Growing Poplar for Growing Markets: An Econometric Study of Lumber Production at the Upper Columbia Mill, Eastern Oregon, USA

Michael Rinaldi
Duke University | Nicholas School of the Environment
Master of Environmental Management 2015
Michael.Rinaldi@duke.edu

Advisors:
Gwen Busby, Economist at GreenWood Resources
Jeffery Vincent, Professor at Duke University

Please Note: Confidential information has been redacted.
1. Executive Summary

GreenWood Resources (GWR) is a global timberland investment and asset management company specializing in the acquisition, development, and management of forestry assets.\(^1\) Assets under GWR management and committed capital total approximately $800 million with the company specializing in fast-growing, short-rotation hardwood tree farms.\(^2\) As of June 2014, GWR manages approximately 57,700 acres in five countries, with an active acquisition program in North America, Latin America, Europe and Asia. The largest North American asset is the Boardman Tree Farm (BTF), a 26,000-acre FSC-certified hybrid poplar plantation in eastern Oregon. The BTF was acquired and built through a GWR investment entity, the Greewood Tree Farm Fund (GTFF). In addition to the tree farm, GTFF built and operates one of the largest hardwood sawmills in North America, the on-site Upper Columbia Mill (UCM), with mill management and lumber sales provided by the Collins Companies.

BTF’s fast-growing trees are a hybrid of cottonwood and poplar developed in the 1980’s for the pulp market and, as a result, have historically had a poor reputation in the lumber industry. Under GWR management, the hybrid poplar was rebranded Pacific Albus\(^*\) and hybrid poplar lumber is currently being sold and marketed as an alternative to alder and aspen. As Pacific Albus becomes increasingly competitive across a range of applications, the market price of Pacific Albus is increasing and converging toward traditionally higher-valued competing species.

Building on past work by GWR, in the fall of 2014 a Duke University master’s student Michael Rinaldi initiated a two-part Masters Project investigating: (1) the statistical relationship between log inputs and lumber outputs at the Upper Columbia Mill; and (2) supply, demand, and price dynamics of Pacific Albus. The following report represents the econometrics portion of this project.

The focus of the present study is an analysis of the data generated by the Comact scanner. We begin by using these data to estimate the lumber recovery factor (LRF), the average volume of lumber output per unit of log volume input. Then, building on this platform, we estimate lumber production by product using log measurement data. Products from the UCM are grouped and sold in three distinct categories or product dimensions: 4/4 x RW (random width), 3 x 7, and 5/4 x RW. Understanding the relationship between mill inputs and outputs is key to understanding the efficiency of mill operations and profitability. Further, the ability to predict lumber production by product makes it possible to generate reliable forecasts of production. Key findings of the econometric analysis include the following:

- An LRF model was created that explained 62% of the variation in the lumber recovery factor
- A Product Recovery Model was developed that explains 73% of the variation in board feet recovered by product dimension and can estimate the board feet recovered from a log by dimension, from just a few key log measurements
- Created an interactive tool for estimating the volume recovery and value of logs with a few measurements

# Table of Contents

1. **EXECUTIVE SUMMARY** ................................................................. 2

2. **INTRODUCTION** ............................................................................ 4

3. **DATA** .......................................................................................... 6
   3.1 Lumber Recovery Factor (LRF) Model ................................. 6
   3.2 Product Recovery (PR) Model ............................................... 6
   3.3 Independent Variables .......................................................... 6

4. **METHODS AND MODELS** ....................................................... 8
   4.1 Lumber Recovery Factor Model ........................................... 9
   4.2 Product Recovery Model ....................................................... 9

5. **RESULTS** .................................................................................. 12
   5.1 Lumber Recovery Factor Model Results ............................. 12
   5.2 Product Recovery Model Results ....................................... 14

6. **DISCUSSION AND CONCLUDING REMARKS** ..................... 16

7. **WORKS CITED** ......................................................................... 18
2. Introduction

Pacific Albus is a hybrid poplar grown in the Pacific Northwest, near Boardman, Oregon. Owned by the GreenWood Tree Farm Fund (GTFF), the Boardman Tree Farm (BTF) is managed by GreenWood Resources (GWR). The FSC-certified plantation is about 26,000 acres and has an estimated capacity to produce up to 80 million board feet per year. The on-site, state-of-the-art, Upper Columbia Mill (UCM) was built in 2008 and is one of the largest hardwood sawmills in North America. Mill management and lumber sales are provided by Collins, the first privately owned forest products company in the United States to have all of its hardwood and softwood forests certified by the Forest Stewardship Council (FSC).

