# STATISTICAL PROCEDURES FOR DEVELOPMENT OF REAL - TIME STATISTICAL PROCESS CONTROL (SPC) IN LUMBER MANUFACTURING

by

CHRISTINA STAUDHAMMER

B.Sc., University of California, Davis, 1990
M.Sc., The University of British Columbia, 2000

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Forestry

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 2004

## Abstract

High raw material costs and reduced allowable forest harvest levels have created challenges for the Canadian lumber industry. Sawlogs typically comprise 75% of all the costs in a sawmill and insufficient log availability is a widespread problem. Thus, maximum product value and yield from every log processed is an urgent priority.

Effective statistical process control (SPC) procedures can greatly enhance product value and yield, ensuring accuracy and minimum waste. However, present procedures are manual in nature. The time and effort required means that only small data samples are collected at infrequent intervals, seriously limiting quality control effectiveness. Attempts to implement automated SPC with non-contact laser range sensors (LRS) have thus far had only limited success. Such systems have given frequent false alarms, prompting tolerances to be set excessively wide. Thus, real problems are often missed for extended periods.

The objective of this research was to establish a system for collecting and processing real-time LRS size control data for automated lumber manufacturing. An SPC system was developed that incorporated multi-sensor data filtering procedures, a model with complex structure, and new control charting procedures. The LRS data were first filtered for measurement errors using techniques from image processing. Non-sawing defects were then removed from the data using a sheet-of-light profiling system and defect recognition algorithm. Defect-free filtered data were modeled in a multi-stage process, which explicitly considered multiple sources of variation and a complex correlative structure. New SPC charts were developed that went beyond traditional size control methods, simultaneously monitoring multiple surfaces and specifically targeting common sawing defects.

Nineteen candidate control charts were evaluated. For some sawing defects (e.g., machine positioning errors and wedge), traditional X-bar and range charts are suggested. These charts were explicitly developed to take into account the components of variance in the model. For other sawing defects (e.g., taper, snipe, flare, and snake), control charts are suggested that are non-traditional. The charts that target these defects were based on the decomposition of LRS measurements into trend, waviness, and roughness.

Applying these methods will lead to process improvements in sawmills, so that machines producing defective material can be identified, allowing prompt repairs to be made.

## Table of Contents

## List of Tables

## List of Figures

## List of Nomenclature and Abbreviations

**ACF** autocorrelation function

**AIC** Akaike's information criteria

**ANOVA** analysis of variance

**AR** autoregressive

**ARIMA** autoregressive integrated moving average

**ARIMA($p,\delta,q$)** ARIMA model with $p$ autoregressive parameters, degree of differencing $\delta$, and $q$ moving average parameters

**ARFIMA** autoregressive fractionally integrated moving average

**ARL** average run length

**ARMA** autoregressive moving average

**$B_1$ - $B_4$** bandsaws #1 - #4

**BB** Bandsaw – Bandsaw saw configuration

**BC** Bandsaw – Chipper-head saw configuration

**BIC** Bayesian information criteria

**cant** a log with one or more squared sides

**canterline** a sawing production line where cants are cut from logs

**CB** Chipper-head – Bandsaw saw configuration

**CCF** cross-correlation function

**CCD** charge-coupled device

**$C_L$** left chipper-head

**CL** centreline of control chart

**classification rate** the proportion of correct classifications from the discriminant functions

**confusion matrix** a listing of the number of observations by group that are classified into each possible group

**COV** components of variance

**$C_R$** right chipper-head

**CT** computed tomography

**DTM** digital terrain mapping

**feedspeed** speed at which lumber is traveling when it comes in contact with saw

**flare** a sawing defect in which a triangular-shaped section is added to the end of the sawn lumber

**flitch** an un-edged board

**gangsaw** a power saw that has several parallel blades making simultaneous cuts

**GIS** geographical information system

**iid** independent and identically distributed

**LCL** lower control limit of a control chart

**LIS** left inside sideboard

**LOS** left outside sideboard

**LRS** laser range sensor

**MA** moving average

**machine positioning problems** sawing defect in which sawn boards are too thick or too thin along their entire lengths (also called setworks problems)

**MINIC** minimum information criteria

**MR chart** control chart for monitoring moving ranges

**$MR_\beta$ chart** control chart for monitoring the moving range of successive board averages

**$MR_{\beta\lambda}$ chart** control chart for monitoring the moving range of successive board × laser position averages

