A Regression-Based Method for Estimating Rip-First Rough-Mill Lumber Yield

Urs Buehlmann
D. Earl Kline
Janice K. Wiedenbeck

Abstract

Estimating yield from lumber cut-up in rip-first rough mills for material management and job costing purposes is uncertain unless simulation models are used. To augment the toolbox for industry practitioners, a novel yield estimation model was derived using linear least squares techniques and data derived from an orthogonal, 2^{20-11} fractional factorial design of resolution V.

The model estimated 450 of 512 cutting bills tested within 1 percent absolute yield. However, cutting bills that do not adhere to the model's framework suffer a larger estimation error. The least squares estimation model thus is a helpful tool in ranking cutting bills that adhere to the model's framework for their expected yield levels and facilitates the selection of part sizes to be included in cutting bills. Further research is needed to make the model useful for a wider range of cutting bills.

In 1966, C. D. Dosker (p. 67 *in* Englerth and Dunmire) contended that "[w]ood is missing out as a raw material in thousands of usages simply because the industry has no information on which a designer, or a user, can determine in advance what the refining costs of wood as a raw material will be." Despite more recent acknowledgments about the industry's lack of ability to obtain "accurate and consistent" (Wiedenbeck and Scheerer 1996:121) yield information to estimate costs and to improve yield (Wiedenbeck and Thomas 1995a), no tool that can easily and reliably obtain this information exists today. The yield estimation matrixes and tables established for different species by several authors (Thomas 1965a, 1965b, 1965c; Schumann and Englerth 1967a, 1967b; Englerth and Schumann 1969; Dunmire 1971; Schumann 1971, 1972, 1973; Wiedenbeck and Thomas 1995a) do not result in accurate yield estimates (Manalan et al. 1980, Yaussy 1986, Thomas et al. 1996, Wiedenbeck and Scheerer 1996, Buehlmann 1998, Buehlmann et al. 1998a, Hoff 2000). Manalan et al. (1980) found differences of as high as 19 percent between yields predicted using the US Department of Agriculture (USDA) Forest Service's FPL 118 yield tables (Englerth and Schumann 1969) and actual yields obtained in rough mills, while Buehlmann et al. (1998a) showed yield differences of as high as 12 percent. At present, the only reliable way for estimating expected yield is to use rough-mill simulation tools (Thomas and Buehlmann 2002). However, performing simulation is time consuming and can require extensive manipulation of the software.

As a response to this lack of reasonable yield estimation tools, Wiedenbeck and Thomas (1995a, 1995b) developed new yield matrixes for rip-first and crosscut-first rough mills. The authors divided 24 cutting bills obtained from members of the Wood Component Manufacturers Association into four groups: (1) short narrow, (2) short wide, (3) long narrow, and (4) long wide. The cutting bills were processed using either rip-first (ROMI RIP; Thomas 1995a and 1995b) or crosscut-first (CORY; Brunner et al. 1990) simulation software. The tests were performed for three different grades: (1) FAS, (2) 1 Common, and (3) 2A Common (National Hardwood Lumber Association 2007). Yields were between a maximum of 82 percent for FAS lumber and short-narrow or long-narrow cutting bills and 29 percent for 2A Common lumber for long-wide cutting bills using rip-first cut-up technology. For crosscut-first technology, simulated yield results ranged from 81 percent

for short-wide cutting bills using FAS lumber to 39 percent for long-wide cutting bills and 2A Common lumber. However, the small number of cutting bills in each category used in the study limits the usefulness of the yield estimates obtained. Also, although the cutting bills were clustered according to similarities in requirements (length and width), large differences in yield persisted within the same group, indicating that small changes in cutting bill requirements can cause large variations in the yield obtained. These findings are consistent with those of Buehlmann (1998) and Buehlmann et al. (1998a), who also found large variability in yield among relatively similar cutting bills.

Three sources contribute to the complexity of the problem of estimating yield: (1) cut-up system related, (2) lumber related, and (3) cutting bill induced. The lumber yield obtained for a given cutting bill changes depending on the cut-up system (rip-first vs. crosscut-first) used (Hall 1978, Harding 1991, Buehlmann et al. 1998b, 1999). Even within the same system, different modes of operation lead to differences in the yield achieved. The lumber used also influences yield. Differences occur between species, grades, and board sequences (Fortney 1994, Buehlmann et al. 1998b, 1999). Another source of influence on yield is differing cutting bill requirements (Buehlmann et al. 1998a, 2003, 2008a, 2008b). In addition, the influences of these various factors on yield are interrelated, thus making it hard to develop a generally applicable model.

Existing yield estimation models do not provide the accuracy and consistency required by industry for planning material requirements or cost calculations (Manalan et al. 1980, Yaussy 1986, Thomas et al. 1996, Wiedenbeck and Scheerer 1996). Therefore, the present study examined a new approach to the problem of lumber yield estimation. Using data from Buehlmann et al. (2008a, 2008b, 2008c, 2008d), a least squares yield estimation model was developed and its performance reviewed.

Methods

The present study used empirical methods and statistical analysis for the development of a yield estimation model based on the method of least squares. Computer-based simulation techniques were used to derive the necessary data for least squares parameter estimation. The goal was to create a least squares model that would estimate yield dependent on cutting bill requirements, lumber grade, and rough-mill set-up.

Lumber cut-up simulation

To design the lumber cut-up simulation tests, the rip-first rough-mill yield simulator's settings as well as the lumber and the cutting bill used were defined to represent industrial rough-mill operations.