Virtually all sawlogs processed at the UCM are sourced from the BTF. This production process is illustrated in Figure 1. Following harvest and delivery to the mill, stems are measured and bucked into logs by the merchandizer. Logs are then transported through a Comact scanner that measures and records each log’s small-end diameter (SED), large end diameter (LED), volume, taper, sweep, and length. Using these measurements, the Comact scanner identifies the optimal mix of boards to be sawn from the log. These products are then cut to their specified dimensions, sorted and visually graded, and dried and packaged for sale. The Comact scanner and operating system are a unique, proprietary system running on a QNX platform. The scanner is programmed to optimize lumber volume recovery at the mill, but its decision making process is unknown to UCM managers and held as a Comact secret.

Figure 1: Production Process

---

4 Collins, 2014
The focus of the present study is an analysis of the data generated by the Comact scanner. We begin by using these data to estimate the lumber recovery factor (LRF), the average volume of lumber output per unit of log volume input. Then, building on this platform, we estimate lumber production by product using log measurement data. Products from the UCM are grouped and sold in three distinct categories or product dimensions: 4/4 x RW (random width), 3 x 7, and 5/4 x RW. Understanding the relationship between mill inputs and outputs is key to understanding the efficiency of mill operations and profitability. Further, the ability to predict lumber production by product makes it possible to generate reliable forecasts of production.

Many studies in the forest economics literature successfully incorporate individual log characteristics such as SED, LED, sweep, taper, and length into statistical models for predicting and simulating product recovery. For example, Barbour et al. (2003) compare actual recovery from Douglas fir and Ponderosa pine to the results predicted by the AUTOSAW sawing simulator and find that the simulator consistently underestimated the volume recovered from the logs. In a separate study, Liu et al. (2007) model lumber value recovery in relation to selected tree characteristics in black spruce using the Optitek sawing simulator. The black spruce models were able to account for more than 92% of the total variation in lumber value recovery. Further Liu et al. (2007) find that DBH has the greatest positive influence on the tree-level product value, followed by tree height.

Similar to the present study, Auty et al. (2014) develop a model to predict the lumber product assortment as a function of tree size variables. The authors model lumber product assortment in black spruce and balsam fir using zero inflated Poisson regression to account for the frequent occurrence of zeros in the count of products by log. They find that models fitted to simulated lumber recovery data tend to over-predict lumber volume and that tree-level models of lumber recovery provide no information about the composition of the assortment of sawn products.

The objective of the present study is to develop an econometric model to estimate Pacific Albus lumber recovery as a function of log measurements. Using a dataset from the UCM Comact scanner, which includes observations from over 240,000 individual logs, models describing lumber recovery across products and recovery by product and grade. These models will have greater explanatory power than previous, data-constrained modeling efforts. Results from the analysis will improve the ability to forecast lumber production and inventory available for sales. Further, the model may eventually be incorporated into a larger framework that incorporates optimal bucking and silvicultural management decisions.

The remainder of the paper is organized as follows. In Section 3, we describe the data. Methods and the econometric models are described in Section 4. Next we present results from model estimation in Section 5. Finally, in Section 6, we provide a discussion of the results and offer concluding remarks.

---

3. Data

Data for the study were collected from the UCM’s Comact scanner between October 16th and December 9th, 2014 and includes observations from 240,539 individual logs. The scanner identifies each log according to the time that it was scanned and records the log’s physical measurements along with a list of products to be sawn from the log, with the depth, width, and length of each product specified. The scanner determines that most logs yield no more than 3–5 pieces of lumber.