**MRI** magnetic resonance imaging

**MSD** moving standard deviation

**n** number of points used in calculating statistic

**OLS** ordinary least squares

**PACF** partial autocorrelation function

**Q chart** control chart based on upper and lower quantiles from a standard or empirical distribution

**$Q_\tau$ chart** control chart for monitoring the slope of measurements within board and side, which is based on upper and lower quantiles from a standard normal distribution

**$Q_{\tau'}$ chart** control chart for monitoring the slope of the last 15 cm of measurements within board and side, which is based on upper and lower quantiles from a standard normal distribution

**$Q_{r_a}$ chart** control chart for monitoring the average roughness of measurements within board and side, which is based on upper and lower quantiles from a standard gamma distribution

**$Q_{r_q}$ chart** control chart for monitoring the RMS roughness of measurements within board and side, which is based on upper and lower quantiles from a standard gamma distribution

**$Q_{r_p}$ chart** control chart for monitoring the peak-to-peak roughness of measurements within board and side, which is based on upper and lower quantiles from a standard gamma distribution

**$Q_{w_a}$ chart** control chart for monitoring the average waviness of measurements within board and side, which is based on upper and lower quantiles from a standard gamma distribution

**$Q_{w_q}$ chart** control chart for monitoring the RMS waviness of measurements within board and side, which is based on upper and lower quantiles from a standard gamma distribution

**$Q_{w_p}$ chart** control chart for monitoring the peak-to-peak waviness of measurements within board and side, which is based on upper and lower quantiles from a standard gamma distribution

**R chart** control chart for monitoring ranges

**$R_{\beta_{grp}}$ chart** control chart for range of subgrouped board averages

**rgb** red green blue

**RIS** right inside sideboard

**$R_{\beta\lambda_{grp}}$ chart** control chart for range of subgrouped averages by board × laser position

**$R_{\lambda_{ind}}$ chart** control chart for monitoring the range of laser position averages within individual boards

**$R_{\lambda_{grp}}$ chart** control chart for monitoring the range of subgrouped averages by laser position

**RMS** root mean square

**RMSE** root mean square error

**ROI1** region above the top LRS (at 106 mm)

**ROI2** region between the two LRSs

**ROI3** region below the bottom LRS (at 22 mm)

**ROS** right outside sideboard

**roughness** the high frequency (short wavelength, or closely spaced) repetitive or random deviations from the "normal" surface

**RR** Circular Saw – Circular Saw configuration

**$R_1$ – $R_{19}$** circular saws #1 - #19

**SARIMA** seasonal autoregressive integrated moving average

**SARFIMA** seasonal autoregressive fractionally integrated moving average

**SAS** statistical software package

**$S_b$ chart** control chart for monitoring the variation between groups of measurements (e.g., boards)

**$S_\beta$ chart** control chart for monitoring the variation due to boards

**$S_{\beta\lambda}$ chart** control chart for monitoring the variation due to the interaction of boards and laser positions

**$S_\lambda$ chart** control chart for monitoring the variation due to laser positions

**$S_w$ chart** control chart for monitoring the variation within groups of measurements (e.g., within boards)

**S1** area within 6.4mm (¼") of the top LRS (100-112 mm from the bottom of the board)

**S2** area within 6.4mm (¼") of the bottom LRS (16-28 mm from the bottom of the board)

**S1**[2] area within 12.8 mm (½") of the top LRS (93-119 mm from the bottom of the board)

**S2**[2] area within 12.8 mm (½") of the bottom LRS (9-35 mm from the bottom of the board)

**snake** sawing defect in which an uneven wave pattern is present on the surface of the board[a]

**snipe** sawing defect in which a triangular-shaped section is removed from the end of sawn lumber

---

[a] For more detailed descriptions and examples of this and other sawing and non-sawing defects, see: Anonymous. 1989. Canadian lumber grading manual. Vancouver, B.C., National Lumber Grades Authority.

**SPC** statistical process control

**SSres** residual sum of squares

**stationarity** a quality of a process in which the statistical parameters (mean and standard deviation) do not change with time

**taper** sawing defect characterized by a gradual increase (or decrease) in thickness along the length of a board

**tear-out** condition in which the saw blade rips the grain on the surface of a workpiece

**UCL** upper control limit of a control chart

**wane** the natural curvature of the edge of a board sawn from a log

**waviness** the medium-to-long frequency (long wavelength) deviations from the "normal" surface

**wedge** sawing defect characterized by a gradual thinning (or thickening) across the width of a board or through its thickness

