Statistical Modelling for Process Control in the Sawmill Industry

Rudolf Beran*

Department of Statistics, University of California, Davis
Davis CA 95616, USA
beran@wald.ucdavis.edu

SUMMARY

In soft-wood sawmills in the western U.S., the green lumber end-product is the result of several distinct sawing operations: an initial breakdown of large diameter logs by a headrig yields boards that are subsequently resawn one or more times by secondary saws of various types. 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.

Shrinkage of green lumber is well-understood. However, because both systematic errors and random errors in thickness accumulate through a sequence of resawing operations, it has not been clear how to separate out the performance of secondary sawing machines. In a pilot study, boards selected “at random” as they came off a headrig were followed through one of more resawings. Initially and at each subsequent stage of the processing, the thickness of each green board was measured at standardized points along each edge. Using the data, this paper develops and validates a physically-based statistical model for the accumulation of systematic and random errors in resawing operations. The model quantifies how thickness variability in a batch of lumber accumulates through a sequence of resawing operations and thereby enables estimation of how much each resawing operation contributes to thickness errors in the end-product. Implications for process control in sequential resawing operations are noted.

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 many western soft-wood mills, the green lumber end-product is the result of several distinct sawing operations: an initial breakdown of the logs by a headrig saw yields boards that are subsequently resawn one or more times by secondary saws. Saws vibrate and may be slightly misaligned. Variability in each of the sequential sawing operations thus contributes to irregularities in thickness of the final green lumber. Because of the interactions among sequential sawing operations, it has not been clear how to evaluate accurately the performance of secondary sawing machines.

Resolution of this problem was the major motivation for the work reported in this paper. Lumber thickness data gathered at a redwood mill in northern California was made available for this study by the U.S. Forest Service. Boards selected “at random” as they came off a headrig were followed through one or more resawings. Initially and at each subsequent stage of the processing, the thickness of each board produced was measured at 8 standardized points, 4 along each edge in facing pairs. The data was coded and served as the empirical basis for this study.

Principal goals of the study were as follows:

• 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 note the implications of the foregoing work for process control in sequential resawing.
• To relate the model to assumptions made in earlier studies of lumber thickness data, such as the one-way random effects model with normally distributed errors (cf. Warren [7]).

2. THE LUMBER THICKNESS DATA

The lumber thickness data analyzed in this paper was collected during a sawmill improvement study carried out at a northern California redwood mill in cooperation with the U.S. Forest Service. Two types of resawing were used in the mill. In a horizontal resaw, the board being cut is pressed flat against a horizontal reference plane. The saw cuts horizontally down the length of the board parallel to the reference
plane, dividing it into an upper offspring board and a lower offspring board. In a vertical splitter resaw, the board being cut is stood vertically on one edge, between two sets of spring loaded rollers. The saw cuts vertically down the length of the board, dividing it into a left offspring board and a right offspring board.

2.1. Design of the study

In the study, two samples of boards were drawn from each of two distinct sequences of sawing operations.

**Sawing sequence for 4 inch lumber.** Boards of nominal 4 inch thickness coming off a headrig 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 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. 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 $k$. 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 are, respectively,

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

The spatially averaged sample mean and sample variance in sawing code $k$ are defined as

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

The initial analysis of how sawing errors propagate 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. 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 Table I are puzzling at first sight because they vary substantially among the sawing codes for 1 inch lumber and 2 inch lumber. A basic problem is to understand why this happens and to estimate quantitatively the random variability in thickness that is contributed by each resawing operation. In Section 3, we construct a statistical model for error propagation through resawing that expresses an idealized physical picture of the process and explains the patterns observed in Tables I and II.

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

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

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

<table>
<thead>
<tr>
<th>Sawing Code k</th>
<th>$\hat{\mu}(k)$</th>
<th>$100\hat{\sigma}^2(k)$</th>
<th>Sawing Code k</th>
<th>$\hat{\mu}(k)$</th>
<th>$100\hat{\sigma}^2(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.91</td>
<td>.222</td>
<td>200</td>
<td>1.89</td>
<td>.075</td>
</tr>
<tr>
<td>130</td>
<td>0.91</td>
<td>.050</td>
<td>210</td>
<td>0.90</td>
<td>.024</td>
</tr>
<tr>
<td>140</td>
<td>0.89</td>
<td>.077</td>
<td>240</td>
<td>0.88</td>
<td>.036</td>
</tr>
</tbody>
</table>

3. A MULTIVARIATE STATISTICAL MODEL

A structured statistical model for random and systematic errors in lumber thickness during sequential sawing is derived in this section. The model expresses an
idealized physical picture of sawing operations. Success of this statistical model in explaining observed patterns of variability in lumber thickness will be evidence for the underlying physical model.