3.1 Lumber Recovery Factor Model

The dependent variable in the LRF model is the board feet of lumber produced per cubic meter of log volume. We index logs by \( i \) and specify the LRF as:

\[
LRF_i = \frac{BF_i}{vol_i} \tag{[1]}
\]

Where:

\( BF_i \) = board feet recovered from log \( i \)

\( vol_i \) = volume of log \( i \), measured in \( m^3 \)

LRF is the most fundamental measure of mill productivity and is a widely used metric within the lumber industry. In the context of the UCM, the LRF is critical to achieving profitable production levels from sustainable harvest volumes from the BTF.

3.2 Product Recovery Model

The dependent variable in the product recovery (PR) model is board feet of lumber recovered by product. Unlike the LRF model where the aim is to explain the efficiency of recovery, the product recovery model seeks to explain the volume of lumber recovered from log \( i=1,2,\ldots,n \) by product dimension \( t=1,2,\ldots,m \), \( BF_{it} \), and the dependent variable is specified as:

\[
\ln(1 + BF_{it}) \tag{[2]}
\]

In this product recovery model, there 718,377 different logs \( (i) \) and three different product dimensions \( (t) \): 4/4 x RW, 3 x 7, and 5/4 x RW. The three product dimensions represent the categories by which products are grouped and sold, and they will be further discussed in Section 3.

Note that to prevent the prediction of negative board feet, we take the natural log of \( BF_{it} \). Additionally, to prevent observations from dropping out when product board feet is equal to 0 we add 1 to \( BF_{it} \) prior to the natural log transformation.

3.3 Independent Variables

For both models, the set of independent variables is similar and are based on log measurements. The four independent variables of interest are all measured in inches and are defined and described in Table 1 and Figure 2.
### Table 1: Independent Variables (values in inches)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Description</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>SED</td>
<td>Small end diameter of log</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LED</td>
<td>Large end diameter of log</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweep</td>
<td>Gradual bend in log</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>Log length</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2:** Diagram of log characteristics
4. Methods and Models

As just explained, the analysis covers two different models answering two different questions. The lumber recovery factor model seeks to answer how different log measurements affect the efficiency of lumber recovery. The product recovery model seeks insight into the question of how log size and form determine how many board feet of each product dimension can be produced from a single log.

In addition to the independent variables presented in Table 1, GWR presents a set of transformed variables for analyses in Table 2. They chose the natural log of SED squared as it was determined to be greatest explanatory power. Instead of running sweep as an independent variable, a sweep factor was created as sweep/SED to provide more context than sweep by itself.

Table 2: Transformed Independent Variables

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnSED^2</td>
<td>ln(SED^2)</td>
</tr>
<tr>
<td>Sweep Factor</td>
<td>(Sweep) / SED</td>
</tr>
<tr>
<td>Log Length</td>
<td>Log length</td>
</tr>
</tbody>
</table>

To create a model that does an even better job of explaining recovery, volumetric independent variables were also created and used in subsequent models. The volume of a log without any sweep is given by the volume of a circular truncated cone, or conical frustrum:

\[ \text{Volume} = \frac{1}{3} \pi \cdot L \cdot (R_1^2 + R_1 \cdot R_2 + R_2^2) \]  \[3\]

Where:

- \( L = \) log length
- \( R_1 = \) short end radius or (SED/2)
- \( R_2 = \) long end radius or (LED/2)

Using Equation 3, two sets of volumetric independent variables were created and are found in Table 3. The first 3 variables interact length, SED, and LED. For the second set of variables, sweep was considered to interact with the previous variables (denoted as X) as Sweep/X, Sweep*X, and X/Sweep. X/Sweep was chosen because of its superior explanatory power over the other two variables.