**X-bar chart** control chart for monitoring average values (e.g., average board thickness)

**X-bar$_{ind}$ chart** control chart for monitoring individual board averages

**X-bar$_{grp}$ chart** control chart for monitoring subgrouped board averages

## List of Symbols

## (in order in which they appear)

### Chapter 1

$b$ number of boards (or other items) measured in periodic samples for SPC

$n$ number of measurements per board (or other item) measured in periodic samples for SPC

$\overline{\overline{X}}$ long-term estimate of the average thickness over all boards and measurement locations

$\hat{\sigma}_{\overline{\overline{X}}}$ long-term estimate of the standard error of the mean

$\hat{\sigma}_w^2$ long-term estimate of the within-board variance

$\hat{\sigma}_b^2$ long-term estimate of the between-board variance

$c_4$ control chart constant that corrects the standard error for bias [b]

$\overline{R}$ average range

$D_3$, $D_4$ control chart constants used in constructing R chart

$B_3$, $B_4$ control chart constant used in constructing $S_w$ chart

$D_{0.001}$, $D_{0.999}$ cumulative probability values for the range at the $0.1^{th}$ and $99.9^{th}$ percentiles

$d_2$ control chart constant that corrects the range for bias

$\chi^2_{(0.001;n-1)}$, $\chi^2_{(0.999;n-1)}$ cumulative probability values for the chi-square distribution with $n$-1 degrees of freedom at the $0.1^{th}$ and $99.9^{th}$ percentiles

$df$ degrees of freedom estimated with Satterthwaite procedure

---

[b] For a discussion of bias correction factors and control chart constants, see, for example, Montgomery (2001).

$MS_b$ between boards (or other grouped measurements) mean squares from a one-way ANOVA

$MS_w$ within board (or other grouped measurements) mean squares from a one-way ANOVA

**Chapter 2**

$S_\sigma$ the "sigma-based probability limit" used in Lee's sigma filter

$l_{km}$ the $m$th LRS measurement from the $k$th sample board

$\bar{l}_{..}$ the average LRS measurement over all boards

$\bar{l}_{k.}$ the average LRS measurement from the $k$th board

$n_k$ the total number of measurements taken on the $k$th board

$w_1$ number of points used in the moving window under filtering Method 1

$w_2$ number of points used in the moving window under filtering Method 2

$MIN$ the minimum number of points needed for re-calculation in Lee's sigma filter

$\tilde{l}_{km}$ moving average of LRS measurements centred around the $m$th point on the $k$th board

$\tilde{\sigma}_{km}$ moving standard deviation (MSD) of LRS measurements centred around the $m$th point on the $k$th board

$min$ the number of points in the $w_2$-point window that fell within the sigma-probability limits

$w_3$ number of points used in the moving window under filtering Method 3

T half the target board thickness value

$\Delta$ maximum allowable deviation from target under Method 3

$\tilde{\sigma}'_{km}$ MSD of LRS measurements centred around the $m$th point on the $k$th board, re-calculated, using the preliminary filtering in Method 3

$\overline{\tilde{\sigma}'_{k\cdot}}$ median value of $\tilde{\sigma}'_{km}$ for $k$th board, re-calculated, using the preliminary filtering in Method 3

$h_{km}$ "hole depth" for $m$th point on the $k$th board

$\hat{\sigma}_{k\cdot}$ standard deviation of measurements from the $k$th board

$\gamma_k(t)$ semivariogram of the LRS measurements from the $k$th board at a distance $t$

**Chapter 3**

$(x,y)$ horizontal and vertical coordinates of pixel locations corresponding to laser line in sheet-of-light profile imaging

$x'$ first derivative along the horizontal direction ($x$), with respect to the vertical direction ($y$)

$num\_miss_1$ number of points missing in ROI1

$num\_miss_2$ number of points missing in ROI2

$num\_miss_3$ number of points missing in ROI3

$num\_miss_{S1}$ number of points missing in S1

$num\_miss_{S2}$ number of points missing in S2

$num\_angled_{S1}$ number of points in S1 where arctan($x'$) > 10 degrees

$num\_angled_{S2}$ number of points in S2 where arctan($x'$) > 10 degrees

$avg\_xprime_{S1}$ average $x'$ value in S1

$avg\_xprime_{S2}$ average $x'$ value in S2

**Chapter 4**

$d$ distance between the Side 1 LRSs and the Side 2 LRSs

$l_{jl_m}$ distance from side $j$ laser $l$ to the board surface at a distance $m$ along the board

$y_{jl_m}$ surface profile value for side $j$ laser $l$ at a distance $m$ along the board

$y_{ijklm}$ profile observation from the $i$th saw configuration, $j$th side, $k$th sample board, $l$th laser location, and $m$th distance along the board