The statistical model makes the following physical assumptions:

• All saws operate 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 flat against a horizontal reference plane. The saw cuts horizontally down the length of the board parallel to the reference plane, dividing it into an upper offspring board and a lower offspring board.

• In a vertical splitter resaw, the board being cut is stood vertically on one edge, between two sets of spring-loaded rollers. The saw cuts vertically down the length of the board, dividing it into a left offspring board and a right offspring board. Spring tension on both sides is ideally close to equal.

• Thickness is measured at the same 4 pairs of facing edge points throughout the sawing sequence.

• Measurement errors are small 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 taking care in measuring thickness. The last assumption reflects separation in the production times of the boards that were selected “at random” for the study.

3.1. A general multivariate model

Let $e_{ij}(k)$ denote the random thickness error at measurement position $j$ on board $i$ in sawing code $k$. Position pairs $(1, 5), (2, 6), (3, 7), (4, 8)$ face across the board. For each $k$, we assume that $x_{ij}(k) = \mu_j(k) + e_{ij}(k)$ for $i \in I(k)$ and $1 \leq j \leq 8$. Here $\mu_j(k)$ denotes the mean thickness at position $j$. Let $x_i(k) = (x_{i1}(k), x_{i2}(k), \ldots, x_{i8}(k))'$ be the measurements on board $i$ arranged as a column vector. Similarly let $\mu(k) = (\mu_1(k), \mu_2(k), \ldots, \mu_8(k))'$ be the vector of positional means and let $e_i(k)$ be the analogously defined vector of positional random thickness errors on board $i$ within code $k$.

The first part of the statistical model for thickness measurements asserts that, within each sawing code $k$,

$$x_i(k) = \mu(k) + e_i(k)$$

and that the $\{e_i(k): i \in I(k)\}$ are independent, identically distributed random vectors.
with mean vector zero and unknown covariance matrix $\Sigma(k)$. Then $E(x_i(k)) = \mu(k)$ and $\text{Cov}(x_i(k)) = \Sigma(k)$. The diagonal elements of $\Sigma(k)$ are the positional variances $\{\sigma_j^2(k): 1 \leq j \leq 8\}$.

The second part of the statistical model, described in the next subsections, draws on the idealized physical model 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 and will be seen to explain the complex patterns of variability reported in Tables I and II.

### 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 modeling for 4 inch lumber from headrig 2 is completely analogous.

For every board number $i$ in sawing code $k$,

\[
\begin{align*}
e_i(110) &= e_i(100) - e_i(120) \\
e_i(111) &= e_i(110) - e_i(112) \\
e_i(121) &= e_i(120) - e_i(122).
\end{align*}
\]

(2)

Here $e_i(100)$ is the vector of random sawing errors introduced into the thickness of board $i$ by headrig 1 and $e_i(120)$ is the random sawing error introduced into the thickness of bottom offspring board $i$ by the first horizontal resaw. Because the board being resawn is pressed flat against a reference plane, the headrig error does not get transmitted to the bottom offspring board. However, as the mathematically necessary relation in the top line of (2) indicates, the random sawing error $e_i(110)$ in the thickness of top offspring board $i$ depends on both the headrig error and the horizontal resaw error. The physical assumptions imply that $e_i(100)$ and $e_i(120)$ are independent random vectors.