Table 3: Volumetric Independent Variables

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LengthxSED^2</td>
<td>Log length*(SED/2)^2</td>
</tr>
<tr>
<td>LengthxLED^2</td>
<td>Log length*(LED/2)^2</td>
</tr>
<tr>
<td>LengthxSEDxLED</td>
<td>Log length*([(SED/2) * (LED/2)]</td>
</tr>
<tr>
<td>Sweep/LengthxSED^2</td>
<td>Sweep/((Log length*(SED/2))^2)</td>
</tr>
<tr>
<td>Sweep/LengthxLED^2</td>
<td>Sweep/((Log length*(LED/2))^2)</td>
</tr>
<tr>
<td>Sweep/LengthxSEDxLED</td>
<td>Sweep/((Log length*(SED/2)*(LED/2))</td>
</tr>
</tbody>
</table>
4.1 Lumber Recovery Factor Model

GWR (2014) present a linear model of LRF as a function of SED, sweep factor, and length for its combination of predictive power, efficiency, simplicity, and ability to satisfy the conditions necessary for statistically valid linear modeling. GWR’s LRF model specification is as follows.\(^9\)

\[
LRF_i = \beta_0 + \beta_1 \ln(SED_i^2) + \beta_2(Sweep Factor_i) + \beta_3(Log Length_i)+u_i
\]

where:
\(\beta_0, \beta_1, \beta_2, \beta_3\) are the parameters to be estimated
\(u_i = \text{error term}\)

The dependent variable of LRF was again chosen as one of the major objectives in lumber manufacturing is to maximize LRF and the yield of the roughmill, while producing a minimum amount of waste.\(^10\) In the present study we estimate **LRF Models** with the following specifications:

1. \[
LRF_i = \beta_0 + \beta_1 SED_i + \beta_2 LED_i + \beta_3 Sweep_i + \beta_4 LogLength_i+u_i
\]
2. \[
LRF_i = \beta_0 + \beta_1 SED_i + \beta_2 LED_i + \beta_3 Sweep_i + \beta_4 LogLength_i + \beta_5 (Length_i \times SED_i^2)
+ \beta_6 (Length_i \times LED_i^2) + \beta_7 (Length_i \times SED_i \times LED_i) + \beta_8 (Sweep_i/Length_i \times SED_i^2)
+ \beta_9 (Sweep_i/Length_i \times LED_i^2) + \beta_{10} (Sweep_i/Length_i \times SED_i \times LED_i)+u_i
\]
3. \[
LRF_i = \beta_0 + \beta_1 \ln(SED_i^2) + \beta_2 (Sweep Factor_i) + \beta_3 (Log Length_i)+u_i
\]
4. \[
LRF_i = \beta_0 + \beta_1 \ln(SED_i^2) + \beta_2 (Sweep Factor_i) + \beta_3 (Length_i) + \beta_4 (Length_i \times SED_i^2)
+ \beta_6 (Length_i \times LED_i^2) + \beta_7 (Length_i \times SED_i \times LED_i) + \beta_8 (Sweep_i/Length_i \times SED_i^2)
+ \beta_9 (Sweep_i/Length_i \times LED_i^2) + \beta_{10} (Sweep_i/Length_i \times SED_i \times LED_i)+u_i
\]

**LRF Model 1** uses the standard independent variables found in **Table 1**, and **Model 2** adds the volumetric variables of **Table 3**. **LRF Model 3** uses GWR’s transformed independent variables found in **Table 2** instead of the standard ones, and **LRF Model 4** adds the volumetric variables of **Table 3**. This specification of the **LRF Models** allows us to see the difference between the independent variable specifications as well as the effect of adding volumetric variables to each model.

4.2 Product Recovery Model

To develop a model that describes how log measurement variables explain the board feet recovery by product, the data were be set up as a panel indexed by log and product type. This is similar to the typical cross-sectional, time-series data format except that the product index takes the place of

---


the time index. The structure of the data were a wide, short panel: there were a large number of logs \((i=1,2,3,...,n)\) but a smaller number of product dimension \((t=1,2,3,...,m)\), where \(m << n\).