Rip-first rough-mill yield simulation.—This study used the USDA Forest Service's ROMI-RIP 1.0 (Thomas 1995a, 1995b) rip-first rough-mill simulator to represent the cut-up of lumber. ROMI-RIP has been shown to be a valid representation of lumber cut-up in industrial plants (Thomas and Buehlmann 2002). The settings used for this study were (1) all-blades movable arbor, (2) dynamic exponential cutting bill part prioritization (Thomas 1996b), (3) smart and unlimited salvage operation (Anderson et al. 1992, Thomas 1996a), (4) no random width and no random length

parts, (5) no finger-jointed or glued-up parts, (6) continuous updating of part counts, (7) end-and-side trim set at 6 mm on all sides, and (8) only clear-2-face (C2F) parts (Thomas 1995a, 1995b). Three replicates of each simulation were performed (Buehlmann 1998). Unless otherwise noted, yields are given in absolute terms and include primary and smart salvage yield (Thomas 1995a, 1995b). The simulation used the dynamic exponential cutting bill part prioritization method developed by Thomas (1996b). No part substitution was allowed; thus, the number of part sizes to be cut toward the end of a simulation run declined as part requirements were met.

Lumber.—No. 1 Common kiln-dried red oak lumber contained in the digital red oak data bank created by Gatchell et al. (1998) was used for the simulation tests. The board size and quality distribution published by Wiedenbeck et al. (2003) was used to create the lumber data files. To avoid biasing the resulting yields, each cutting bill's part quantity was set such that it required at least 150 boards (Buehlmann 1998, Buehlmann and Zuo 2008).

Cutting bills.—The ''Buehlmann'' cutting bill (Buehlmann et al. 2008c, 2008d), thought to represent cutting bills typically used in industrial operations of the secondary wood product industries, was used for this study. Table 1 (the first five columns) shows the details of this cutting bill. As described by Buehlmann et al. (2008c, 2008d), this cutting bill was designed to represent all cutting sizes between 127 and 2,159 mm in length and between 25 and 121 mm in width by clustering (e.g., grouping) individual part sizes to the 20 standard sizes shown in Table 1. For validation of the least squares yield estimation model, two additional sets of cutting bills were created: (1) five cutting bills using the Buehlmann cutting bill but uniform random part quantities (Table 1, columns 6 through 10 [labeled 1 through 5]) and (2) five cutting bills from the literature and industry that were clustered to fit the Buehlmann cutting bill (see Buehlmann 1998:213–215, appendices G and H).

Statistical analysis

Least squares regression modeling techniques were used to build a yield estimation model based on the standardized and simplified Buehlmann cutting bill (Buehlmann et al. 2008c, 2008d).

Least squares model.—Assumptions for ordinary least squares models include that factor levels are known constants, the observed responses are random variables, and the random error terms are independently, identically, and normally distributed with a mean of zero and a common variance (Ott 1993, Montgomery 2005). The data used for the creation and validation of the Buehlmann cutting bill (Buehlmann et al. 2008c, 2008d) and for the assessment of the influence of cutting bill requirements on lumber yield (Buehlmann et al. 2008a, 2008b) were used to estimate the parameters of the linear least squares yield estimation model.

The parameter estimates were attained using the General Linear Model procedure in the Statistical Analysis System (SAS) software package (SAS Institute 1996). The data used were screened to find outliers. Residual plots of the fitted model were used to reveal the presence of any lack of fit with the model data that had previously gone undetected. Because of the orthogonal experimental design, multi-collinearity was nonexistent. The data were fitted to the

Table 1.—*Part sizes and part quantity requirements of the original Buehlmann cutting bill and five randomly selected cutting bill quantities.*

Part no.	Part name	Length (mm)	Width (mm)	Cutting bill (no. of pieces)					
				Original[a]	1	2	3	4	5
1	L_1W_1	254	38	341	44	15	126	195	173
2	L_2W_1	445	38	742	600	633	347	536	649
3	L_3W_1	699	38	1,083	726	627	443	1045	835
4	L_4W_1	1,207	38	608	437	607	424	330	209
5	L_5W_1	1,842	38	258	92	215	132	36	211
6	L_1W_2	254	57	379	69	166	251	204	305
7	L_2W_2	445	57	746	395	173	297	687	502
8	L_3W_2	699	57	1,200	1,013	1,196	936	107	834
9	L_4W_2	1,207	57	654	88	499	484	552	355
10	L_5W_2	1,842	57	246	206	3	85	101	49
11	L_1W_3	254	89	114	18	75	36	93	75
12	L_2W_3	445	89	254	10	242	245	156	71
13	L_3W_3	699	89	365	76	349	16	139	309
14	L_4W_3	1,207	89	221	151	165	32	99	71
15	L_5W_3	1,842	89	142	81	25	78	73	126
16	L_1W_4	254	108	123	115	72	121	47	14
17	L_2W_4	445	108	248	110	113	244	50	107
18	L_3W_4	699	108	395	365	89	172	62	273
19	L_4W_4	1,207	108	213	43	106	15	46	7
20	L_5W_4	1,842	108	100	70	63	18	24	42

[a] Buehlmann et al. (2008d).

following model:

$$Y_m = \beta_0 + \sum_{i=1}^{4} \sum_{j=1}^{3} \beta_{ij}(L_iW_j)$$

$$+ \sum_{i=1}^{4} \sum_{j=1}^{3} \left\{ \sum_{k\neq i} \sum_{l\neq j} \beta_{ijkl}(L_iW_j)(L_kW_j) \right.$$

$$+ \sum_{k\neq i} \beta_{ijkj}(L_iW_j)(L_kW_j)$$

$$\left. + \sum_{l\neq j} \beta_{ijil}(L_iW_j)(L_iW_l) \right\} + \varepsilon_m \quad (1)$$

where Y_m is the yield for the *m*th observation, β_0 is the intercept, β_{ij} is the parameter estimate for part size *ij*, L_iW_j is the part quantity required by part size *ij* scaled from -1 (when quantity is zero) to $+1$ (when quantity is maximum), β_{ijkl} is the interaction between L_iW_j and L_kW_l, and ε_m is the random error.