Because the rationale behind the second and third lines in (2) is completely analogous, the random vectors $e_i(110)$ and $e_i(112)$ are also independent, as are the random vectors $e_i(120)$ and $e_i(122)$. Equation (2) plus the assumption of stability in the performance of the horizontal resaw yields the following relationships for the positional variances:

\[
\begin{align*}
\sigma_j^2(110) &= \sigma_j^2(100) + \sigma_j^2(120) \\
\sigma_j^2(111) &= \sigma_j^2(110) + \sigma_j^2(112) \\
\sigma_j^2(121) &= \sigma_j^2(120) + \sigma_j^2(122) \\
\sigma_j^2(112) &= \sigma_j^2(122).
\end{align*}
\]

(3)
The physical picture of the vertical splitter resaw indicates that, for every board number $i$ in sawing code $k$,

$$
e_i(113) = \lambda(110)e_i(110) + d_i(110)$$
$$e_i(114) = (1 - \lambda(110))e_i(110) - d_i(110)$$
$$e_i(113) = \lambda(120)e_i(120) + d_i(120)$$
$$e_i(114) = (1 - \lambda(120))e_i(120) - d_i(120).$$

Here $\lambda(k)$ is a constant between 0 and 1. The value $\lambda(k) = 1/2$ means that the splitter divides the random thickness error in the parent board equally among the left and right offspring boards. This ideal will happen only if the springs on both sides of the vertical splitter have perfectly equal tension. For an actual splitter, $\lambda(k)$ is an unknown constant between 0 and 1 to be estimated from the data. The assumption of stability in the operation of the vertical splitter implies that

$$\lambda(110) = \lambda(120).$$

The error vector $d_i(k)$ represents random variability in the path of the vertical splitter saw as it rips through board $i$ in sawing code $k$. 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 $T(k)$ are denoted by $\{\tau^2_j(k): 1 \leq j \leq 8\}$. The random vectors $e_i(110)$ and $d_i(110)$ are physically and therefore statistically independent, as are $e_i(120)$ and $d_i(120)$. It follows from (4) and stability of the vertical splitter resawing process that

$$\sigma^2_j(113) = \lambda^2(110)\sigma^2_j(110) + \tau^2_j(110)$$
$$\sigma^2_j(114) = (1 - \lambda(110))^2\sigma^2_j(110) + \tau^2_j(110)$$
$$\sigma^2_j(123) = \lambda^2(120)\sigma^2_j(120) + \tau^2_j(120)$$
$$\sigma^2_j(124) = (1 - \lambda(120))^2\sigma^2_j(120) + \tau^2_j(120)$$
$$\tau^2_j(110) = \tau^2_j(120).$$

For process control purposes, the key variances in equations (3) and (6) are $\sigma^2_j(k)$ for $k = 100, 120, 112, 122$ and $\tau^2_j(k)$ for $k = 110, 120$. Indeed, at measurement position $j$, $\sigma^2_j(100)$ measures variability introduced by headrig 1; $\sigma^2_j(120)$ measures additional variability introduced by the first horizontal resaw into 2 inch lumber; $\sigma^2_j(112) = \sigma^2_j(122)$ measures additional variability introduced by the second horizontal resaw into 1 inch lumber; $\tau^2_j(110) = \tau^2_j(120)$ measures additional variability introduced by the vertical splitter resaw into 1 inch lumber. Moreover, $\lambda(110) = \lambda(120)$ measures the fraction of the parent board thickness error that the vertical splitter transmits to the offspring boards on the 3 side. Thus, the random errors introduced by the
various saws involved in the sequential resawing can be monitored by estimating these parameters.

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,

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

(7)

Stability of the sawing process entails the equalities

$$
\begin{align*}
\alpha_j(110) &= \alpha_j(120) \\
\beta_j(110) &= \beta_j(120).
\end{align*}
$$

(8)

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 $\{\mu_j(k): 1 \leq j \leq 8\}$ and the positional variances $\{\sigma_j^2(k): 1 \leq j \leq 8\}$ will vary only slightly with $j$. In such circumstances, the spatially averaged mean and variance, defined as