In 2014, over 50 unique product dimensions were produced at the UCM, but many of these comprised only a small percentage of annual production volume. For the period between October 16\textsuperscript{th} and December 9\textsuperscript{th}, 2014, the Comact scanner reported 26 unique products with depths and widths as shown in \textit{Table 4}.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Product Dimension & Frequency & Percent & Mean BF & Total BF & % BF \\
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\end{tabular}
\caption{Comact Scanner Data by Product}
\end{table}

We do not report product length above because lumber is sold as random length. Additionally, lumber is sold as random width (RW). As such, the 26 dimensions are grouped and sold in three distinct categories: 4/4 x RW, 3 x 7, and 5/4 x RW. These three categories were used as the product dimension index for the PR model panel regression models. Their observation frequency is summarized below in \textit{Table 5}.
Table 5: Product dimensions $t=1,2,3$

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To allow the log measurement coefficients to vary across product classes, the independent variables were interacted with dummies for the three product dimension classes. The standard errors were clustered by log to provide consistent estimates of the standard errors.

In the present study we estimate the board feet of lumber recovered by product dimension for the following PR Model specifications:

1. $\ln (1 + BF_{it}) = \beta_0 + \beta_1 \text{Dimension}_t + \beta_2 (SED_i \ast \text{Dimension}_t) + \beta_3 (LED_i \ast \text{Dimension}_t) + \beta_4 (\ln \text{Length}_i \ast \text{Dimension}_t) + \beta_5 (\text{Sweep}_i \ast \text{Dimension}_t) + u_{it}$

2. $\ln (1 + BF_{it}) = \beta_0 + \beta_1 \text{Dimension}_t + \beta_2 (SED_i \ast \text{Dimension}_t) + \beta_3 (LED_i \ast \text{Dimension}_t) + \beta_4 (\ln \text{Length}_i \ast \text{Dimension}_t) + \beta_5 (\text{Sweep}_i \ast \text{Dimension}_t) + \beta_6 (\text{Length}_i \ast SED_i^2 \ast \text{Dimension}_t) + \beta_7 (\text{Length}_i \ast LED_i^2 \ast Dimension_t) + \beta_8 (\text{Length}_i \ast SED_i \ast LED_i \ast \text{Dimension}_t) + \beta_9 (\text{Sweep}_i / (\text{Length}_i \ast SED_i^2) \ast \text{Dimension}_t) + \beta_{10} (\text{Sweep}_i / (\text{Length}_i \ast LED_i^2) \ast \text{Dimension}_t) + \beta_{11} (\text{Sweep}_i / (\text{Length}_i \ast SED_i \ast LED_i) \ast \text{Dimension}_t) + u_{it}$

3. $\ln (1 + BF_{it}) = \beta_0 + \beta_1 \text{Dimension}_t + \beta_2 (\ln (SED_i^2) \ast \text{Dimension}_t) + \beta_3 (\text{Sweep Factor}_i \ast \text{Dimension}_t) + \beta_4 (\ln \text{Length}_i \ast \text{Dimension}_t) + u_{it}$

4. $\ln (1 + BF_{it}) = \beta_0 + \beta_1 \text{Dimension}_t + \beta_2 (\ln (SED_i^2) \ast \text{Dimension}_t) + \beta_3 (\text{Sweep Factor}_i \ast \text{Dimension}_t) + \beta_4 (\ln \text{Length}_i \ast \text{Dimension}_t) + \beta_5 (\text{Length}_i \ast SED_i^2 \ast \text{Dimension}_t) + \beta_6 (\text{Length}_i \ast LED_i^2 \ast \text{Dimension}_t) + \beta_7 (\text{Length}_i \ast SED_i \ast LED_i \ast \text{Dimension}_t) + \beta_8 (\text{Sweep}_i / (\text{Length}_i \ast SED_i^2) \ast \text{Dimension}_t) + \beta_9 (\text{Sweep}_i / (\text{Length}_i \ast LED_i^2) \ast \text{Dimension}_t) + \beta_{10} (\text{Sweep}_i / (\text{Length}_i \ast SED_i \ast LED_i) \ast \text{Dimension}_t) + u_{it}$