Model validation.—The validity of the least square model was tested by comparing the estimated yield from the model with results obtained from simulation runs using the ROMI RIP software (Thomas 1995a, 1995b). In particular, three steps were used to validate the model. First, the model was compared with the data on which it was built (i.e., the results from the 512 cutting bills tested from the fractional factorial design). Second, the model was tested using the original cutting bill part sizes (Table 1), but the part quantities were determined as a uniform randomly distributed number. This step therefore tested the model's accuracy when only part quantities are changed. In Table 1, column 5 shows the original quantities for the Buehlmann cutting bill, and columns 6 through 10 display the quantities of the five cutting bills created using a random number generator to create required part quantities. Third, cutting bills used in actual rough mills were used to test the

practical usefulness of the least squares model. By analyzing these results, the possible estimation error could be quantified and assigned to the different sources of error for the model. The total error thus was partitioned into the following parts:

$$\text{Error}_{\text{Total}} = \text{Error}_{\text{Clustering}} + \text{Error}_{\text{Scaling}} + \text{Error}_{\text{Model}} \quad (2)$$

where $\text{Error}_{\text{Clustering}}$ is the error caused by the clustering of parts within part size ranges (part groups), $\text{Error}_{\text{Scaling}}$ is the error caused by the necessary scaling of the part quantities obtained from the cutting bill to fit the yield estimation model's framework (e.g., proportionally adjusting part quantities to levels consistent with those of the original model), and $\text{Error}_{\text{Model}}$ is the error caused by lack of fit of the model.

Results

Using results from Buehlmann et al. (2008a), a least squares model was built. Buehlmann et al. (2008a, 2008b) performed 512 tests with three replicates per test using the ROMI RIP simulation tool (Thomas 1995a, 1995b). The average yield found for these tests was 65.09 percent, with a standard deviation of 3.59 percent. The maximum yield obtained was 70.81 percent; the minimum was 48.63 percent. Thus, a 23.22 percent yield difference exists as a result of the different cutting bill requirements tested.

Least squares model

Based on the results from the 2^{20-11} fractional factorial design (Box et al. 1978) used by Buehlmann et al. (2008a, 2008b), the parameter estimates were attained using the method of least squares. All 20 main effects (i.e., the 20 parts) were significant at the 95 percent level of significance. Of the 190 secondary interactions, 113 were also significant at the 95 percent level. The secondary interactions help to explain the variability observed (Buehlmann et al. 2008a).
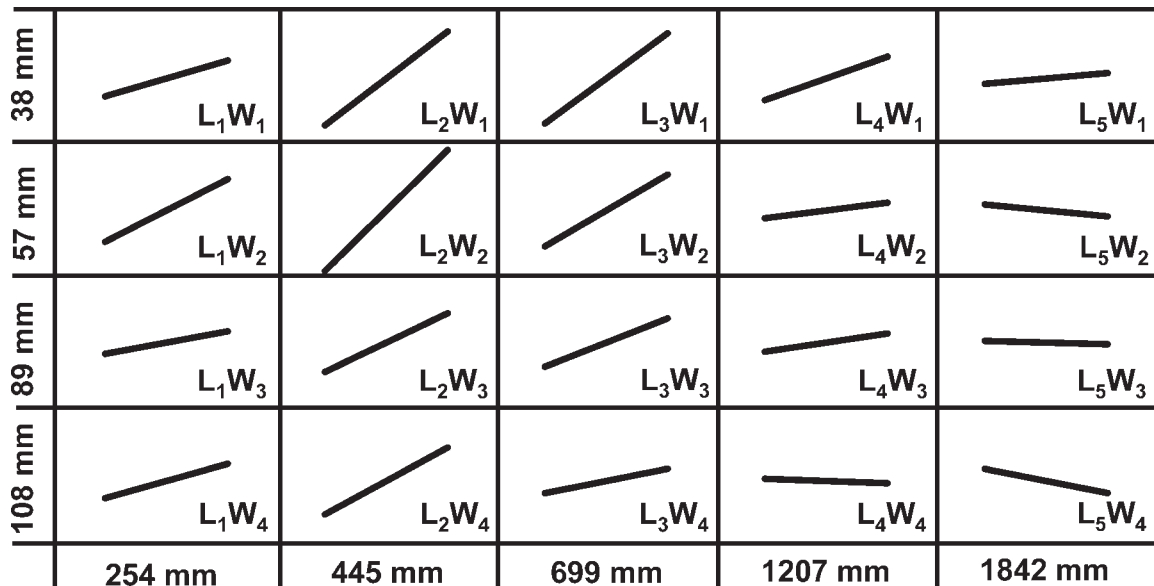
Figure 1.—Yield slopes for the 20 part sizes reflecting the average influence on yield of each individual part size.

