mean thickness of certain offspring boards and distance-based correlations among the random sawing errors on these boards. These findings cast serious doubts on the use of the one-way random effects model in Warren's [8] pioneering statistical studies of lumber thickness measurements.

2. THE LUMBER THICKNESS DATA

2.1. Collection and coding of the measurements

Section 1 defines the various sawing operations. Boards selected "at random" after breakdown of a log by a headrig and divider were followed through one or more resaws into thinner boards. This was done for two distinct sawing sequences:

Sawing sequence for 4 inch lumber. Boards of nominal 4 inch thickness coming off a headrig and divider were followed through two resaws, first into 2 inch lumber, then into 1 inch lumber. The first resawing operation was a horizontal resaw. The 2 inch bottom offspring boards were distinguished from the 2 inch top offspring boards. Half of the bottom 2 inch boards and half of the top 2 inch boards were then followed through a vertical splitter yielding 1 inch lumber; the remaining 2 inch boards were sent through another horizontal resaw into 1 inch lumber. The first sample consisted of 50 boards from headrig 1, the second sample of 41 boards from headrig 2.

Sawing sequence for 2 inch lumber. Boards of nominal 2 inch thickness coming off a headrig and divider were followed through a vertical splitter yielding 1 inch lumber. The first sample consisted of 97 boards from headrig 1, the second sample of 100 boards from headrig 2.

The thickness of every board produced in these two experiments was measured by micrometer to .01 inch at eight standardized points, 4 along each edge in facing pairs. The measurements on each board were labelled by board number, by position on the board, and by a three digit sawing code xyz that identifies the headrig and resawing sequence producing the board.

For the 4 inch lumber sample, the first digit x identifies the headrig:

$x = 1$ for boards from headrig 1

$x = 2$ for boards from headrig 2

The second digit y describes the observed offspring of the second sawing operation:

$y = 0$ if no second sawing operation has been performed

$y = 1$ if the board is the top offspring from a horizontal resaw
$y = 2$ if the board is the bottom offspring from a horizontal resaw
$y = 3$ if the board is the left offspring from a vertical splitter
$y = 4$ if the boards is the right offspring from a vertical splitter

The third digit z describes analogously the observed offspring of the third sawing operation.

Sawing codes for the 2 inch lumber sample are defined in the same way. Thus, care is needed to distinguish between sawing code 100 (say) referring to 4 inch boards from headrig 1 and sawing code 100 referring to 2 inch boards from headrig 1.

Listing of the data by board number and by sawing code showed that some records were missing for some resawing sequences. These boards and their offspring were dropped in analysis of the data. The net sample sizes after omission of incomplete records were

4 inch lumber:

49 boards in codes 100, 110, 120
24 boards in codes 111, 112, 121, 122
25 boards in codes 113, 114, 123, 124
40 boards in codes 200, 210, 220
20 boards in codes 211, 212, 221, 222, 213, 214, 223, 224

2 inch lumber:

96 boards in codes 100, 130, 140
98 boards in codes 200, 230, 240

2.2. Assumptions on the sawing processes and measurements

Expert advice from those who collected the lumber thickness measurements supports the following assumptions:

- All saws operated stably, without bursts of erratic behavior or appreciable mechanical wear during the processing of the sampled boards.
- In a horizontal resaw, the board being cut is pressed completely flat against the fixed horizontal bed of rollers.
- In a vertical splitter resaw, the spring tensions on the two sets of rollers that hold the board on edge are very nearly equal.
- Thickness is measured at the same 4 pairs of facing edge points throughout the sawing sequence.
- Errors in the measurements are negligible relative to the actual variability in lumber thickness.
- Sawing errors that occur on a particular board are independent of those on the other boards sampled.

The first three assumptions are satisfied in a well-controlled sawmill. Fulfilling the next two assumptions is a matter of care in measuring thickness. The last assumption reflects separation in the production times of the boards that were selected "at random" for the study.

2.3. Spatially averaged sample means and sample variances

Let $x_{ij}(k)$ denote the measured thickness of board i at edge position j in sawing code $k = xyz$. The values of j range from 1 to 8, with values 1 to 4 labelling the standardized measurement positions along one edge of the board and values 5 to 8 labelling the respective facing positions along the other edge of the board. The two board edges were systematically identified and tracked through resaws.

Let $I(k)$ denote the set of board numbers in sawing code k for which we have thickness measurements and let $n(k)$ be the number of such boards. The sample mean and sample variance of the thickness measurements at position j in sawing code k are, respectively,

$$\hat{\mu}_j(k) = n^{-1}(k) \sum_{i \in I(k)} x_{ij}(k), \qquad \hat{\sigma}_j^2(k) = (n(k) - 1)^{-1} \sum_{i \in I(k)} (x_{ij}(k) - \hat{\mu}_j(k))^2,$$

for $1 \le j \le 8$. The *spatially averaged* sample mean and sample variance in sawing code k are defined as

$$\hat{\hat{\mu}}(k) = 8^{-1} \sum_{j=1}^{8} \hat{\mu}_j(k), \qquad \hat{\hat{\sigma}}^2(k) = 8^{-1} \sum_{j=1}^{8} \hat{\sigma}_j^2(k).$$

The initial analysis of how sawing errors accumulate will be based on these spatially averaged statistics. Their values for the 4 inch and 2 inch lumber experiments are reported in Tables I and II.

Table I. Spatially averaged sample means and variances for the 4 inch lumber

Sawing Code k	$\hat{\hat{\mu}}(k)$	$100\hat{\hat{\sigma}}^2(k)$	Sawing Code k	$\hat{\hat{\mu}}(k)$	$100\hat{\hat{\sigma}}^2(k)$
100	3.91	.248	200	3.90	.241
110	1.88	.281	210	1.87	.287
120	1.89	.047	220	1.89	.053
111	0.84	.339	211	0.83	.374
112	0.91	.075	212	0.91	.084
121	0.84	.131	221	0.84	.127
122	0.91	.070	222	0.91	.072
113	0.91	.062	213	0.91	.058
114	0.87	.103	214	0.87	.077
123	0.91	.015	223	0.91	.015
124	0.88	.032	224	0.88	.026

Table II. Spatially averaged sample means and variances for the 2 inch lumber

Sawing Code k	$\hat{\bar{\mu}}(k)$	$100\hat{\bar{\sigma}}^2(k)$	Sawing Code k	$\hat{\bar{\mu}}(k)$	$100\hat{\bar{\sigma}}^2(k)$
100	1.91	.222	200	1.89	.075
130	0.91	.050	230	0.90	.024
140	0.89	.077	240	0.88	.036

In Tables I and II, the spatially averaged sample means behave as one might expect: a small amount of average lumber thickness is converted to sawdust in either horizontal or vertical resawing. The spatially averaged sample variances in both tables differ remarkably across the sawing codes. Top offspring boards from horizontal resaws consistently have much larger variances than the corresponding bottom offspring boards. Left offspring boards from vertical resaws consistently have smaller variances than the corresponding right offspring boards. These unexpected patterns in the sample variances puzzled those who gathered the measurements. Section 3 resolves the puzzle through statistical modeling of resawing errors.