As with LRF, PR Model 1 uses the standard independent variables found in Table 1, and Model 2 adds the volumetric variables of Table 3. PR Model 3 uses GWR’s transformed independent variables found in Table 2 instead of the standard ones, and PR Model 4 adds the volumetric variables of Table 3. This specification of the PR Models allows us to see the difference between the independent variable specifications as well as the effect of adding volumetric variables to each model.
5. Results

All of the statistical analysis, with the exception of the preliminary analysis, was done using STATA 13 software. Both models were estimated by multivariate regression, with coefficients allowed to vary by product dimension in the product recovery model.

5.1 Lumber Recovery Factor Model Results

The results of the four LRF models described in the methods are found below in Table 6. The independent variables are listed in the first column, and their coefficients and t-statistics are grouped by model in the following columns. The R² and number of observations are found in the bottom two rows. The coefficients are indicated as significant at the 95% confidence level with one *, at the 99% confidence level with two **, and at the 99.9% confidence level with three ***.

From the table, equations can be constructed to predict the LRF. For example, the equation for LRF Model 3 is found below in equation [5].

\[
LRF_i = XX + XX(\ln(SED_i^2)) - XX(Sweep Factor_i) + XX(Log Length_i)
\]  

Comparing LRF Model 1 and 2, the R² increased as expected when the volumetric variables were added, increasing from 0.551 to 0.604. The coefficients on SED both had a large positive coefficient that was significant at a 99.9% confidence level, while the coefficients on Sweep were both negative and significant. The coefficients on Log Length were both positive and small while remaining significant at a 99.9% confidence level. A surprise though, was that the coefficient on LED changed from a small positive number in Model 1 to a negative number of the same magnitude in Model 2.

In LRF Model 3 and 4 transformed independent variables were used and the resulting R² increased when volumetric variables were added, from 0.581 to 0.619. The shared coefficients were all similar in magnitude and sign between the two models. The coefficient on the other hand was small, positive, and insignificant for LRF Model 3, but large, negative, and significant for LRF Model 4.

Using the transformed independent variables for LRF Models 3 and 4 resulted in higher R²s with one less variable than LRF Model 1 and 2.
Table 6: Lumber Recovery Factor Model Results
5.2 Product Recovery Model Results

The results of the three product recovery models described in the methods are found below in Table 7. Again, independent variables are listed in the first column, and their coefficients and T-statistics are grouped by model in the following columns. The R^2 and number of observations are found in the bottom two rows. The coefficients are indicated as significant at the 95% confidence level by one *, at the 99% confidence level by two **, and at the 99.9% confidence level by three ***.

From the table, equations for each product dimension can be constructed to predict the board feet recovered. For example, the equations for Model 3 are found below in equations [6-8].

Measurements for the independent variables can be inserted into each equation to predict recovery, expressed in ln(1+BF). To get to BF, this number is simply exponentiated and one is subtracted from it.

Dimension 1: 4/4 x RW
\[
\ln(1 + BF_{1i}) = XX + XX(\ln SED_{1i}^2) - XX(\text{Sweep Factor}_{1i}) + XX(\text{Log Length}_{1i})
\] [6]

Dimension 2: 3 x 7
\[
\ln(1 + BF_{2i}) = (XX) + (XX)(\ln SED_{2i}^2) + (XX)(\text{Sweep Factor}_{2i}) + (XX)(\text{Log Length}_{2i})
\] [7]

Dimension 3: 5/4 x RW
\[
\ln(1 + BF_{3i}) = (XX) + (XX)(\ln SED_{3i}^2) + (XX)(\text{Sweep Factor}_{3i}) + (XX)(\text{Log Length}_{3i})
\] [8]

The coefficients of the standard independent variables in PR Model 1 were all significant at a 99.9% confidence level, with the exception of LED for Dimension 2, and their sign and magnitude reflected the results from the LRF model. PR Model 2 features the addition of volumetric independent variables. Adding these variables increased the explanatory power of the model, as the R^2 increased from 0.713 to 0.731. The coefficients and their impact on BF recovered remained consistent across the two models with SED having a positive coefficient and Sweep having a negative coefficient.