$b_i$ number of boards from the the $i$th saw configuration

$n_{ijkl}$ total number of profile observations from the the $i$th saw configuration, $j$th side, $k$th sample board, $l$th laser location

$\mu_{ij}$ mean profile of the $i$th saw configuration and $j$th side

$\beta_{ijk}$ $k$th board effect from the $i$th saw configuration and $j$th side

$\lambda_{ijl}$ $l$th laser location effect from the $i$th saw configuration and $j$th side

$\beta\lambda_{ijkl}$ interaction of the $k$th sample board and $l$th laser location from the $i$th saw configuration and $j$th side

$\varepsilon_{ijklm}$ error associated with the $m$th measurement from the $l$th laser location, $k$th sample board, in the $i$th saw configuration and $j$th side

$\sigma_{\beta_{ij}}^2$ variance of $\beta_{ijk}$ from the $i$th saw configuration and $j$th side

$\sigma_{\lambda_{ij}}^2$ variance of $\lambda_{ijl}$ from the $i$th saw configuration and $j$th side

$\sigma_{\beta\lambda_{ij}}^2$ variance of $\beta\lambda_{ijkl}$ from the $i$th saw configuration and $j$th side

$\sigma_{\varepsilon_{ij}}^2$ variance of $\varepsilon_{ijklm}$ from the $i$th saw configuration and $j$th side

$\rho_{ijkl_t}$ ACF of profile data from $i$th saw configuration, $j$th side, $k$th sample board, $l$th laser location at lag $t$

$\bar{y}_{ijkl\cdot}$ mean profile for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$AIC_{ijkl}$ Akaike's Information Criteria for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$\ell(\hat{\boldsymbol{\psi}}_{ijkl} \mid \mathbf{y}_{ijkl})$ empirical maximized log likelihood function

$\hat{\boldsymbol{\psi}}_{ijkl}$ row vector of estimated model parameters for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$\mathbf{y}_{ijkl}$ column vector of profile measurements from the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$K$ number of parameters in the model

$f$ shorthand for $ijkl$

$\hat{\mu}_f$ estimated mean value of the $f$th profile

$\delta_f$ degree of differencing for $f$th series

$\phi_f$ autoregressive parameter of $f$th series

$\theta_f$ moving average parameter of $f$th series

$B$ backshift operator

$\alpha_f$ intercept of $f$th series

$\upsilon_{fm}$ white noise error process of $f$th series

$\sigma_{\upsilon_f}^2$ variance of $\upsilon_{fm}$

$\nabla$ difference operator

$\Phi_f$ autoregressive parameter for cyclical (seasonal) behaviour of $f$th series

$\Theta_f$ moving average parameter for cyclical (seasonal) behaviour of $f$th series

$s_f$ cycle length of $f$th series

$\delta'_f$ degree of differencing in cyclical (seasonal) behaviour of $f$th series

$\alpha$ significance level

$z_{fm}$ $m$th profile measurement from the $f$th series, with autocorrelation removed

$v_{ijklm}$ white noise error process

$\sigma_{v_{ij}}^2$ variance of $v_{ijklm}$ for $i$th saw configuration and $j$th side

$\overline{n}_{ij\cdot\cdot}$ average number of observations per board and laser position for $i$th saw configuration and $j$th side

$\sigma_{\bar{y}_{ijkl\cdot}}$ standard error of the mean value, $\overline{y}_{ijkl\cdot}$

$\sigma_{\bar{y}_{ijk\cdot\cdot}}$ standard error of the mean value, $\overline{y}_{ijk\cdot\cdot}$

$\overline{y}_{ijk\cdot\cdot}$ mean profile observation for $i$th saw configuration, $j$th side, and $k$th board

$\sigma_{\bar{y}_{ij\cdot\cdot\cdot}}$ standard error of the mean value, $\overline{y}_{ij\cdot\cdot\cdot}$

$\overline{y}_{ij\cdot\cdot\cdot}$ mean profile observation for $i$th saw configuration and $j$th side

$\hat{y}_{ijklm}$ predicted value of $y_{ijklm}$

$\hat{\beta}_{ijk}$ estimated value of $\beta_{ijk}$

$\hat{\lambda}_{ijl}$ estimated value of $\lambda_{ijl}$

$\hat{\beta}\lambda_{ijkl}$ estimated value of $\beta\lambda_{ijkl}$

$\hat{\alpha}_f$ estimated value of $\alpha_f$

$\hat{\theta}_f$ estimated value of $\theta_f$

$\hat{\phi}_f$ estimated value of $\phi_f$

$\hat{\sigma}_{\upsilon_f}$ estimated value of $\sigma_{\upsilon_f}$

$\hat{s}_f$ estimated value of $s_f$

$\hat{\delta}_f$ estimated value of $\delta_f$

$\hat{\delta}'_f$ estimated value of $\delta'_f$

$\hat{\Phi}_f$ estimated value of $\Phi_f$

$\hat{\Theta}_f$ estimated value of $\Theta_f$

**Chapter 5**

$\bar{y}_{i.k\cdot\cdot}$ average profile value for $i$th saw configuration and $k$th board