The 20 main effects were able to explain 78 percent of the variability observed (i.e., $R^2 = 0.78$). The R^2 value for the model containing all main effects and all interactions was 0.95. However, because the coefficient of determination (R^2) can always be increased by adding additional regressor variables (Mays 1995), adjusted R^2 is considered to better represent the ability of the model to account for the variability observed. The adjusted R^2 for the model containing all 210 terms (20 main effects and 190 secondary interactions) was 0.94. Forward selection at the 95 percent level of significance created a model containing all 20 main effects and 113 secondary interactions. The adjusted R^2, however, was approximately the same as that for the full model, with a value of 0.95. Similar R^2 values were obtained for the backward, the forward, and the stepwise procedure (Ott 1993). Given that the additional terms in the full model, with all main effects and all secondary interactions included, did not decrease the adjusted R^2 markedly, the full model was chosen as the appropriate one. Thus, the model used was the one presented in Equation 1.

Values for the main effects ranged from +1.60 for part size L_2W_2 to −0.32 for part size L_5W_4, whereas values for the secondary interactions varied from +0.27 for the interaction between L_2W_4 and L_4W_4 to −0.39 for the interactions between L_3W_1 and L_4W_1 (all values highly significant at the 99% level). The values for the 20 main effects and the 190 secondary interactions can be found in Buehlmann (1998, tables 5.4 and A-6). Figure 1 displays the yield slopes of the 20 main effects. In this figure, the right end of the slope within each part size represents the yield contribution of a particular part size when the maximum part quantity for this part is called for. The left end of the slope represents the yield contribution of a part size when no parts of this size are required (quantity of zero). Total yield is then calculated as the sum of the individual part size's yield contributions. The intercept (i.e., overall average yield) was 65.09 percent and intersects all the slopes in the middle between zero and maximum quantity. In other words, if all part quantities are set at 50 percent quantity,

then the model estimates the yield to be 65.09 percent (Buehlmann 2008a).

Validation of the least squares model

The validation of the least squares model was accomplished through (1) using cutting bills from the original fractional factorial design, (2) using cutting bill part quantities set randomly, and (3) using cutting bills from industry.

Validation based on data from the original fractional factorial design.—Replicating the 512 test settings used in the fractional factorial design employed to create the least squares model, 303 of the yield results estimated by the model had an estimation error of smaller than 0.50 percent absolute yield, eight observations had an error of larger than 2 percent, and the maximum estimation error observed was 4.27 percent. Overall, 468 of the 512 estimated yields were within the 95 percent confidence interval, and 44 were outside. The average of all errors cancelled out as expected (i.e., the average estimation error was 0.00 percent). From a practical standpoint, the model performs very well when the input data adhere to the framework set forth for the model. This is supported by the fact that the model estimated the yield of a given cutting bill within an accuracy 1 percent absolute yield in 88 percent of the cases.

Validation using randomly determined part quantities.—Five tests were performed to determine the accuracy of the yield estimation model when the part quantities were no longer either zero or maximum quantity for a given part size but, instead, were random quantities between zero and maximum quantity (Table 1, columns 6 through 10). The average estimation error for these five cutting bills was 2.19 percent, with the maximum error being 2.89 percent and the minimum error 1.62 percent (Table 2). All the estimated yields were significantly different from the simulated yields ($\alpha = 0.05$). However, the model ranked the cutting bills successfully within an accuracy of less than half a percent yield. In other words, even though the absolute yield result had an average error of 2.19 percent, the model still

Table 2.—Individual yield (%) results and summary statistics for the five tests with random part quantities.[a]

	Test 1	Test 2	Test 3	Test 4	Test 5	Average
Simulation	68.73	68.45	69.48	70.11	70.08	69.37
LS model[b]	65.95	66.56	66.59	68.35	68.46	67.18
Difference	2.78*	1.89*	2.89**	1.76*	1.62*	2.19***

[a] Asterisks indicate statistical significance (* = significant at 95% level, ** = significant at 99% level, *** = not tested).
[b] LS model = least squares model.

distinguished the relative yield of different cutting bills within an accuracy of approximately half a percent yield. The ranking, from highest to lowest, of yield results for the five tests using simulation is test 4, test 5, test 3, test 1, and test 2. The ranking using the yield estimation model is test 5, test 4, test 3, test 2, and test 1. However, yields from tests 4 and 5 were only 0.03 percent different and those from tests 1 and 2 only 0.28 percent apart, while the difference to test 3 is 0.72 percent between tests 1 and 3 and 0.63 percent between tests 3 and 5. Thus, the cutting bills were ranked correctly into fairly narrow pockets of yield levels.

Validation using industrial cutting bills.—Finally, five tests were performed using actual cutting bills from industry. Table 3 shows the average results obtained from these tests. Also shown are the standard deviations and the maximum and minimum values observed. In Table 3, lines 1 through 4 show the results for the original simulation (line 1), the results for the simulation with parts clustered to the 20 standard sizes of the Buehlmann cutting bill (line 2), the results for the simulation with parts clustered and scaled (line 3), and the estimated average yield for the five cutting bills from the least squares yield estimation model (line 4). Lines 5 through 10 show the differences between the four yield results presented on lines 1 through 4. Lines 11 and 12 show the number of different part sizes in the original and clustered cutting bills.

The average yield obtained from the original cutting bills tested, as shown in Table 3, was 67.61 percent. When the original cutting bill parts were clustered to the standard sizes of the Buehlmann cutting bill, the average yield decreased by 1.82 percent, on average, for the five cutting bills tested (Table 3, line 5). This was expected, because the average number of parts in the cutting bill decreased from 20 (line 11) for the full cutting bill to nine (line 12) when the parts were clustered. Yield from a cutting bill generally increases when the number of parts to be cut increases (Buehlmann et

al. 1998a, Thomas and Brown 2003, Buehlmann et al. 2008b).