3. A MULTIVARIATE STATISTICAL MODEL

This section constructs a statistical model for the accumulation of random and systematic errors in lumber thickness during sequential resawing. The model expresses an idealized physical understanding of the sawing operations and will be seen to explain the patterns observed in Tables I and II.

3.1. A general multivariate model

Let $x_i(k) = (x_{i1}(k), x_{i2}(k), \ldots, x_{i8}(k))'$ be the thickness measurements on board i, arranged as a column vector. The first part of the statistical model for thickness measurements asserts that, for board i in sawing code k,

$$x_i(k) = \mu(k) + e_i(k), \tag{1}$$

where $\mu(k)$ is a vector constant and the $\{e_i(k) \colon i \in I(k)\}$ are independent, identically distributed random vectors with mean vector zero and unknown covariance matrix $\Sigma(k)$. Then $\mathrm{E}(x_i(k)) = \mu(k)$ and $\mathrm{Cov}(x_i(k)) = \Sigma(k)$. The components of $\mu(k)$ are the *positional means* $\{\mu_j(k) \colon 1 \leq j \leq 8\}$ for boards in sawing code k. The components of $e_i(k)$ are the random thickness errors found on boards in sawing code k. The diagonal elements of $\Sigma(k)$ are the *positional variances* $\{\sigma_j^2(k) \colon 1 \leq j \leq 8\}$ for boards in that code.

The second part of the statistical model, developed in the next subsections, draws on a physical undersatnding of resawing to develop relations among the positional variances in parent and offspring boards and among the positional means in parent and offspring boards. These relations link random and systematic thickness errors in the various sawing codes.

3.2. Relations among positional means and variances for 4 inch lumber

Consider the sawing sequence that starts with 4 inch lumber from headrig 1 that is subsequently resawn twice, the first time horizontally. The modeling for resawn 4 inch lumber from headrig 2 is completely analogous.

Positional variances. During a horizontal resaw of board i in sawing code 100, the downward pressure on the board ensures the following:

- The bottom surface of board i is pushed flat against the fixed horizontal bed of rollers.
- Consequently, the vector $e_i(100)$ of random sawing errors in the thickness of board i in code 100 is expressed as random fluctuations in the top surface of that board. These fluctuations are inherited by the top surface of offspring board i in code 110.
- Because the bottom surface of offspring board i in code 120 is held flat against the rollers, the horizontal resaw introduces the vector $e_i(120)$ of random sawing errors into the top surface of that board.
- Compensatingly, the horizontal resaw introduces the vector $-e_i(120)$ of random sawing errors into the bottom surface of board i in code 110. The top surface of board i in code 110 exhibits the errors $e_i(100)$, as noted in the second bulleted point. Thus, the overall random thickness error vector $e_i(110)$ for board i in code 110 must equal the top surface error vector $e_i(100)$ plus the bottom surface error vector $-e_i(120)$. This reasoning shows that $e_i(110) = e_i(100) - e_i(120)$.

The same argument applies to subsequent horizontal resawing of boards from codes 110 and 120. Consequently, for board number i,

$$\begin{aligned} e_i(110) &= e_i(100) - e_i(120) \\ e_i(111) &= e_i(110) - e_i(112) \\ e_i(121) &= e_i(120) - e_i(122). \end{aligned} \tag{2}$$

Moreover, the random error vectors $e_i(100)$ and $e_i(120)$ are independent because the headrig saw and the first resaw are physically independent. Similarly, the random vectors $e_i(110)$ and $e_i(112)$ are independent, as are the random vectors $e_i(120)$ and $e_i(122)$.

Equations (2) yield the following relationships for the positional variances:

$$\begin{aligned}\sigma_j^2(110) &= \sigma_j^2(100) + \sigma_j^2(120)\\ \sigma_j^2(111) &= \sigma_j^2(110) + \sigma_j^2(112)\\ \sigma_j^2(121) &= \sigma_j^2(120) + \sigma_j^2(122).\end{aligned} \tag{3}$$

The assumption of stability in the operation of the horizontal resaw implies that its variability is unchanged when resawing boards in either code 110 or 120. Thus, we expect

$$\sigma_j^2(112) = \sigma_j^2(122). \tag{4}$$

For $k = 110$ or 120, let the vector $d_i(k)$ represent random thickness errors made by the vertical splitter saw on its *left* side as it rips through board i in sawing code k. The left side yields the offspring boards in sawing codes 113 or 123. The $\{d_i(k): i \in I(k)\}$ are independent, identically distributed random vectors with mean vector zero and unknown covariance matrix $T(k)$. The diagonal elements of $T(k)$ define the *positional splitter variances* $\{\tau_j^2(k): 1 \leq j \leq 8\}$. The random vectors $e_i(110)$ and $d_i(110)$ are statistically independent because the first horizontal resaw and the subseqent vertical splitter resaw are physically independent operations. Similarly, $e_i(120)$ and $d_i(120)$ are independent random vectors.

When a board in code k is vertically resawn, let $\lambda(k)$ denote the fraction of board thickness error that is transmitted to output from the left side of the vertical splitter. The physical picture of a vertical splitter resaw indicates that, for board i,

$$\begin{aligned}e_i(113) &= \lambda(110)e_i(110) + d_i(110)\\ e_i(114) &= (1-\lambda(110))e_i(110) - d_i(110)\\ e_i(123) &= \lambda(120)e_i(120) + d_i(120)\\ e_i(124) &= (1-\lambda(120))e_i(120) - d_i(120).\end{aligned} \tag{5}$$

The value $\lambda(110) = 1/2$ means that the splitter divides the random thickness error in the parent board equally among the left and right offspring boards. This will happen only if the spring tensions on the two sets of rollers in the vertical splitter are perfectly equal (cf. the description of a vertical splitter resaw in Section 1). Other values of $\lambda(110)$ model the effects of unequal tension in the two sets of springs. For instance, if the left side of the splitter, touching offspring boards in code 113, is firmer than the right side of the splitter, touching offspring boards in code 114, then the thickness errors in each parent board are transmitted more to their offspring boards in code 114. In this event, $\lambda(110) < 1/2$. The converse interpretation holds if $\lambda(110) > 1/2$. The third and fourth identities in (5) are the analogous relations for vertical splitter resaws of parent boards in sawing code 120.