$\bar{y}_{i\cdot g\cdots}$ average profile value for $i$th saw configuration and $g$th group of boards

$MR(\bar{y}_{i\cdot k\cdot\cdot})_k$ moving range of successive board averages, $\bar{y}_{i.k\cdot\cdot}$, in $i$th saw configuration, $k$th to ($k$+1)th board

$R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G}$ range of $\bar{y}_{i\cdot gk\cdot\cdot}$ board averages in $g$th group and $i$th saw configuration

$G$ subgroup size

$\bar{y}_{i \cdot gk \cdot}$ average profile value for $k$th board, in $i$th saw configuration and $g$th group of boards

$S^2_{\beta_{ijg}}$ variation due to board effects for $i$th saw configuration, $j$th side, $g$th group of boards

$R(\bar{y}_{ijkl \cdot})\big|^2_{l=1}$ range of $\bar{y}_{ijkl \cdot}$ laser position averages for $k$th board in $i$th saw configuration and $j$th side

$R(\bar{y}_{ijg \cdot l \cdot})\big|^2_{l=1}$ range of $\bar{y}_{ijg \cdot l \cdot}$ laser position averages in $g$th group, $i$th saw configuration, and $j$th side

$\bar{y}_{ijg \cdot l \cdot}$ average profile value for $l$th laser position in $i$th saw configuration, $j$th side, and $g$th group of boards

$MR(\bar{y}_{ijkl \cdot})_k$ moving range of $k$th to $(k+1)$th successive board × laser position averages, $\bar{y}_{ijkl \cdot}$, in $i$th saw configuration, $j$th side, and $l$th laser position

$R(\bar{y}_{ijgkl \cdot})\big|^G_{k=1}$ range of $\bar{y}_{ijgkl \cdot}$ board × laser position averages in $l$th laser position and $g$th group, $i$th saw configuration, and $j$th side

$\bar{y}_{ijgkl \cdot}$ average profile value for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position, in $g$th group of boards

$S^2_{\lambda_{ijg}}$ variation due to laser position for $i$th saw configuration, $j$th side, and $g$th group of boards

$S^2_{\beta\lambda_{ijg}}$ variation due to board × laser position interaction for $i$th saw configuration, $j$th side, and $g$th group of boards

$\hat{\tau}_{1_{ijk}}$ estimated slope of measurements in the horizontal direction along the board from the $i$th saw configuration, $j$th side, and $k$th board

$\hat{\tau}'_{1_{ijk}}$ estimated slope of measurements in the horizontal direction along the board, for the last 15 cm of board from the $i$th saw configuration, $j$th side, and $k$th board

$\overline{r_{ijk\cdot}^{(a)}}$ average of arithmetic average roughness values for $i$th saw configuration, $j$th side, and $k$th board

$\overline{r_{ijk\cdot}^{(q)}}$ average of RMS roughness values for $i$th saw configuration, $j$th side, and $k$th board

$\overline{r_{ijk\cdot}^{(p)}}$ average of peak-to-peak roughness values for $i$th saw configuration, $j$th side, $k$th board

$\overline{w_{ijk\cdot}^{(a)}}$ average of arithmetic average waviness values for $i$th saw configuration, $j$th side, and $k$th board

$\overline{w_{ijk\cdot}^{(q)}}$ average of RMS waviness values for $i$th saw configuration, $j$th side, and $k$th board

$\overline{w_{ijk\cdot}^{(p)}}$ average of peak-to-peak waviness values for $i$th saw configuration, $j$th side, and $k$th board

$T_i$ target surface profile value (half the thickness value) for the $i$th saw configuration

$\hat{\sigma}_{(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})/2}$ standard error of the average profile value by board for the $i$th saw configuration

$\hat{\sigma}^2_{\beta_{ij}ij}$ estimate of $\sigma^2_{\beta_{ij}ij}$

$\hat{\sigma}^2_{\lambda ij}$ estimate of $\sigma^2_{\lambda ij}$

$\hat{\sigma}^2_{\beta\lambda_{ij}}$ estimate of $\sigma^2_{\beta\lambda_{ij}}$

$\hat{\sigma}_{(\bar{y}_{i1g\cdot\cdot}+\bar{y}_{i2g\cdot\cdot})/2}$ standard error of the average profile value for a subgroup of boards in the $i$th saw configuration