The scaling of the part quantities influenced yield by, on average, 0.85 percent (Table 3, line 7). This result is obtained by subtracting the average yield of the clustered and scaled cutting bills (line 3) from the average yield of the clustered cutting bills (line 2). This result was thought to have a minor influence on yield, because the proportions of the part quantity remain the same in the full cutting bill compared to the scaled one. However, as this observation reveals, yield does not change in a completely proportional fashion when the part quantities for a cutting bill are changed proportionally (i.e., scaled).

The error term caused by the least squares estimation model, however, was the largest of all the error terms observed. On average, for the five cutting bills used, the model's estimation error was 7.62 percent (Table 3, line 10). This result is obtained when the average yield from the least squares estimation model (line 4) is subtracted from the average yield of the clustered and scaled simulation tests (line 3). Equation 3 displays the breakdown of the average total error of the five cutting bills tested:

$$10.30\% \ (\mathrm{Error}_{\mathrm{Total}}) = 1.82\% \ (\mathrm{Error}_{\mathrm{Clustering}}) \\ + 0.85\% \ (\mathrm{Error}_{\mathrm{Scaling}}) \\ + 7.62\% \ (\mathrm{Error}_{\mathrm{Model}}) \quad (3)$$

The average total estimation error for this case was 10.30 percent (Table 3, line 8). The maximum error observed was 12.19 percent, and the minimum error was 8.01 percent.

Of the total average estimation error, 1.82 percent was caused by the clustering of parts (18% of the total average estimation error), another 0.85 percent can be attributed to scaling of the part quantities (8% of the total average estimation error), and the remaining 7.62 percent to the least squares yield estimation model (74% of the total average

Table 3.—Average yield (%) results and maximum and minimum values obtained when testing the five "real" cutting bills.

Test	Observation	Average ± SD	Max.	Min.
1	Simulation	67.61 ± 3.11	72.39	64.40
2	Clustered simulation	65.79 ± 3.14	70.16	62.64
3	Clustered and scaled simulation	64.94 ± 2.69	68.13	62.09
4	LS model[a]	57.31 ± 3.17	60.20	52.71
5	Simulation − clustered simulation [test 1 − test 2]	1.82 ± 3.12	4.48	2.28
6	Simulation − clustered and scaled simulation [test 1 − test 3]	2.68 ± 3.31	5.58	2.99
7	Clustered simulation − clustered and scaled simulation [test 2 − test 3]	0.85 ± 0.77	2.04	0.00
8	Simulation − LS model [test 1 − test 4]	10.30 ± 1.69	12.19	8.01
9	Clustered simulation − LS model [test 2 − test 4]	8.48 ± 4.32	15.26	4.85
10	Clustered and scaled simulation − LS model [test 3 − test 4]	7.62 ± 4.24	14.68	3.92
11	No. of parts in cutting bill	20 ± 15	36	7
12	No. of part sizes used	9 ± 4	16	6

[a] LS model = least squares model.

Table 4.—*Results of the lack-of-fit test based on the 512 tests (three replicates).*[a]

Source	SS	df	MS	F	P
Lack of fit	841.32	301	2.80	23.09	0.0001
Pure error	123.99	1,024	0.12		
Total error	965.31	1,325	0.73		

[a] SS = sum of squares, MS = mean square.

estimation error). Only a quarter of the total error thus can be attributed to the clustering and scaling of the original cutting bills to the Buehlmann cutting bill, which supports the concept that was used to derive the Buehlmann cutting bill (Buehlmann et al. 2008c). Three-quarters of the error results from the yield estimation model. However, despite the rather large estimation error, the model classified four out of five cutting bills correctly as to their ranking of the yield level achieved.

Discussion

The yield estimation model works well for cutting bills whose requirements conform to the framework used in the present study. Especially when testing cutting bills with part quantities that are either maximum or zero quantity according to the part quantity distribution created in Buehlmann et al. (2008d), the estimation error was rather small. Even when using uniform random numbers between zero and maximum quantity to establish part quantities for individual part sizes, the error size was still acceptable. Moreover, even for actual cutting bills, the ranking of cutting bills in terms of the relative yield achieved compared to others was quite successful. The model ranked four out of five cutting bills correctly. The model thus can be used for testing the influence of different cutting bill requirements on yield as compared to others. By so doing, cutting orders can be broken down into several individual cutting bills such that overall yield is maximized, as suggested by Buehlmann et al. (1998a).

Attempts were undertaken to improve the accuracy of the yield estimation model for cutting bills from industrial operations. To establish the source of the error when estimating the actual cutting bills from industry, a lack-of-fit test was performed (Neter et al. 1996). Lack-of-fit tests break the error from the model down into two components, one being the noise from the differences in yield obtained between the replicates of each test and the other being the error caused by the least squares model not being perfect. Using the data from the original 2^{20-11} fractional factorial design (Box et al. 1978), the squared error for each of the 1,536 data points obtained (512 tests with three replicates each) was calculated using Equation 4:

$$\sum_{k=1}^{k=3} \sum_{j=1}^{j=4} \sum_{i=1}^{i=5} (Y_{ijk} - \hat{Y}_{ij})^2$$
$$= \sum_{k=1}^{k=3} \sum_{j=1}^{j=4} \sum_{i=1}^{i=5} (Y_{ijk} - \overline{Y}_{ij})^2 + \sum_{i=1}^{i=4} \sum_{j-1}^{j=5} (\overline{Y}_{ij} - \hat{Y}_{ij})^2 \quad (4)$$

where Y_{ijk} is the yield result of replicate ijk, $\hat{Y}_{ij}$ is the estimated value ij, and $\overline{Y}_{ij}$ is the average yield obtained of the three replicates.