It follows from (5) that

$$\begin{aligned}\sigma_j^2(113) &= \lambda^2(110)\sigma_j^2(110) + \tau_j^2(110)\\ \sigma_j^2(114) &= (1-\lambda(110))^2\sigma_j^2(110) + \tau_j^2(110)\\ \sigma_j^2(123) &= \lambda^2(120)\sigma_j^2(120) + \tau_j^2(120)\\ \sigma_j^2(124) &= (1-\lambda(120))^2\sigma_j^2(120) + \tau_j^2(120).\end{aligned} \tag{6}$$

The assumption of stability in the operation of the vertical splitter implies that the spring tensions on the rollers are unchanged as is the variability of the saw blade when resawing boards in either code 110 or 120. Thus, we expect

$$\lambda(110) = \lambda(120) \qquad \tau_j^2(110) = \tau_j^2(120). \tag{7}$$

For process control purposes, the key parameters in equations (3) and (6) are $\sigma_j^2(k)$ for $k = 100, 120, 112, 122$ plus $\tau_j^2(k)$ and $\lambda(k)$ for $k = 110, 120$. It follows from the foregoing analysis that, at measurement position j, $\sigma_j^2(100)$ measures variability introduced by headrig 1; $\sigma_j^2(120)$ measures additional variability introduced by the first horizontal resaw into 2 inch lumber; and $\sigma_j^2(112)$ or $\sigma_j^2(122)$ measures additional variability introduced by the second horizontal resaw into 1 inch lumber. Moreover, $\tau_j^2(110)$ or $\tau_j^2(120)$ measures additional variability introduced by the vertical splitter resaw into 1 inch lumber; and $\lambda(110)$ or $\lambda(120)$ measures the fraction of the parent board thickness error that the vertical splitter transmits to left offspring boards, that is, to boards in code 113 or 123. By estimating these parameters, the random errors introduced by the various saws involved in the sequential resawing can be monitored.

Positional means. Let $\alpha_j(k)$ denote the loss in mean thickness at position j caused by horizontal resawing of boards in code k. This is the mean thickness lost to sawdust. Let $\beta_j(k)$ denote the loss in mean thickness at position j due to vertical resawing of boards in code k. Evidently,

$$\begin{aligned}\mu_j(100) &= \mu_j(110) + \mu_j(120) + \alpha_j(100)\\ \mu_j(110) &= \mu_j(111) + \mu_j(112) + \alpha_j(110)\\ \mu_j(120) &= \mu_j(121) + \mu_j(122) + \alpha_j(120)\\ \mu_j(110) &= \mu_j(113) + \mu_j(114) + \beta_j(110)\\ \mu_j(120) &= \mu_j(123) + \mu_j(124) + \beta_j(120).\end{aligned} \tag{8}$$

Stability of the resawing processes implies that the loss of mean thickness incurred by a horizontal resaw or a vertical resaw does not change during the experiment. Thus, we expect

$$\alpha_j(110) = \alpha_j(120) \qquad \beta_j(110) = \beta_j(120). \tag{9}$$

The systematic losses in thickness introduced by the various saws involved in the sequential resawing can be monitored by estimating the parameters $\alpha_j(k)$ for $k = 100, 110, 120$ and $\beta_j(k)$ for $k = 110, 120$.

3.3. Checking the model on the 4 inch lumber data

In a well-controlled saw-mill, the positional means $\{\mu_j(k)\colon 1 \le j \le 8\}$ and the positional variances $\{\sigma_j^2(k)\colon 1 \le j \le 8\}$ will vary only slightly with j. In such circumstances, the spatially averaged mean and variance, defined as

$$\bar{\mu}(k) = 8^{-1}\sum_{j=1}^{8}\mu_j(k), \qquad \bar{\sigma}^2(k) = 8^{-1}\sum_{j=1}^{8}\sigma_j^2(k),$$

may serve as adequate summaries. The values of $\bar{\mu}(k)$ and $\bar{\sigma}^2(k)$ are estimated by the spatially averaged sample mean $\hat{\bar{\mu}}(k)$ and sample variance $\hat{\bar{\sigma}}^2(k)$ recorded in Table I. Because of spatial averaging, the sampling variability of these estimators is relatively low, even for the relatively small samples sizes available for many of the sawing codes in this study.

Let $\bar{\alpha}(k)$ denote the average of the $\{\alpha_j(k)\colon 1 \le j \le 8\}$ and let $\bar{\beta}(k)$ similarly denote the average of the $\{\beta_j(k)\colon 1 \le j \le 8\}$. From equations (8) and (9),

$$\begin{aligned}
\bar{\mu}(100) &= \bar{\mu}(110) + \bar{\mu}(120) + \bar{\alpha}(100)\\
\bar{\mu}(110) &= \bar{\mu}(111) + \bar{\mu}(112) + \bar{\alpha}(110)\\
\bar{\mu}(120) &= \bar{\mu}(121) + \bar{\mu}(122) + \bar{\alpha}(120)\\
\bar{\mu}(110) &= \bar{\mu}(113) + \bar{\mu}(114) + \bar{\beta}(110)\\
\bar{\mu}(120) &= \bar{\mu}(123) + \bar{\mu}(124) + \bar{\beta}(120)\\
\bar{\alpha}(110) &= \bar{\alpha}(120)\\
\bar{\beta}(110) &= \bar{\beta}(120).
\end{aligned} \tag{10}$$

Let $\bar{\tau}^2(k)$ be the average of the $\{\tau_j^2(k)\colon 1 \le j \le 8\}$. It follows immediately from equations (3), (4), (6) and (7) that

$$\begin{aligned}
\bar{\sigma}^2(110) &= \bar{\sigma}^2(100) + \bar{\sigma}^2(120)\\
\bar{\sigma}^2(111) &= \bar{\sigma}^2(110) + \bar{\sigma}^2(112)\\
\bar{\sigma}^2(121) &= \bar{\sigma}^2(120) + \bar{\sigma}^2(122)\\
\bar{\sigma}^2(112) &= \bar{\sigma}^2(122)
\end{aligned} \tag{11}$$

and that

$$\begin{aligned}
\bar{\sigma}^2(113) &= \lambda^2(110)\bar{\sigma}^2(110) + \bar{\tau}^2(110)\\
\bar{\sigma}^2(114) &= (1 - \lambda(110))^2\bar{\sigma}^2(110) + \bar{\tau}^2(110)\\
\bar{\sigma}^2(123) &= \lambda^2(120)\bar{\sigma}^2(120) + \bar{\tau}^2(120)\\
\bar{\sigma}^2(124) &= (1 - \lambda(120))^2\bar{\sigma}^2(120) + \bar{\tau}^2(120)\\
\bar{\tau}^2(110) &= \bar{\tau}^2(120)\\
\lambda(110) &= \lambda(120).
\end{aligned} \tag{12}$$

The first four equations in (12) may be solved algebraically to give

$$\begin{aligned} \lambda(110) &= .5 + .5[\bar{\sigma}^2(113) - \bar{\sigma}^2(114)]/\bar{\sigma}^2(110) \\ \bar{\tau}^2(110) &= \bar{\sigma}^2(113) - \lambda^2(110)\bar{\sigma}^2(110) \end{aligned} \tag{13}$$

and the analogous formulae for $\lambda(120)$ and $\bar{\tau}^2(120)$.

A useful test of the model for sequential resawing of the 4 inch lumber is to check whether the spatially averaged sample means and variances in Table I *approximately* satisfy equations (10) through (12). The board thicknesses in this study were measured to the nearest .01 inch. It is apparent from Tables I and II that the random sawing errors and the losses in mean thickness to sawdust were all much smaller than 1.00 inch in magnitude. It follows that the measurements contain at most two significant figures of information about the random sawing errors and the mean losses.