$\bar{y}_{ijg\cdot\cdot}$ average profile value for $i$th saw configuration and $j$th side, in $g$th group of boards

$\overline{MR(\bar{y}_{i\cdot k\cdot\cdot})_k}_{\ i}$ average of $MR(\bar{y}_{i\cdot k\cdot\cdot})_k$ values in the $i$th saw configuration

$\overline{R(\bar{y}_{i\cdot gk\cdot\cdot})\big|_{k=1}^{G}}_{\ i}$ average of the $R(\bar{y}_{i\cdot gk\cdot\cdot})\big|_{k=1}^{G}$ values in the $i$th saw configuration

$MS_{\beta_{ijg}}$ mean squares due to boards, for the $i$th saw configuration, $j$th side, $g$th group of boards

$MS_{\beta\lambda_{ijg}}$ mean squares due to board × laser interaction, for the $i$th saw configuration, $j$th side, $g$th group of boards

$\overline{n}_{ijg\cdot l}$ average number of observations for the $i$th saw configuration, $j$th side, $g$th group of boards, and $l$th laser position

$df(\beta_{ij})_G$ estimated degrees of freedom of the Chi-square distribution for $\sigma^2_{\beta_{ij}}$ in the $i$th saw configuration and $j$th side, with subgroup size $G$

$MS_{\beta_{ij}}$ mean squares due to boards, for the $i$th saw configuration and $j$th side

$MS_{\beta\lambda_{ij}}$ mean squares due to board × laser interaction, for the $i$th saw configuration and $j$th side

$\overline{R(\overline{y}_{ijkl\cdot})\big|^2_{l=1}}_{ij}$ average of all $R(\overline{y}_{ijkl\cdot})\big|^2_{l=1}$ values for the $i$th saw configuration and $j$th side

$\overline{R_l(\overline{y}_{ijg\cdot l\cdot})\big|^2_{l=1}}_{ij}$ average of all $R(\overline{y}_{ijg\cdot l\cdot})\big|^2_{l=1}$ values for the $i$th saw configuration and $j$th side

$\overline{MR(\overline{y}_{ijkl\cdot})_k}_{ijl}$ average of all $MR(\overline{y}_{ijkl\cdot})_k$ values for the $i$th saw configuration, $j$th side, and $l$th laser position

$\overline{R(\overline{y}_{ijgkl\cdot})\big|^G_{k=1}}_{ijl}$ average of the $R(\overline{y}_{ijgkl\cdot})\big|^G_{k=1}$ values in the $i$th saw configuration, $j$th side, and $l$th laser position

$MS_{\lambda_{ijg}}$ mean squares for laser, for the $i$th saw configuration, $j$th side, $g$th group of boards

$\overline{n}_{ijgk\cdot}$ average number of observations for the $i$th saw configuration, $j$th side, $g$th group, and $k$th board

$df(\lambda_{ij})_G$ estimated degrees of freedom of the Chi-square distribution for $\sigma^2_{\lambda_{ij}}$ in the $i$th saw configuration, and $j$th side, with subgroup size $G$

$MS_{\varepsilon_{ij}}$ mean squares for residual, for the $i$th saw configuration and $j$th side

$\overline{n}_{ijg\cdot\cdot}$ average number of observations for the $i$th saw configuration, $j$th side, and $g$th group

$df(\beta\lambda_{ij})_G$ estimated degrees of freedom of the Chi-square distribution for $\sigma^2_{\beta\lambda_{ij}}$ in the $i$th saw configuration and $j$th side, with subgroup size $G$

$x_1$ horizontal position of the laser measurement along the board (cm)

$x_2$ vertical position of the laser measurement on the board (cm)

$\tau_{0_{ijk}}$, $\tau_{1_{ijk}}$, and $\tau_{2_{ijk}}$ coefficients from a regression of $y_{ijklm}$ versus $x_1$ and $x_2$, for the $i$th saw configuration, $j$th side, and $k$th board

$\zeta_{ijklm}$ prediction error associated with the a regression of $y_{ijklm}$ versus $x_1$ and $x_2$

$\sigma_{\zeta_{ij}}^{\ 2}$ variance of $\zeta_{ijklm}$ for the $i$th saw configuration and $j$th side