Table 4 shows the results of the lack-of-fit test based on the 512 tests (three replicates each). The squared error caused by lack of fit is almost seven times larger than the squared error caused by the noise in the system. Lack of fit thus is a highly significant ($\alpha < 0.01$) contributor to the total error observed, as indicated by the P value of 0.0001. This suggests that the variability of the average yield from the three replicates is not captured perfectly by the model.

Residual analyses did not reveal a pattern that would allow us to conclusively point to a source (or sources) of the errors observed (Fig. 2). However, Figure 2 shows that the larger yield estimation errors occur when either few or many part sizes are required to be cut by the cutting bill. When 9 to 16 part sizes are required, the residual is smaller. Relatively few residuals exceed 2 percent error. Also, one has to be aware that the residuals are calculated from individual observations (i.e., the replicates) and not on the average yield of the three replicates, to which the least squares model was fitted. When analyzing the errors of the average yield from the three replicates versus the estimated value of the model, only 8 of the 512 observations were larger than 2.00 percent. Thus, despite the lack of fit of the model, the least squares estimation model was able to estimate the expected yield quite accurately.

Attempts to improve the accuracy of the least squares model, as suggested by Draper and Smith (1981) and Neter et al. (1996), by using either a higher-order polynomial model or transformation failed. Polynomial models up to the third order for the main effects were tested with no significant improvement of the lack-of-fit term. Neither was transformation of the data successful. The following transformations were tested: natural log, exponential, square root, logit, reciprocal, power *k,* and combinations of these. However, none of these transformations reduced the lack of fit significantly.

Future research should explore other techniques for building an estimation model to create a simple yet reliable tool for the industry. For example, neural networks offer a way to deal with nonlinear data. Also, neural networks are nonparametric models, i.e., they do not require prior knowledge of the function being estimated. These two features of neural networks make them a superior candidate for solving the problem at hand. For the yield estimation problem, nonlinearity is most likely present, and the function that relates cutting bill characteristics to yield is unknown. Another interesting aspect of neural networks is that they can be trained to solve problems (Burke 1991). Fuzzy systems are also a technique with potential. As was observed when the errors caused by the least squares estimation model were analyzed, in some instances a small quantity of parts of a particular size seems to have a disproportionally large influence on yield. Fuzzy systems offer a way to address nonlinearity and interactions implicit in a set of data. They do not, however, require the specification of a nonlinear dynamic system, the acquisition of a representative set of training samples, and the encoding of the training samples by repeated learning cycles, as is the case for neural networks. Fuzzy systems require only that a ''rule matrix'' be partially filled by an expert (Kosko 1992). Future research could show if fuzzy systems would be able to capture the nonlinear response of the dependent variable (part yield) from different levels of the independent variables (required part quantities). The wood industry would benefit greatly from having a simple yet accurate
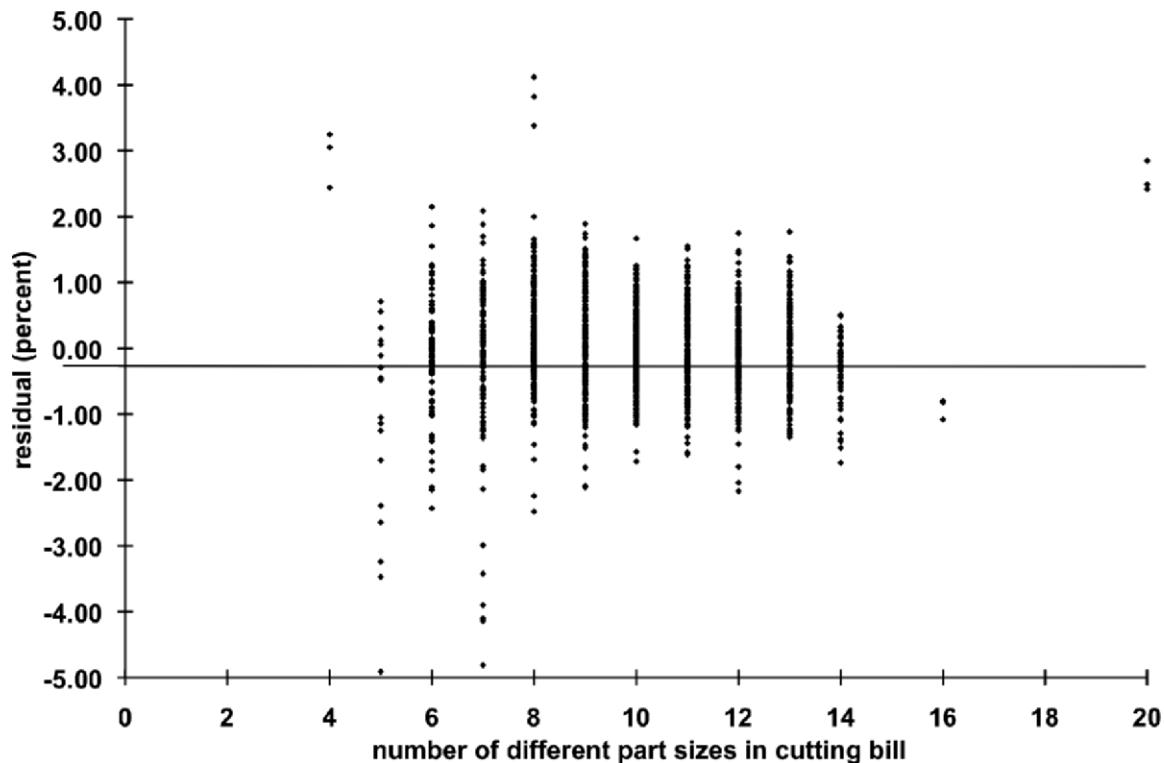
Figure 2.—Residual plot for the 512 tests (1,536 data points given three replicates per test).

yield estimation tool that would allow the industry to better estimate raw material costs and requirements before execution. Also, such an estimation tool could be used to compose cutting bills for maximum overall yield, thus lowering costs and raw material requirements.