Fitting the first five equations in (10) to Table I yields the following estimated mean sawing losses for resawn lumber from headrig 1:

$$\begin{aligned} &\hat{\bar{\alpha}}(100) = .14 && \\ &\hat{\bar{\alpha}}(110) = .13 && \hat{\bar{\alpha}}(120) = .14 \\ &\hat{\bar{\beta}}(110) = .10 && \hat{\bar{\beta}}(120) = .10 \end{aligned}$$

Similarly, for resawn lumber from headrig 2:

$$\begin{aligned} &\hat{\bar{\alpha}}(200) = .14 && \\ &\hat{\bar{\alpha}}(210) = .13 && \hat{\bar{\alpha}}(220) = .14 \\ &\hat{\bar{\beta}}(210) = .09 && \hat{\bar{\beta}}(220) = .10. \end{aligned}$$

Bearing in mind that thickness measurements were made to the nearest .01 inch, the foregoing estimates nearly satisfy the equalities predicted in the last two equations of (10) and the analogous equations for headrig 2 lumber. Thus, a horizontal resaw loses to sawdust about .14 inch of spatially averaged mean thickness while a vertical splitter resaw loses about .10 inch. Closer examination of the data indicates that this mean thickness loss is nearly constant across all 8 measurement positions.

From Table I, $\hat{\bar{\sigma}}^2(100) = .248 \times 10^{-2}$ while $\hat{\bar{\sigma}}^2(200) = .241 \times 10^{-2}$. Thus, the random sawing errors from headrig 2 in the 4 inch lumber experiment have essentially the same variability as those from headrig 1.

According to the four variance equalities in (11), the spatially averaged sample variances for headrig 1 in Table I should satisfy the following approximate equalities:

$$\begin{aligned} .281 &\approx .248 + .047 \\ .339 &\approx .281 + .075 \\ .131 &\approx .047 + .070 \\ .075 &\approx .070. \end{aligned}$$

For headrig 2, the analogous approximate equalities linking Table I to the model are:

$$\begin{aligned} .287 &\approx .241 + .053 \\ .374 &\approx .287 + .084 \\ .127 &\approx .053 + .072 \\ .084 &\approx .072. \end{aligned}$$

It is thus apparent that the physically based statistical model for horizontal resaws largely explains the otherwise puzzling pattern of variability observed in the offspring boards obtained from one or two sequential horizontal resaws. We cannot expect closer agreement, given that lumber thickness measurements were made to .01 inch and that sample sizes were modest.

Substituting spatially averaged sample variances into (13) suggests the estimators

$$\begin{aligned} \hat{\lambda}(110) &= .5 + .5[\hat{\bar{\sigma}}^2(113) - \hat{\bar{\sigma}}^2(114)]/\hat{\bar{\sigma}}^2(110) \\ \hat{\bar{\tau}}^2(110) &= \max\{\hat{\bar{\sigma}}^2(113) - \hat{\lambda}^2(110)\hat{\bar{\sigma}}^2(110), 0\}. \end{aligned}$$

The positive part adjustment on the right side of the second equation above reduces the mean squared error of the naive, possibly negative, estimator that omits this step. Analogous formulae hold for $\hat{\lambda}(120)$, $\hat{\bar{\tau}}^2(120)$, and the counterparts for headrig 2. For the 4 inch lumber sequence from headrig 1, Table I yields

$$\begin{aligned} \hat{\lambda}(110) &= .427 \qquad & \hat{\bar{\tau}}^2(110) &= .0108 \times 10^{-2} \\ \hat{\lambda}(120) &= .319 \qquad & \hat{\bar{\tau}}^2(120) &= .0102 \times 10^{-2}. \end{aligned}$$

Note that $\hat{\bar{\tau}}^2(110)$ nearly equals $\hat{\bar{\tau}}^2(120)$ as predicted by (12). The values of $\hat{\lambda}(110)$ and $\hat{\lambda}(120)$ are both less than the 1/2 expected from an ideal splitter. This indicates that the left side of the vertical splitter, touching the code 113 or 123 boards, is firmer than the right side, touching the code 114 or 124 boards. (See the discussion that follows equation (5)). Similarly, for the 4 inch lumber sequence from headrig 2,

$$\begin{aligned} \hat{\lambda}(210) &= .467 \qquad & \hat{\bar{\tau}}^2(210) &= \max\{-.00456 \times 10^{-2}, 0\} = 0 \\ \hat{\lambda}(220) &= .396 \qquad & \hat{\bar{\tau}}^2(220) &= .00668 \times 10^{-2}. \end{aligned}$$

Again, the left side of the vertical splitter, touching the code 213 or 223 boards, is firmer than the right side, touching the 214 or 224 boards. In fact, the same vertical splitter was used to divide parent boards stemming from headrig 1 and headrig 2. While it is possible that the vertical splitter was less variable at the time it resawed boards from headrig 2, it is also likely that the differing estimates of $\bar{\tau}^2$ in the preceding two displays reflect the limitations of estimates based on lumber thickness measurements to the nearest .01 inch.

The statistical model of Section 3.2 thus explains quantitatively how systematic and random errors in lumber thickness accumulate in resawing 4 inch boards. Physically based, the model and its estimated parameter values provide a sound basis for understanding the thickness errors contributed by each resawing operation.

3.4. Relations among positional means and variances for 2 inch lumber

Consider next the sawing sequence that starts with 2 inch lumber from headrig 1 that is sent through the vertical splitter. The modeling for 2 inch lumber from headrig 2 is completely analogous. The argument used to derive equations (5) and (6) in Section 3.3 now yields the relations

$$\begin{aligned} e_i(130) &= \lambda(100)e_i(100) + d_i(100) \\ e_i(140) &= (1-\lambda(100))e_i(100) - d_i(100) \end{aligned}$$

and

$$\begin{aligned} \sigma_j^2(130) &= \lambda^2(100)\sigma_j^2(100) + \tau_j^2(100) \\ \sigma_j^2(140) &= (1-\lambda(100))^2\sigma_j^2(100) + \tau_j^2(100). \end{aligned} \tag{14}$$

Here $\tau_j^2(100)$ measures the variability introduced at position j by the vertical splitter and $\lambda(100)$ is the fraction of parent board thickness error that is transmitted to output from the left side of the splitter. The equation relating mean thicknesses of parent and offspring boards is

$$\mu_j(100) = \mu_j(130) + \mu_j(140) + \beta_j(100), \tag{15}$$

where $\beta_j(100)$ is the loss to sawdust in mean thickness at position j caused by the vertical resaw.

3.5. Checking the model on the 2 inch lumber data

It follows immediately from equation (15) that the spatially averaged means and variances satisfy

$$\bar{\mu}(100) = \bar{\mu}(130) + \bar{\mu}(140) + \bar{\beta}(100) \tag{16}$$

and from equation (14) that

$$\begin{aligned} \bar{\sigma}^2(130) &= \lambda^2(100)\bar{\sigma}^2(100) + \bar{\tau}^2(100) \\ \bar{\sigma}^2(140) &= (1-\lambda(100))^2\bar{\sigma}^2(100) + \bar{\tau}^2(100). \end{aligned} \tag{17}$$

Fitting equation (16) and its analog for headrig 2 to Table II yields the following estimated mean sawing losses for 2 inch lumber from headrigs 1 and 2:

$$\hat{\bar{\beta}}(100) = .11 \qquad \hat{\bar{\beta}}(200) = .11.$$

As might be expected, these estimated spatially averaged mean thickness losses due to vertical splitting of the 2 inch lumber from the two headrigs are equal. Moreover, they are close to the already discussed estimated spatially averaged mean thickness losses due to vertical splitting in the second resaw of 4 inch lumber from the two headrigs.