$\hat{\sigma}_{\hat{\tau}1_{ijk}}$ standard error of the estimated $\hat{\tau}_{1_{ijk}}$ values for the $i$th saw configuration and $j$th side

$\overline{\hat{\tau}}_{1_{ij\cdot}}$ average of all $\hat{\tau}_{1_{ijk}}$ values

$\hat{\sigma}_{\hat{\tau}'1_{ijk}}$ standard error of the estimated $\hat{\tau}'_{1_{ijk}}$ values for the $i$th saw configuration and $j$th side

$\overline{\hat{\tau}}'_{1_{ij\cdot}}$ average of all $\hat{\tau}'_{1_{ijk}}$ values

$y'_{ijklm}$ de-trended surface profile for $i$th saw configuration, $j$th side, $k$th board, $l$th laser position, $m$th distance along the board

$w_{ijklm}$ waviness for $i$th saw configuration, $j$th side, $k$th board, $l$th laser position, $m$th distance along the board

$r_{ijklm}$ roughness for $i$th saw configuration, $j$th side, $k$th board, $l$th laser position, $m$th distance along the board

$r_{ijkl}^{(a)}$ arithmetic average roughness for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$\bar{r}_{ijkl\cdot}$ average of all roughness values ($r_{ijklm}$) for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$r_{ijkl}^{(q)}$ root mean square (RMS) roughness for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$r_{ijkl}^{(p)}$ peak-to-peak roughness for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$w_{ijkl}^{(a)}$ arithmetic average waviness for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$w_{ijkl}^{(q)}$ root mean square (RMS) waviness for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$w_{ijkl}^{(p)}$ peak-to-peak waviness for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$\bar{w}_{ijkl\cdot}$ average of all waviness values ($w_{ijklm}$) for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$\overline{w_{ijk\cdot}^{(p)}}$ average value of $w_{ijk1}^{(p)}$ and $w_{ijk2}^{(p)}$

$Q_{ij}\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{0.1\%}$, $Q_{ij}\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{99.9\%}$, and $Q_{ij}\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{50\%}$ lower 0.1% quantile, upper 0.1% quantile, and median of distribution of peak-to-peak waviness by board for $i$th saw configuration and $j$th side

$Q_{ij}\left(\overline{w_{ijk\cdot}^{(a)}}\right)_{0.1\%}$, $Q_{ij}\left(\overline{w_{ijk\cdot}^{(a)}}\right)_{99.9\%}$, and $Q_{ij}\left(\overline{w_{ijk\cdot}^{(a)}}\right)_{50\%}$ lower 0.1% quantile, upper 0.1% quantile, and median of distribution of average waviness by board for $i$th saw configuration and $j$th side

$Q_{ij}\left(\overline{w_{ijk\cdot}^{(q)}}\right)_{0.1\%}$, $Q_{ij}\left(\overline{w_{ijk\cdot}^{(q)}}\right)_{99.9\%}$, and $Q_{ij}\left(\overline{w_{ijk\cdot}^{(q)}}\right)_{50\%}$ lower 0.1% quantile, upper 0.1% quantile, and median of distribution of RMS waviness by board for $i$th saw configuration and $j$th side

$Q_{ij}\left(\overline{r_{ijk\cdot}^{(a)}}\right)_{0.1\%}$, $Q_{ij}\left(\overline{r_{ijk\cdot}^{(a)}}\right)_{99.9\%}$, and $Q_{ij}\left(\overline{r_{ijk\cdot}^{(a)}}\right)_{50\%}$ lower 0.1% quantile, upper 0.1% quantile, and median of distribution of average roughness by board for $i$th saw configuration and $j$th side

$Q_{ij}\left(\overline{r_{ijk\cdot}^{(q)}}\right)_{0.1\%}$, $Q_{ij}\left(\overline{r_{ijk\cdot}^{(q)}}\right)_{99.9\%}$, and $Q_{ij}\left(\overline{r_{ijk\cdot}^{(q)}}\right)_{50\%}$ lower 0.1% quantile, upper 0.1% quantile, and median of distribution of RMS roughness by board for $i$th saw configuration and $j$th side

$Q_{ij}\left(\overline{r_{ijk\cdot}^{(p)}}\right)_{0.1\%}$, $Q_{ij}\left(\overline{r_{ijk\cdot}^{(p)}}\right)_{99.9\%}$, and $Q_{ij}\left(\overline{r_{ijk\cdot}^{(p)}}\right)_{50\%}$ lower 0.1% quantile, upper 0.1% quantile, and median of distribution of peak-to-peak roughness by board for $i$th saw configuration and $j$th side