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

may serve as adequate summaries. The values of $\tilde{\mu}(k)$ and $\tilde{\sigma}^2(k)$ are estimated by the spatially averaged sample mean $\tilde{\mu}(k)$ and sample variance $\tilde{\sigma}^2(k)$ 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 $\tilde{\alpha}(k)$ denote the average of the $\{\alpha_j(k): 1 \leq j \leq 8\}$ and let $\tilde{\beta}(k)$ similarly
denote the average of the \( \{ \beta_j(k) : 1 \leq j \leq 8 \} \). From equations (7) and (8),

\[
\begin{align*}
\bar{\mu}(100) &= \bar{\mu}(110) + \bar{\mu}(120) + \bar{\alpha}(100) \\
\bar{\mu}(110) &= \bar{\mu}(111) + \bar{\mu}(112) + \bar{\alpha}(110) \\
\bar{\mu}(120) &= \bar{\mu}(121) + \bar{\mu}(122) + \bar{\alpha}(120) \\
\bar{\mu}(110) &= \bar{\mu}(113) + \bar{\mu}(114) + \bar{\beta}(110) \\
\bar{\mu}(120) &= \bar{\mu}(123) + \bar{\mu}(124) + \bar{\beta}(120) \\
\bar{\alpha}(110) &= \bar{\alpha}(120) \\
\bar{\beta}(110) &= \bar{\beta}(120). 
\end{align*}
\]  

(9)

Let \( \bar{\tau}^2(k) \) be the average of the \( \{ \tau^2_j(k) : 1 \leq j \leq 8 \} \). The equations of Section 3.2 imply that

\[
\begin{align*}
\bar{\sigma}^2(110) &= \bar{\sigma}^2(100) + \bar{\sigma}^2(120) \\
\bar{\sigma}^2(111) &= \bar{\sigma}^2(110) + \bar{\sigma}^2(112) \\
\bar{\sigma}^2(121) &= \bar{\sigma}^2(120) + \bar{\sigma}^2(122) \\
\bar{\sigma}^2(112) &= \bar{\sigma}^2(122) 
\end{align*}
\]  

(10)

and that

\[
\begin{align*}
\bar{\sigma}^2(113) &= \lambda^2(110)\bar{\sigma}^2(110) + \bar{\tau}^2(110) \\
\bar{\sigma}^2(114) &= (1 - \lambda(110))^2\bar{\sigma}^2(110) + \bar{\tau}^2(110) \\
\bar{\sigma}^2(123) &= \lambda^2(120)\bar{\sigma}^2(120) + \bar{\tau}^2(120) \\
\bar{\sigma}^2(124) &= (1 - \lambda(120))^2\bar{\sigma}^2(120) + \bar{\tau}^2(120) \\
\bar{\tau}^2(110) &= \bar{\tau}^2(120) \\
\lambda(110) &= \lambda(120). 
\end{align*}
\]  

(11)

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 (9) through (11). Fitting the first five equations in (9) to Table I yields the following estimated mean sawing losses for resawn lumber from headrigs 1 and 2:

\[
\begin{align*}
\hat{\alpha}(100) &= .14 \\
\hat{\alpha}(110) &= .13 \\
\hat{\alpha}(120) &= .14 \\
\hat{\beta}(110) &= .10 \\
\hat{\beta}(120) &= .10 
\end{align*}
\]

and

\[
\begin{align*}
\hat{\alpha}(200) &= .14 \\
\hat{\alpha}(210) &= .13 \\
\hat{\alpha}(220) &= .14 \\
\hat{\beta}(210) &= .09 \\
\hat{\beta}(220) &= .10. 
\end{align*}
\]

Note that the equalities predicted in the last two equations of (9) are nearly satisfied up to .01 inch by these estimates. 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 (but see Section 5.3 for a closer assessment).

From Table I, \( \hat{\sigma}^2(100) = 0.248 \times 10^{-2} \) while \( \hat{\sigma}^2(200) = 0.241 \times 10^{-2} \). Thus, the random sawing errors from headrig 2 in the 2 inch lumber experiment have essentially the same variability as those from headrig 1.

According to the four variance equalities in (10), the spatially averaged sample variances for headrig 1 in Table I should satisfy the following approximate equalities:

\[
\begin{align*}
0.281 & \approx 0.248 + 0.047 \\
0.339 & \approx 0.281 + 0.075 \\
0.131 & \approx 0.047 + 0.070 \\
0.075 & \approx 0.070.
\end{align*}
\]

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

\[
\begin{align*}
0.287 & \approx 0.241 + 0.053 \\
0.374 & \approx 0.287 + 0.084 \\
0.127 & \approx 0.053 + 0.072 \\
0.084 & \approx 0.072.
\end{align*}
\]

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.