From Table II, $\hat{\bar{\sigma}}^2(100) = .222 \times 10^{-2}$ while $\hat{\bar{\sigma}}^2(200) = .075 \times 10^{-2}$. Thus, the random sawing errors from headrig 2 in the 2 inch lumber experiment are considerably smaller than those from headrig 1. By reasoning akin to that in Section 3.3, relations (17) lead to the estimators

$$\hat{\lambda}(100) = .5 + .5[\hat{\bar{\sigma}}^2(130) - \hat{\bar{\sigma}}^2(140)]/\hat{\bar{\sigma}}^2(100)$$
$$\hat{\bar{\tau}}^2(100) = \max\{\hat{\bar{\sigma}}^2(130) - \hat{\lambda}^2(100)\hat{\bar{\sigma}}^2(100), 0\}$$

and the analogous expressions for their headrig 2 counterparts. For the 2 inch lumber sequence from headrigs 1 and 2, Table II yields

$$\hat{\lambda}(100) = .439 \qquad \hat{\bar{\tau}}^2(100) = .00718 \times 10^{-2}$$
$$\hat{\lambda}(200) = .420 \qquad \hat{\bar{\tau}}^2(200) = .0108 \times 10^{-2}.$$

Here again, the left side of the vertical splitter, touching the code 130 or 230 boards, is firmer than the right side, touching the code 140 or 240 boards. The same vertical splitter was used to divide 2 inch parent boards stemming from headrig 1 and headrig 2. While it is possible that the vertical splitter was less variable at the time it resawed boards from headrig 1, it is also likely that the differing estimates of $\bar{\tau}^2$ in the preceding two displays reflect the limitations of estimates based on lumber thickness measurements to the nearest .01 inch.

Much as for the 4 inch lumber, the statistical model of Section 3.4 explains quantitatively how systematic and random errors in lumber thickness accumulate in resawing 2 inch headrig output.

4. IMPLICATIONS FOR PROCESS CONTROL

The modeling and data analysis of Section 3 inform several aspects of process control in the sawmill industry:

- the monitoring of individual saws;
- the assessment of priorities in sawmill improvement projects;
- the setting of target thicknesses in cutting green lumber.

This section develops these points.

4.1. Monitoring saw performance

The overall sawing process is under good control only if the headrig and subsequent resaws are individually under good control. The statistical modeling and data analysis in Section 3 provides a template for what can be done routinely to monitor the performance of individual saws. The basic steps are:

a) Follow randomly selected boards through all sawing operations, taking thickness measurements at every stage of the process. This is accomplished best by automated equipment that sends precise thickness measurements taken at many points along each edge of a board directly to a computer for analysis.

b) Check as in Sections 3.3 and 3.5 whether the estimates of spatially averaged mean and variance approximately satisfy the mathematical relations expected under the physical model for resawing. Failure in these relations with an adequate sample size would point to poorly controlled sawing.
c) If the model fits as expected, concentrate on the key parameter estimates that reveal performance of individual saws. For instance, in the 4 inch lumber sequence:
   - The pair $\hat{\bar{\mu}}(100)$, $\hat{\bar{\sigma}}^2(100)$ summarizes the performance of headrig 1.
   - The pair $\hat{\bar{\mu}}(120)$, $\hat{\bar{\sigma}}^2(120)$ summarizes the performance of the first horizontal resaw from 4 inches to 2 inches.
   - The pair $\hat{\bar{\mu}}(112)$, $\hat{\bar{\sigma}}^2(112)$ and the pair $\hat{\bar{\mu}}(122)$, $\hat{\bar{\sigma}}^2(122)$ both summarize the performance of the second horizontal resaw from 2 inches to 1 inch. Ordinarily both pairs will be nearly equal.
   - The triple $\hat{\bar{\mu}}(113)$, $\hat{\lambda}(110)$, $\hat{\bar{\tau}}^2(110)$ and the triple $\hat{\bar{\mu}}(123)$, $\hat{\lambda}(120)$, $\hat{\bar{\tau}}^2(120)$ both summarize the performance of the vertical splitter acting on the two inch lumber. Ordinarily both triples will be nearly equal.
d) Determine the need for maintenance on individual saws by referring the foregoing estimates to control charts.

Brown [4] and Whitehead [9] proposed control charts for monitoring thickness variations between and within boards generated by a saw. Neither author took into account how the method of sampling the boards and the locations of the thickness measurement sites affect the joint distribution of the thickness errors. For a detailed critique, see Beran [3]. We will see in Section 5 that the assumption of normally distributed errors is also questionable.

4.2. Priorities in sawmill improvement

The insight we have gained into how random errors accumulate through resawing helps to determine priorities for sawmill improvements. In particular:

- A vertical splitter under good control divides sawing errors in the input board nearly equally between the two offspring boards, while adding small errors of equal magnitude to each offspring. Comparing the spatially averaged variance estimates for sawing codes 110, 120 in Table I (horizontal resaw) with those for sawing codes 130, 140 in Table II (vertical resaw) indicates the potential superiority of a vertical resaw in controlling sawing variability for *both* sets of offspring boards.
- The variability of a horizontal resaw affects both lower and upper offspring boards. However, only the upper offspring boards are affected by errors in previous sawing operations, making sawing errors in upper offspring more variable. Thus, the target thickness for upper offspring boards in a horizontal resaw should be set larger to offset this greater variability. The next section describes how.

4.3. Setting target thickness for green lumber

An ideal procedure for setting the target thickness in each sawing code would take into simultaneous account the variability and mean losses introduced at each point of each board. Because sawing errors at measurement points are correlated and are not quite normally distributed (see Section 5), this is not easily accomplished. As a practical substitute, we may choose target thickness for green lumber so as to control the probability of insufficient thickness at each individual measurement point. By target thickness, we intend the mean thickness to be achieved at that point. A procedure for so doing will be illustrated below for 1 inch boards obtained by two successive horizontal resaws of 4 inch lumber from headrig 1. The argument given here assumes that the positional means and variances do not depend on the measurement position and that the thickness errors are normally distributed.