For PR Model 3, the coefficients were all significant at a 99% confidence level. The coefficient of lnSED2 was positive for each product dimension, meaning that it increases board feet recovered as SED increases. The coefficient on Sweep Factor was also negative for each dimension as well, reflecting the fact that as sweep increases, board feet recovered decreases.

PR Model 4 features the addition of volumetric independent variables. Adding these variables increased the explanatory power of the model, as the R^2 increased from 0.710 to 0.733. The coefficients and their impact on BF recovered remained consistent across the two models with SED having a positive coefficient and Sweep having a negative coefficient.
Table 7: Product Recovery Model Results
6. Discussion and Concluding Remarks

Using Comact scanner data, volumetric variables, and transformed independent variables, an LRF model was created that explained 62% of the variation in the lumber recovery factor at the Upper Columbia Mill. This increased the robustness of the preliminary analysis done by GWR and enables us to better predict how SED, sweep, and log length affects the ability of the mill to efficiently convert logs into lumber.

A Product Recovery Model was developed that can estimate the board feet recovered from a log by dimension, from just a few measurements (the independent variables). The model explains 73% of the variation in board feet recovered by product dimension. The product recovery model has the potential to estimate the volume of individual products recovered from a stand of trees, given a set of bucking assumptions, before they are even cut. Together with market price data by product, the model might also provide an estimate of stand value.

The major determinants of Pacific Albus product value recovery are volume, dimension, and grade. Prices vary by both dimension and grade, and volume determines how much total value is recovered per grade. The product recovery model successfully addresses both volume and product dimension but was unable to address grade. With Pacific Albus, grading is done visually and is nearly impossible to predict without looking a finished product. Instead of trying to predict grade and value, Pacific Albus sales for 2014 were analyzed, and average weighted price was constructed for each dimension as seen in Table 7.

Table 7: Grade Volume/Total Dimension Volume and Average Realization by Dimension in 2014

<table>
<thead>
<tr>
<th>Grade</th>
<th>Volume</th>
<th>Total Dimension</th>
<th>Average Realization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This weighted price gives a more accurate value of each dimension by using the percentages of the volume sold of each grade by dimension and the average price per MBF by dimension. Combined with the Product Recovery Model, a log valuation tool was created in Excel and is seen below in Figure 3. In it, the PR model predicts BF by dimension based on the independent variables. The separate dimension volumes are then multiplied by the average realization per product dimension to give the total value recovered from the log.

**Figure 3**: Log Valuation Tool, Product Recovery Model 3 Results

The model is interactive and the weighted price can be adjusted based on sales expectations by re-allocating the volume percentages of each grade by dimension. Both the predicted volume and weighted averages become more accurate when applied to a large number of trees, as the price is based on averages.

Mill and forest managers could use this Log Valuation Tool to predict the volume and value of a stand of trees or logs before they enter the mill. Knowing this information can help managers make better-informed decisions when it comes to tree farming and mill operations. Before it can be used, the model’s accuracy should be tested by comparing the predicted value against actual sales of a group of trees. Currently, this is a difficult task as logs aren’t tracked from Comact scanner to sale. Furthermore, controlling directly for the different Pacific Albus clones harvested, instead of fixed effects, could enhance the model by accounting for specific characteristics associated with each clone.

If accurate, a model could then be built to determine how tree growth affects log value. One of the most difficult decisions at the Boardman Tree Farm is the choice between harvesting a tree now as opposed to letting it grow in size and value. If a Pacific Albus growth model were incorporated with the Log Valuation tool, a cost benefit analysis could be done to compare the value of immediate harvest against growth and future harvest. These modeling efforts could lead to a fully optimized forest-to-product-market system, maximizing the efficiency of forest and mill operations.
7. Works Cited