Applying the first four relations in (11) to the spatially averaged sample variances for headrig 1 yields the estimators

\[
\hat{\lambda}(110) = 0.5 + 0.5[\hat{\sigma}^2(113) - \hat{\sigma}^2(114)]/\hat{\sigma}^2(110)
\]

\[
\hat{\tau}(110) = \hat{\sigma}^2(113) - \hat{\lambda}(110)^2\hat{\sigma}^2(110)
\]

and analogous expressions for \( \hat{\lambda}(120), \hat{\tau}(120) \), and the counterparts for headrig 2. For the 4 inch lumber sequence from headrig 1, Table I yields

\[
\begin{align*}
\hat{\lambda}(110) & = 0.43 \\
\hat{\lambda}(120) & = 0.32
\end{align*}
\]

\[
\begin{align*}
\hat{\tau}(110) & = 0.011 \times 10^{-2} \\
\hat{\tau}(120) & = 0.010 \times 10^{-2}.
\end{align*}
\]

Note that \( \hat{\tau}^2(110) \) nearly equals \( \hat{\tau}^2(120) \) as predicted by (11). The values of \( \hat{\lambda}(110) \) and \( \hat{\lambda}(120) \) are both less than the 1/2 expected from an ideal splitter. This suggests
that the 3 side of the vertical splitter is firmer than the 4 side, because of unequal spring tension on the two sides. Similarly, for the 4 inch lumber sequence from headrig 2,

\[ \hat{\lambda}(210) = .47 \quad \hat{\tau}^2(210) = .000 \times 10^{-2} \]
\[ \hat{\lambda}(220) = .40 \quad \hat{\tau}^2(220) = .007 \times 10^{-2}. \]

Again we see that the 3 side of the vertical splitter is firmer than the 4 side. While it is possible that the vertical splitter was less variable at the time it resawed boards from headrig 2, it is likely that the differing estimates of \( \hat{\tau}^2 \) in the preceding two displays reflect the limitations of 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. An argument just like the one already developed for sawing codes 113, 114, 123, 124 yields the variance relations

\[ \sigma_j^2(130) = \lambda^2(100)\sigma_j^2(100) + \tau_j^2(100) \]
\[ \sigma_j^2(140) = (1 - \lambda(100))^2\sigma_j^2(100) + \tau_j^2(100). \]  

(12)

Here \( \tau_j^2(100) \) measures the variability introduced at position \( j \) by the vertical splitter and \( \lambda(100) \) is the splitting fraction. The equation relating mean thicknesses of parent and offspring boards is

\[ \mu_j(100) = \mu_j(130) + \mu_j(140) + \beta_j(100), \]  

(13)

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

From equations (12) and (13), the spatially averaged means and variances satisfy

\[ \bar{\mu}(100) = \bar{\mu}(130) + \bar{\mu}(140) + \bar{\beta}(100) \]  

(14)

and

\[ \bar{\sigma}^2(130) = \lambda^2(100)\bar{\sigma}^2(100) + \bar{\tau}^2(100) \]
\[ \bar{\sigma}^2(140) = (1 - \lambda(100))^2\bar{\sigma}^2(100) + \bar{\tau}^2(100). \]  

(15)
Fitting equation (14) 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:

\[ \hat{\beta}(100) = .11 \quad \hat{\beta}(200) = .11. \]

As might be expected, these estimated spatially averaged mean thickness losses due to the vertical splitting of the 2 inch lumber are equal and are close to the estimates obtained from the 4 inch lumber resawing sequence.

From Table II, \( \hat{\sigma}^2(100) = .222 \times 10^{-2} \) while \( \hat{\sigma}^2(200) = .075 \times 10^{-2} \). Thus, the random sawing errors from headrig 2 in the 2 inch lumber experiment are considerably smaller than those from headrig 1. Applying the relations (15) to the spatially averaged variances for headrig 1 yields the estimators

\[
\begin{align*}
\hat{\lambda}(100) &= .5 + .5(\hat{\sigma}^2(130) - \hat{\sigma}^2(140))/\hat{\sigma}^2(100) \\
\hat{\tau}(100) &= \hat{\sigma}^2(130) - \hat{\lambda}(100)^2\hat{\sigma}^2(110)
\end{align*}
\]

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

\[
\begin{align*}
\hat{\lambda}(100) &= .44 \\
\hat{\tau}(100) &= .007 \times 10^{-2} \\
\hat{\lambda}(200) &= .42 \\
\hat{\tau}(200) &= .000 \times 10^{-2}.
\end{align*}
\]

Again we see that the 3 side of the vertical splitter is firmer than the 4 side. 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 \( \hat{\tau}^2 \) in the preceding two displays reflect the limitations of 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. Through its physical character, the model and its estimated parameters provide a sound basis for understanding the thickness errors contributed by each resawing operation.