Suppose that the minimal required thickness of the nominally 1 inch green lumber is .86 inch. This figure recognizes the market definition of 1 inch finished dry lumber and includes an allowance for shrinkage in drying and for planer loss. Let θ denote the target mean thickness. The observed lumber thickness is generically $\theta + E$, where E is the random sawing error. For the present purpose, we take E to be normally distributed with mean 0 and variance σ^2 that depends on the particular 1 inch sawing code. We assume that the process is under good control so that the target mean thickness can be set accurately. The requirement that the observed lumber thickness be at least .86 with probability c is expressed by

$$c = \mathrm{P}(E + \theta \geq .86) = \mathrm{P}(-E \leq \theta - .86) = \Phi[\sigma^{-1}(\theta - .86)],$$

where Φ denotes the standard normal cumulative distribution function. Hence the target mean thickness is

$$\theta = .86 + \sigma\Phi^{-1}(c).$$

Let α be the average loss in thickness incurred at each measurement position by a horizontal resaw. Let $\sigma^2(k)$ be the variance at each measurement position in sawing code k. As we have seen, this variance depends considerably upon the sawing code. The foregoing calculation of target thickness for 1 inch lumber generates Table III of target thicknesses by sawing code.

Table III. Target mean thicknesses in sawing codes
to achieve in each 1 inch code a measured thickness $\geq .86$ with probability c

Sawing Code k	Target Thickness
111	$.86 + \Phi^{-1}(c)\sigma(111)$
112	$.86 + \Phi^{-1}(c)\sigma(112)$
121	$.86 + \Phi^{-1}(c)\sigma(121)$
122	$.86 + \Phi^{-1}(c)\sigma(122)$
110	$1.72 + \alpha + \Phi^{-1}(c)[\sigma(111) + \sigma(112)]$
120	$1.72 + \alpha + \Phi^{-1}(c)[\sigma(121) + \sigma(122)]$
100	$3.44 + 3\alpha + \Phi^{-1}(c)[\sigma(111) + \sigma(112) + \sigma(121) + \sigma(122)]$

For the 4 inch data from headrig 1, reasonable values (see Section 3.3 and Table I) are $\alpha = .14$, $\sigma^2(111) = .00339$, $\sigma^2(112) = \sigma^2(122) = .00073$, and $\sigma^2(121) = .00131$. When $c = .95$, then $\Phi^{-1} = 1.645$. With these choices, Table III yields

Table IV. Target mean thicknesses in observed sawing codes
to achieve in each 1 inch code a measured thickness $\geq .86$ with probability .95

Sawing Code	111	112	121	122	110	120	100
Target mean thickness	.96	.90	.92	.90	2.00	1.96	4.10

These target thicknesses ensure, under the assumptions specified, that about 95% of the boards measured at position j in sawing codes 111, 112, 121, 122 will be at least .86 inch thick at that point. Two remarks:

- Non-normality in the distribution of actual board distributions entails that these target thicknesses may be off a bit. After collecting sufficient lumber thickness data, the normal distribution could be replaced by the observed empirical distribution of thickness errors.
- It is not necessarily a good policy to require 95% reliability in all four horizontal resawing codes. It may be more economical to seek higher reliability in the least variable codes, 112 and 122, moderate reliability in code 121, and lower reliability in code 111. Table III is easily modified to handle any desired pattern of reliabilities once the economic calculations have been made.

5. POSITIONAL STATISTICAL ANALYSES

The main finding of this paper is that fitting a physically based statistical model to measured board thicknesses provides sound quantitative insight into the propagation of thickness errors through lumber resawing. The model thereby enables effective quality control of individual saws, determination of target mean thicknesses by sawing code, and setting priorities for sawmill improvements. The relatively simple data analysis in Section 3 advanced these goals. This section outlines results from more detailed statistical analyses of the random and systematic thickness errors that occur in each measurement position during the sawing process. These results challenge two assumptions made in earlier studies of lumber thickness data: that the thickness errors are normally distributed and that the thickness errors satisfy a one-way random effects model (cf. Warren [8]). In addition, the positional analyses provide effective estimation of departures from mean board flatness in each sawing code. This is useful for identifying misaligned saws.

5.1. Marginal distribution of the random thickness errors

The data for sawing code k generates the residuals $r_{ij}(k) = x_{ij}(k) - \hat{\mu}_j(k)$ for $i \in I(k)$, $1 \leq j \leq 8$. Qnorm plots of these residuals reveal notable qualitative differences among the sawing codes. Sawing error distributions whose tails are fatter than those of a normal distribution are indicated by the residual plots for the following codes: in the 4 inch lumber sequences, codes 100, 110, 111, 113, 114 and their counterparts from headrig 2; in the 2 inch lumber sequences, all sawing codes. The shape of the residual plot for the parent headrig is inherited by the plots for the offspring sawing codes listed above. Residual plots for the other offspring codes exhibit no notable departures from normality.

The error accumulation model of Section 3 suggests an interpretation for this pattern of roughly normal and strikingly non-normal residual plots. The thickness errrors introduced into the 4 inch boards by headrig 1 or 2 are non-normally distributed; the errors introduced by the first and second horizontal resaws are roughly normal; the errors introduced by the vertical splitter are also roughly normal. However, when strikingly non-normal errors are added to normal errors, the result is not normally distributed. Thus, equations (2) and (5) explain how the non-normality of the errors created by headrig 1 or 2 is transmitted to some subsequent sawing codes but not to others. Similar reasoning explains the inheritance of non-normality seen in the residual plots for the 2 inch sawing sequence.

5.2. Covariance matrix of lumber thicknesses

In some earlier studies of lumber thickness errors (cf. Warren [8]), statistical methods developed for the one-way random effects model (cf. Scheffé [7]) were used to analyze thickness measurements made on boards within a given sawing code. Since the one-way random effects model is a severe restriction of the general multivariate error model described at the beginning of Section 3, the validity of this assumption is open to question.

In the notation of Section 3, the one-way random effects model specifies that $\mu_j(k) = \bar{\mu}(k)$, $\sigma_j^2(k) = \bar{\sigma}^2(k)$ for $1 \le j \le 8$; and that the 8×8 covariance matrix $\Sigma(k)$ of the error vector $e_i(k) = (e_{i1}(k), e_{i2}(k), \ldots, e_{i8}(k))'$ has the form

$$\Sigma_o(k) = \begin{pmatrix} A & B & B & B & B & B & B & B \\ B & A & B & B & B & B & B & B \\ B & B & A & B & B & B & B & B \\ B & B & B & A & B & B & B & B \\ B & B & B & B & A & B & B & B \\ B & B & B & B & B & A & B & B \\ B & B & B & B & B & B & A & B \\ B & B & B & B & B & B & B & A \end{pmatrix}, \tag{18}$$

where $A > B > 0$ both depend on sawing code k and $A = \sigma^2(k)$. It is customary to write $A = \sigma_w^2(k) + \sigma_b^2(k)$ and $B = \sigma_b^2(k)$, where $\sigma_w^2(k)$ is the variance *within* boards and $\sigma_b^2(k)$ is the variance *between* boards for sawing code k. In essence, the one-way random effects model specifies that the positional means are equal, that the positional variances are equal, and that the correlation between the thickness errors at any two distinct measurement sites is positive and the same.