Summary and Conclusions

Linear least squares estimation of yield based on cutting bill requirements is a viable concept provided the cutting bills used adhere to the framework (i.e., part sizes and part quantities) established for the model. Within this framework, cutting bill yield could be estimated within 1 percent accuracy for 88 percent of the cutting bills tested. Furthermore, the model was able to rank actual cutting bills in order of expected level of yield correctly in 80 percent of the cases. This allows use of the yield estimation model for assessing cutting bills in terms of their potential for higher yield as compared with other cutting bills. Moreover, large cutting orders can be broken down into several individual cutting bills such that maximum overall yield is achieved. Also, for cutting bills with requirements similar to the cutting bills used to create the model, more accurate job costing can be achieved compared with other models that do not use simulation to predict part yields resulting from lumber cut-up.

More work needs to be done to fit the model to cutting bills from industry, where the parts have to be clustered and scaled to the standardized and simplified cutting bill. The estimation error for such cutting bills was 10.30 percent, on average, for the five cutting bills tested. The error of the model was 7.62 percent, while clustering and scaling only contributed 1.82 and 0.85 percent, respectively, to the total estimation error. More work is needed to improve the existing least squares model for more general yield estimation purposes. Neural networks or fuzzy systems may provide appropriate models for the general yield estimation problem, because they offer advantages over linear or nonlinear least squares estimation.

Acknowledgments

Literature Cited

Anderson, J. D., C. D. Brunner, and A. G. Maristany. 1992. Effect of sawing stages on fixed-width, fixed-length dimension yield. *Forest Prod. J.* 42(11/12):74–78.

Box, G. E. P., W. G. Hunter, and J. S. Hunter. 1978. Statistics for Experimenters. John Wiley & Sons, New York. 653 pp.

Brunner, C. C., D. A. Butler, A. G. Maristany, and D. VanLeeuwen. 1990. Optimal clear-area sawing patterns for furniture and millwork blanks. *Forest Prod. J.* 40(3):51–56.

Buehlmann, U. 1998. Understanding the relationship of lumber yield and cutting bill requirements: A statistical approach. Doctoral dissertation. Virginia Polytechnic Institute and State University, Blacksburg.

Buehlmann, U., D. E. Kline, and J. K. Wiedenbeck. 1998a. Understanding the relationship of lumber yield and cutting bill requirements: A statistical approach to the rough mill yield estimation problem. Presented at the Annual Meeting of the Forest Products Society, June 23, 1998, Merida, Mexico.

Buehlmann, U., D. E. Kline, J. K. Wiedenbeck, and R. Noble. 2008a. The influence of cutting bill requirements on lumber yield using a fractional factorial design—Part I: Linearity and least squares. *Wood Fiber Sci.* 40(4):599–609.

Buehlmann, U., D. E. Kline, J. K. Wiedenbeck, and R. Noble. 2008b. The influence of cutting bill requirements on lumber yield using a fractional factorial design—Part II: Correlation and number of part sizes. *Wood Fiber Sci*. 40(4):610–619.

Buehlmann, U., D. E. Kline, J. K. Wiedenbeck, and R. Noble. 2008c. Validation of the standardized and simplified cutting bill. *Wood Fiber Sci*. 40(2):202–213.

Buehlmann, U., J. K. Wiedenbeck, and D. E. Kline. 1998b. Character-marked furniture: Potential for lumber yield increase in rip-first rough mills. *Forest Prod. J*. 48(4):43–50.

Buehlmann, U., J. K. Wiedenbeck, and D. E. Kline. 1999. Character-marked furniture: Potential for lumber yield increase in crosscut-first rough mills. *Forest Prod. J*. 49(2):65–72.

Buehlmann, U., J. K. Wiedenbeck, and D. E. Kline. 2003. Influence of cutting bill requirements on lumber yield in a rip-first rough mill. *Wood Fiber Sci*. 35(2):187–200.

Buehlmann, U., J. K. Wiedenbeck, R. Noble, and D. E. Kline. 2008d. Creating a standardized and simplified cutting bill using group technology. *Wood Fiber Sci*. 40(1):29–41.

Buehlmann, U. and X. Zuo. 2008. Investigating the influence of lumber sample subsets on simulated rough mill part yields. *Forest Prod. J*. 40(3):84–90.

Burke, L. I. 1991. Introduction to artificial neural systems for pattern recognition. *Comput. Oper. Res*. 18(2):211–220.

Draper, N. R. and H. Smith. 1981. Applied Regression Analysis. 2nd ed. John Wiley & Sons, New York. 709 pp.

Dunmire, D. E. 1971. Predicting yields from Appalachian red oak logs and lumber. *In:* Oak Symposium Proceedings, USDA Forest Service, Northeastern Forest Experiment Station, Upper Darby, Pennsylvania.

Englerth, G. H. and D. E. Dunmire. 1966. Programming for lumber yield. *Forest Prod. J*. 16(9):67–69.