4. IMPLICATIONS FOR PROCESS CONTROL

The modeling and data analysis of Section 3 bear on several aspects of process control in the sawmill industry: the monitoring of individual saws; the assessment of priorities in sawmill improvement projects; and 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 \( \hat{\mu}(100), \hat{\sigma}^2(100) \) summarizes the performance of headrig 1.
- The pair \( \hat{\mu}(120), \hat{\sigma}^2(120) \) summarizes the performance of the first horizontal resaw from 4 inches to 2 inches.
- The pair \( \hat{\mu}(112), \hat{\sigma}^2(112) \) and the pair \( \hat{\mu}(122), \hat{\sigma}^2(122) \) both summarize the performance of the second horizontal resaw from 2 inches to 1 inch. Ordinarily both pairs will be nearly equal.
- The triple \( \hat{\mu}(113), \hat{\lambda}(110), \hat{\tau}^2(110) \) and the triple \( \hat{\mu}(123), \hat{\lambda}(120), \hat{\tau}^2(120) \) 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 [8] 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 thickness measurement sites affects 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 the same
sawing codes 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.
  Only the upper offspring boards are affected by errors in previous sawing opera-
tions, making sawing errors in this code 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 \( \theta \) 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 \( \sigma^2 \) 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

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

where \( \Phi \) denotes the standard normal cumulative distribution function. Hence the
target mean thickness is

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

<table>
<thead>
<tr>
<th>Sawing Code $k$</th>
<th>Target Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>$.86 + \Phi^{-1}(c)\sigma(111)$</td>
</tr>
<tr>
<td>112</td>
<td>$.86 + \Phi^{-1}(c)\sigma(112)$</td>
</tr>
<tr>
<td>121</td>
<td>$.86 + \Phi^{-1}(c)\sigma(121)$</td>
</tr>
<tr>
<td>122</td>
<td>$.86 + \Phi^{-1}(c)\sigma(122)$</td>
</tr>
<tr>
<td>110</td>
<td>$1.72 + \alpha + \Phi^{-1}(c)[\sigma(111) + \sigma(112)]$</td>
</tr>
<tr>
<td>120</td>
<td>$1.72 + \alpha + \Phi^{-1}(c)[\sigma(121) + \sigma(122)]$</td>
</tr>
<tr>
<td>100</td>
<td>$3.44 + 3\alpha + \Phi^{-1}(c)[\sigma(111) + \sigma(112) + \sigma(121) + \sigma(122)]$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Sawing Code</th>
<th>111</th>
<th>112</th>
<th>121</th>
<th>122</th>
<th>110</th>
<th>120</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target mean thickness</td>
<td>.96</td>
<td>.90</td>
<td>.92</td>
<td>.90</td>
<td>2.00</td>
<td>1.96</td>
<td>4.10</td>
</tr>
</tbody>
</table>

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 point of this paper is that fitting a physically based statistical model to the data provides sound insight into the propagation of lumber thickness errors through lumber resawing. The model thereby enables quantitative 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 supported 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 [7]). In addition, the positional analyses provide a way of assessing departures from mean flatness in each sawing code.

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 errors 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. How-
ever, when strikingly non-normal errors are added to normal errors, the result is not normally distributed. Thus, equations (2) and (4) 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 [7]), statistical methods developed for the one-way random effects model (cf. Scheffé [6]) 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
\[ \mu_j(k) = \bar{\mu}(k), \quad \sigma^2_j(k) = \bar{\sigma}^2(k) \] for \( 1 \leq j \leq 8 \), and that the covariance matrix \( \Sigma(k) \) of the error vector \( e_i(k) \) has the form