This last assumption seems unrealistic as a description of dependence among sawing errors in this study. On physical grounds, it appears likely that the correlation between two sawing errors is positive but decreases as the distance between the measurement positions increases. Recall that measurement positions 1 to 4 are sequenced along one edge of a board, that measurement positions 5 to 8 are sequenced along the other edge of the board, and that positions 1 and 5, 2 and 6, 3 and 7, 4 and 8 are facing pairs across the width of the board. Moreover, the distance between adjacent measurement positions along either edge is much greater than the width of the board. Thus, the correlation between errors at sites 1 and 2 should be virtually the same as the correlation between errors at sites 1 and 6, and so forth. These considerations generate the homogeneous covariance matrix model

$$\Sigma_h(k) = \begin{pmatrix} A & B & C & D & E & B & C & D \\ B & A & B & C & B & E & B & C \\ C & B & A & B & C & B & E & B \\ D & C & B & A & D & C & B & E \\ E & B & C & D & A & B & C & D \\ B & E & B & C & B & A & B & C \\ C & B & E & B & C & B & A & B \\ D & C & B & E & D & C & B & A \end{pmatrix}, \tag{19}$$

where $A > E > B > C > D > 0$ each depend on the sawing code k.

Example: Sawing code 111 in the 4 inch lumber sequence. Let

$$\hat{\mu}(k) = n^{-1}(k) \sum_{i \in I(k)} x_i(k)$$

denote the vector of positional sample means. The sample covariance matrix

$$\hat{\Sigma}(k) = (n(k)-1)^{-1} \sum_{i \in I(k)} (x_i(k) - \hat{\mu}(k))(x_i(k) - \hat{\mu}(k))'$$

has for sawing code 111 the value

$$100\hat{\Sigma}(111) = \begin{pmatrix} .300 & .238 & .142 & .086 & .241 & .190 & .091 & .023 \\ .238 & .406 & .167 & .131 & .145 & .288 & .055 & -.020 \\ .142 & .167 & .369 & .110 & .014 & .072 & .165 & .026 \\ .086 & .131 & .110 & .232 & .000 & .059 & .047 & .141 \\ .241 & .145 & .014 & .000 & .397 & .279 & .153 & .069 \\ .190 & .288 & .072 & .059 & .279 & .363 & .154 & .062 \\ .091 & .055 & .165 & .047 & .153 & .154 & .232 & .131 \\ .023 & -.020 & .026 & .141 & .069 & .062 & .131 & .236 \end{pmatrix}. \quad (20)$$

To fit the one-way random effects covariance matrix to the data, we estimate A in (18) by averaging the diagonal elements of $\hat{\Sigma}(111)$ and estimate B in (18) by averaging the off-diagonal elements of $\hat{\Sigma}(111)$. This procedure, which is consistent under the one-way random effects model, yields the estimate

$$100\hat{\Sigma}_o(111) = \begin{pmatrix} .317 & .116 & .116 & .116 & .116 & .116 & .116 & .116 \\ .116 & .317 & .116 & .116 & .116 & .116 & .116 & .116 \\ .116 & .116 & .317 & .116 & .116 & .116 & .116 & .116 \\ .116 & .116 & .116 & .317 & .116 & .116 & .116 & .116 \\ .116 & .116 & .116 & .116 & .317 & .116 & .116 & .116 \\ .116 & .116 & .116 & .116 & .116 & .317 & .116 & .116 \\ .116 & .116 & .116 & .116 & .116 & .116 & .317 & .116 \\ .116 & .116 & .116 & .116 & .116 & .116 & .116 & .317 \end{pmatrix}. \quad (21)$$

Note that matrix (21) is positive definite.

Inspection suggests that $\hat{\Sigma}(111)$ in (20) lacks the structure of the estimated covariance matrix $\hat{\Sigma}_o(111)$ in (21) obtained under the one-way random effects model. A formal likelihood ratio test of the normal one-way random effects model versus the normal general multivariate model strongly rejects the former. Morrison [6], p. 250 gives the procedure. This finding makes questionable Warren's [8] use of the one-way random effects model in his pioneering statistical study of lumber thickness measurements. It casts doubt on later quality control procedures for lumber thickness that rely on this model (cf. Whitehead [9]).

Sawing code 111 labels the top offspring of top offspring through two horizontal resaws. As discussed in Section 5.1, the residual plot for code 111 points to some non-normality in the thickness errors. A nonparametric bootstrap version of the likelihood ratio test, constructed by the method developed in Beran [1], still rejects the one-way random effects model. The bootstrap method obtain the critical value for the test statistic by resampling from the empirical distribution of the data after that is adjusted by linear transformation to have sample covariance matrix $\hat{\Sigma}_o(111)$.

Having seen that the one-way random effects model does not fit the sawing code 111 measurements, we consider the more general homogenous covariance matrix (19). To fit this covariance structure to the data, we estimate A, B, C, D, E by averaging over the relevant entries in $\hat{\Sigma}(111)$. This procedure, which is consistent under the homogeneous covariance matrix model, yields the estimate

$$100\hat{\Sigma}_h(111) = \begin{pmatrix} .317 & .134 & .079 & .044 & .209 & .134 & .079 & .044 \\ .134 & .317 & .134 & .079 & .134 & .209 & .134 & .079 \\ .079 & .134 & .317 & .134 & .079 & .134 & .209 & .134 \\ .044 & .079 & .134 & .317 & .044 & .079 & .134 & .209 \\ .209 & .134 & .079 & .044 & .317 & .134 & .079 & .044 \\ .134 & .209 & .134 & .079 & .134 & .317 & .134 & .079 \\ .079 & .134 & .209 & .134 & .079 & .134 & .317 & .134 \\ .044 & .079 & .134 & .209 & .044 & .079 & .134 & .317 \end{pmatrix}. \qquad (22)$$

Note that matrix (22) is positive definite.

Inspection indicates that $\hat{\Sigma}_h(111)$ is closer than $\hat{\Sigma}_o(111)$ to the sample covariance matrix $\hat{\Sigma}(111)$ in (20). In recent unpublished work, Liao [6] has shown that $\hat{\Sigma}_h(111)$ lies within various 95% nonparametric bootstrap confidence sets for the unknown covariance matrix $\Sigma(111)$. Thus, the physical understanding that motivates the homogeneous covariance matrix (19) is supported by analysis of the data for sawing code 111. The next subsection will use the homogeneous covariance matrix model in studying whether the positional mean lumber thicknesses are equal (the ideal) or exhibit a trend that reflects saw misalignments.

5.3. Mean vector of lumber thicknesses

In a well controlled sawmill, the positional mean lumber thicknesses will be very nearly equal. The analysis described in this section provides a way of determining on the study data whether this is the case. For this purpose, we relabel the positional means in sawing code k as a two-way layout in which the two factors describe the geometrical location of each thickness measurement position. Let

$$\nu_{1j}(k) = \mu_j(k), \qquad \nu_{2j}(k) = \mu_{j+4}(k), \qquad 1 \le j \le 4.$$

The $\{\nu_{1j}(k)\colon 1 \le j \le 4\}$ report the positional means at the four measurement positions along the first edge of a board while the $\{\nu_{2j}(k)\colon 1 \le j \le 4\}$ report the positional means at the four facing measurement positions along the second edge of a board.