Englerth, G. H. and D. R. Schumann. 1969. Charts for calculating dimension yields from hard maple lumber. Research Paper FPL 118. USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin.

Gatchell, C. J., R. E. Thomas, and E. S. Walker. 1998. 1998 Data bank for kiln-dried red oak lumber. Research Paper NE-245. USDA Forest Service, Northeastern Forest Experiment Station, Radnor, Pennsylvania.

Hall, S. P. 1978. The effects of one cross-cut-first line and one rip first-line on rough mill costs and yields. Master's thesis. Virginia Polytechnic Institute and State University, Blacksburg.

Harding, O. V. 1991. Development of a decision software system to compare rip-first and crosscut-first yields. Doctoral dissertation. Mississippi State University, Mississippi State.

Hoff, K. 2000. Limitations of lumber-yield nomograms for predicting lumber requirements. General Technical Report NE-270. USDA Forest Service, Northeastern Forest Experiment Station, Newton Square, Pennsylvania.

Kosko, B. 1992. Neural Networks and Fuzzy Systems. A Dynamical System Approach to Machine Intelligence. Prentice Hall, Englewood Cliffs, New Jersey.

Manalan, B. A., G. R. Wells, and H. A. Core. 1980. Yield deviations in hardwood dimension stock cutting bills. *Forest Prod. J*. 40(1):40–42.

Mays, D. P. 1995. Class notes: Statistics in research II. Virginia Polytechnic Institute and State University, Blacksburg.

Montgomery, D. C. 2005. Design and Analysis of Experiments. 6th ed. John Wiley & Sons, Hoboken, New Jersey. 660 pp.

National Hardwood Lumber Association (NHLA). 2007. Rules for the measurement and inspection of hardwood and cypress. NHLA, Memphis, Tennessee.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wassermann. 1996. Applied Linear Statistical Models. Times Mirror Higher Education Group and Richard D. Irwin, Chicago. 1,399 pp.

Ott, R. L. 1993. An Introduction to Statistical Methods and Data Analysis. 4th ed. Duxbury Press, Belmont, California. 1,173 pp.

SAS Institute. 1996. SAS User's Guide. SAS Institute, Cary, North Carolina.

Schumann, D. R. 1971. Dimension yields from Black Walnut lumber. Research Paper FPL-162. USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin.

Schumann, D. R. 1972. Dimension yields from alder lumber. Research Paper FPL-170. USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin.

Schumann, D. R. 1973. Dimension stock yields from lumber of three hardwood species. *Forest Prod. J*. 23(3):17–21.

Schumann, D. R. and G. H. Englerth. 1967a. Yields of random width dimension from 4/4 hard maple lumber. Research Paper 81. USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin.

Schumann, D. R. and G. H. Englerth. 1967b. Dimension stock yields of specific width cuttings from 4/4 hard maple lumber. Research Paper FPL-85. USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin.

Thomas, R. E. 1995a. ROMI RIP: ROugh MIll RIP-first simulator user's guide. General Technical Report NE-202. USDA Forest Service, Northeastern Forest Experiment Station, Radnor, Pennsylvania.

Thomas, R. E. 1995b. ROMI RIP: ROugh MIll RIP-first simulator. General Technical Report NE-206. USDA Forest Service, Northeastern Forest Experiment Station, Radnor, Pennsylvania.

Thomas, R. E. 1996a. ROMI RIP: An analysis tool for rip-first rough-mill operations. *Forest Prod. J*. 46(2):57–60.

Thomas, R. E. 1996b. Prioritizing parts from cutting bills when gang-ripping first. *Forest Prod. J*. 46(10):61–66.

Thomas, R. E. and J. Brown. 2003. Determining the impact of sorting capacity on rip-first rough mill yield. *Forest Prod. J*. 53(7):54–60.

Thomas, R. E. and U. Buehlmann. 2002. Validation of the ROMI-RIP Rough Mill Simulator. *Forest Prod. J*. 52(2):23–29.

Thomas, R. E., J. K. Wiedenbeck, and K. Hoff. 1996. Make informed decisions in the rough mill using new and improved computer simulation programs. *In:* Proceedings of CIFAC '96 International Symposium; Wood Machining Institute, Berkeley, California, pp. 47–56.

Thomas, R. J. 1965a. The yield of dimension stock: Volume 1. Technical Report 24a. North Carolina State University, Raleigh.

Thomas, R. J. 1965b. The yield of dimension stock: Volume 2. Technical Report 24a. North Carolina State University, Raleigh.

Thomas, R. J. 1965c. Analysis of yield of dimension stock from standard lumber grades. *Forest Prod. J*. 15(7):285–288.

Wiedenbeck, J. K., J. Brown, N. Bennett, and E. Rast. 2003. Hardwood lumber widths and grades used by the furniture and cabinet industries: Results of a 14-mill survey. *Forest Prod. J*. 53(4):72–80.

Wiedenbeck, J. K. and C. Scheerer. 1996. A report on rough mill yield practices and performance—How well are we doing? *In:* Twenty-Fourth Annual Hardwood Symposium of the Hardwood Research Council. pp. 121–128.

Wiedenbeck, J. K. and R. E. Thomas. 1995a. Don't gamble your fortunes—Focus on rough mill yield. *Wood Wood Prod*. 100(7): 148–149.

Wiedenbeck, J. K. and R. E. Thomas. 1995b. Rough mill study identifies yield improvement opportunities. *Millwork Manuf*. 12(3):22–23.

Yaussy, D. A. 1986. Green lumber grade yields from factory grade logs of three oak species. *Forest Prod. J*. 36(5):53–56.