\[
\Sigma_0(k) = \begin{pmatrix}
A & B & B & B & B & B & B \\
B & A & B & B & B & B & B \\
B & B & A & B & B & B & B \\
B & B & B & A & B & B & B \\
B & B & B & B & A & B & B \\
B & B & B & B & B & A & B \\
B & B & B & B & B & B & A
\end{pmatrix},
\]

where \( A > B > 0 \) both depend on sawing code \( k \) and \( A = \sigma^2(k) \). It is customary to write \( A = \sigma^2_w(k) + \sigma^2_b(k) \) and \( B = \sigma^2_b(k) \), where \( \sigma^2_w(k) \) is the variance within boards and \( \sigma^2_b(k) \) 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 along one edge of a board and that measurement positions 5 to 8 face these along the other edge 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 is virtually the same as the correlation between er-
rors at sites 1 and 6, and so forth. These considerations generate the homogeneous covariance matrix model


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


Inspection suggests that this sample covariance matrix lacks the structure of $\hat{\Sigma}_o(111)$, the estimated covariance matrix 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 [5], p. 250 gives the procedure. 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 critical value is found by resampling from the empirical distribution after that is adjusted by linear transformation to have covariance matrix $\hat{\Sigma}_o(111)$.
To fit the homogenous covariance matrix (17) to the data for sawing code 111, we may estimate $A$, $B$, $C$, $D$, $E$ by averaging over the relevant entries in $\hat{\Sigma}(111)$. This procedure yields the homogeneous estimate

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

Note that this matrix is positive definite. Inspection indicates that $\hat{\Sigma}_h(111)$ is closer than $\hat{\Sigma}_o(111)$ to the sample covariance matrix above. In recent unpublished work, Shanmei Liao, a student of the author, has shown that $\hat{\Sigma}_h(111)$ lies within various 95% nonparametric confidence sets for the unknown covariance matrix $\Sigma(111)$. Thus, the physical understanding that motivates the homogeneous covariance matrix (14) is supported by analysis of the data for sawing code 111.

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 that indicates more explicitly the location of each thickness measurement position. Let

$$\nu_{ij}(k) = \mu_j(k), \quad \nu_{2j}(k) = \mu_{j+4}(k), \quad 1 \leq j \leq 4.$$

The $\{\nu_{ij}(k): 1 \leq j \leq 4\}$ report the positional means at the four measurement positions along the first edge of a board while the $\{\nu_{2j}(k): 1 \leq j \leq 4\}$ 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 $\{\nu_{ij}(k): 1 \leq i \leq 2, 1 \leq j \leq 4\}$ satisfy $\nu_{ij} = c + a_i + b_j + g_{ij}$ with $a_+ = b_+ = g_{i+} = g_{+j} = 0$.
- **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 covariance \( \Sigma(k) \) accurately. We therefore impose on the general model the restriction that \( \Sigma(k) \) has the homogeneous structure (17). 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 \( 8 \times 8 \) covariance matrix to the 5 parameters \( A, B, C, D, E \).

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

\[
\hat{\mu}_S(k) = \arg\min_{\mu \in S} (\hat{\mu}(k) - \mu)' \hat{\Sigma}_h^{-1}(k)(\hat{\mu}(k) - \mu).
\]

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

\[
(n(k)/8)E[(\hat{\mu}_S(k) - \mu)' \hat{\Sigma}_h^{-1}(k)(\hat{\mu}_S(k) - \mu)],
\]

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 (18) is the estimated risk of \( \hat{\mu}_S(k) \):

\[
(n(k)/8)(\hat{\mu}_S(k) - \hat{\mu}(k)' \hat{\Sigma}_h^{-1}(k)(\hat{\mu}_S(k) - \hat{\mu}(k))) + 2 \text{dim}(S) - 8.
\]

Here \( \text{dim}(S) \) 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 (19) converges to the true risk (18) 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 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 modest 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

<table>
<thead>
<tr>
<th>Submodel</th>
<th>Full</th>
<th>Additive</th>
<th>Wedge</th>
<th>Ripple</th>
<th>Flat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated risk</td>
<td>1.00</td>
<td>.42</td>
<td>.27</td>
<td>1.28</td>
<td>1.12</td>
</tr>
</tbody>
</table>

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. Because the model quantifies how sawing errors accumulate through resawing operations, it enables separate estimation 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.
ACKNOWLEDGEMENT

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.

REFERENCES