As in two-way analysis of variance, consider five possible submodels for these positional means:

- *Unrestricted.* The $\{\nu_{ij}(k)\colon 1 \le i \le 2, 1 \le j \le 4\}$ satisfy $\nu_{ij} = c + a_i + b_j + g_{ij}$ with $a_+ = b_+ = g_{i+} = g_{+j} = 0$. The subscript $+$ designates summation over all values of the subscript it replaces. For instance, $a_+ = \sum_{i=1}^{2} a_i$ and $g_{i+} = \sum_{j=1}^{4} g_{ij}$.
- *Additive.* The $\{\nu_{ij}(k)\}$ satisfy $\nu_{ij} = c + a_i + b_j$ with $a_+ = b_+ = 0$.
- *Wedge.* The $\{\nu_{ij}(k)\}$ satisfy $\nu_{ij} = c + a_i$ with $a_+ = 0$.
- *Ripple.* The $\{\nu_{ij}(k)\}$ satisfy $\nu_{ij} = c + b_j$ with $b_+ = 0$.
- *Flat.* The $\{\nu_{ij}(k)\}$ satisfy $\nu_{ij} = c$.

The constants c, $\{a_i\}$, $\{b_j\}$ and $\{g_{ij}\}$ depend on the sawing code k. The label for each submodel describes the mean shape of a board whose positional means satisfy that submodel. For a sawmill manager, the ideal shape is Flat.

Which submodel best describes what is happening in the data? We approach this question by assuming that the general multivariate model of Section 3 holds. Because many of the sawing codes observed contain 25 or fewer boards, it is not possible to estimate the unrestricted covariance matrix $\Sigma(k)$ accurately. We therefore impose on the general model the restriction that $\Sigma(k)$ has the homogeneous structure (19). The reasonability of this assumption was examined in Section 5.2. Fitting the homogeneous covariance structure reduces the number of covariance matrix parameters to be estimated from 36 in a general 8×8 covariance matrix to the 5 parameters A, B, C, D, E in (19).

As competing fits to the mean data, we consider the generalized least squares fits to each of the five submodels specified above. The generalized least squares estimator of the mean vector $\mu(k)$ for submodel S is

$$\hat{\mu}_S(k) = \operatorname*{argmin}_{\mu \in S} (\hat{\mu}(k) - \mu)' \hat{\Sigma}_h^{-1}(k)(\hat{\mu}(k) - \mu).$$

The normalized quadratic risk of the estimator $\hat{\mu}_S(k)$ is

$$(n(k)/8)\mathrm{E}[(\hat{\mu}_S(k) - \mu(k))' \Sigma_h^{-1}(k)(\hat{\mu}_S(k) - \mu(k))], \tag{23}$$

the expectation being computed *under the Unrestricted submodel* that puts no limitations on the components of $\mu(k)$. Note that the risk of the estimator $\hat{\mu}(k)$ is 1. We seek a submodel estimator that achieves smaller risk through variance-bias tradeoff.

A surrogate for the unknown risk in (23) is the estimated risk of $\hat{\mu}_S(k)$:

$$(n(k)/8)(\hat{\mu}_S(k) - \hat{\mu}(k))' \hat{\Sigma}_h^{-1}(k)(\hat{\mu}_S(k) - \hat{\mu}(k)) + 2\dim(S) - 8. \tag{24}$$

Here $\dim(S)$ is the dimension of the space to which submodel S restricts the mean vector $\mu(k)$. By extension of arguments in Beran [2], it is seen that the estimated risk (24) converges to the true risk (23) of the submodel estimator $\hat{\mu}_S(k)$, under

Unrestricted model asymptotics in which the number of measurement positions and the number of sampled boards both tend to infinity. Thus, the submodel fit that has smallest estimated risk approximates, in risk, the submodel fit that has the smallest (unknown) risk. This result provides a rationale for relying on the submodel fit with smallest estimated risk.

Example: Sawing code 111 in the 4 inch lumber sequence. Table V reports the estimated risks for each of the five submodel fits described above. The clear winner with smallest estimated risk is the Wedge submodel fit, in which the estimated mean thickness on the first edge is .831 inch while the estimated mean thickness on the second edge is .845 inch. This departure from flatness points to slight saw misalignment. The statistical technique described in this section—essentially a signal recovery technique—is an effective way of scrutinizing positional averages to check saw alignments.

Table V. Estimated risks of submodel fits to mean thicknesses in code 111

Submodel	Full	Additive	Wedge	Ripple	Flat
Estimated risk	1.00	.42	.27	1.28	1.12

6. DISCUSSION

This paper has developed and validated on study data a multivariate statistical model for lumber thickness measurements. The model is physically based, expressing an idealized physical understanding of horizontal resaws and of vertical splitter resaws and it quantifies how sawing errors accumulate through resawing operations. The model thereby enables estimation from board thickness measurements of how much thickness error, systematic or random, is contributed by each resawing operation. Section 3 described how spatially averaged sample means and sample variances of board thicknesses can be analyzed to quantify the performance characteristics of each separate resaw. Implications for process control—monitoring the performance of each sawing operation, determining priorities for sawmill improvement, and setting target thicknesses for green lumber in each sawing code—were developed in Section 4.

Section 5 gave techniques for positional analysis of lumber thickness measurements. Proposed and validated on the study data was the idea that sawing errors along a board are positively correlated, the amount of correlation decreasing as the distance between the thickness measurement sites increases. The positional mean thicknesses on boards in each sawing code form a two-way layout, the two factors being the edge and the measurement site along that edge. Unlike classical two-way layouts, positional thickness measurements are positively correlated as just described. A model selection technique based on estimated risks determined which of five submodel fits to the mean thicknesses was most trustworthy in approximating the unknown mean thicknesses. As an example, the mean shape of boards in sawing code 111 was found to be slightly wedged, indicating a small saw misalignment. The positional analyses of section 5 thus provide a more detailed way of monitoring the performance of each sawing operation.

ACKNOWLEDGEMENTS

Conversations with Dean Huber of the U.S. Forest Service in San Francisco greatly assisted the author's understanding of sawmill operations and of the manner in which the study data was collected. Comments by a referee prompted clearer exposition of this understanding.

REFERENCES

1. Beran, R. Simulated power functions. *Annals of Statistics* 1986 **14**:151–173.
2. Beran, R. REACT scatterplot smoothers: superefficiency through basis economy. *Journal of the American Statistical Association* 2000 **95**:155–171.
3. Beran, R. Control charts for monitoring lumber thickness. Unpublished technical report.
4. Brown, TD. Determining lumber target sizes and monitoring sawing accuracy. *Forest Products Journal* 1979; **29**:48–54.
5. Liao, S. Personal communication.
6. Morrison, DF. *Multivariate Statistical Methods* (second edition). McGraw: New York, 1976.
7. Scheffé, H. *The Analysis of Variance*. John Wiley: New York, 1959.
8. Warren, WG. How to calculate target thickness for green lumber. *Information Report VP-X-112*, Canadian Forestry Service, 1973.
9. Whitehead, JC. Procedures for developing a lumber-size control system. *Information report VP-X-184*, Canadian Forestry Service, 1978.