$B_{ijk}$ random board effect for $i$th saw configuration, $j$th side, and $k$th board

$L_{ijl}$ random laser position effect for $i$th saw configuration, $j$th side, and $l$th laser position

$BL_{ijkl}$ random board × laser effect for $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position

$\tilde{\bar{y}}_{ijkl}$ simulated average profile for the $k$th board and $l$th laser position in the $i$th saw configuration and $j$th side

$e_{ijklm}$ simulated autocorrelated errors for $i$th saw configuration, $j$th side, $k$th board, $l$th laser position, and $m$th measurement along the board

$\tilde{y}_{ijklm}$ simulated $m$th profile for the $k$th board and $l$th laser position in the $i$th saw configuration and $j$th side

$u_{ijkl_m}$ random number for $i$th saw configuration, $j$th side, $k$th board, $l$th laser position, and $m$th measurement along the board

$\bar{\hat{\sigma}}^2_{\upsilon_{ij\cdot\cdot}}$ average of the estimated white noise error process variance parameters for $i$th saw configuration and $j$th side

$\bar{\hat{\phi}}_{ij\,ij}$ average of the estimated autoregressive parameters for $i$th saw configuration and $j$th side

$\bar{\hat{\theta}}_{ij}$ average of the estimated moving average parameters for $i$th saw configuration, $j$th side

$\overline{\hat{\alpha}}_{ij}$ average of the estimated intercept parameters for $i$th saw configuration, and $j$th side

$\Delta_m$ simulated machine positioning defect deviation

$\Delta_w$ simulated wedge defect deviation

$\Delta_t$ simulated taper defect deviation

$\Delta_f$ simulated snipe/flare defect deviation

$P$ period of simulated snake defect

$A$ amplitude of simulated snake defect

**Appendix 1**

$p$ number of AR parameters in an ARMA or ARIMA process for a single series

$q$ number of MA parameters in an ARMA or ARIMA process for a single series

$\phi_1, \phi_2, \ldots, \phi_p$ AR parameters for a single series

$\theta_1, \theta_2, \ldots, \theta_q$ MA parameters for a single series

$\alpha$ intercept for a single series

$\upsilon_m$ white noise error process for a single series

$\sigma_\upsilon^2$ variance of $\upsilon_m$

$\delta$ degree of differencing in an ARIMA process for a single series

$P$ number of seasonal AR parameters in a SARIMA process for a single series

$Q$ number of seasonal MA parameters a SARIMA process for a single series

$\Phi_1, \Phi_2, \ldots, \Phi_P$ seasonal AR parameters for a single series

$\Theta_1, \Theta_2, \ldots, \Theta_Q$ seasonal MA parameters for a single series

$\delta$ degree of seasonal differencing in a SARIMA process for a single series

$p_f$ number of AR parameters in an ARIMA process for the $f$th series

$q_f$ number of MA parameters in an ARIMA process for the $f$th series

$\delta_f$ degree of differencing in an ARIMA process for the $f$th series

$P_f$ number of seasonal AR parameters in a SARIMA process for the $f$th series

$Q_f$ number of seasonal MA parameters a SARIMA process for the $f$th series

$s_f$ seasonal period of the $f$th series

$\delta_f$ degree of seasonal differencing in an SARIMA process for the $f$th series

$I(\omega_m)$ periodogram of ARFIMA process for a single series

$\omega_m$ frequency: $\omega = m\pi/n$, $m=0, 1, \ldots, n$

$n$ total number of observations in a single series

$f_u(\omega_m)$ spectral density of an ARMA($p,q$) process

$g(\omega_m)$ spectral density of an ARFIMA process

$f(\omega_m)$ $\frac{\sigma_\upsilon^2}{2\pi}\left(4\sin^2\frac{\omega_m}{2}\right)^{-\delta}$

$Y_m$ $\ln(I(\omega_m))$

$X_{1_m}$ $\ln(4\sin^2(\omega_m/2))$

$c_0, c_1$ regression coefficients from fitting $X_{1_m}$ versus $Y_m$

$X_{2_m}$ $\ln(4\sin^2(s\omega_m/2))$

$\phi_{1_f}$ first-order AR parameter for *f*th series

$\theta_{1_f}$ first-order MA parameter for *f*th series

$\hat{\theta}_{1_f}$ estimated value of $\theta_{1_f}$

$\hat{\phi}_{1_f}$ estimated value of $\phi_{1_f}$

$\Phi_{1_f}$ first-order seasonal AR parameter for *f*th series

$\Theta_{1_f}$ first-order seasonal MA parameter for *f*th series

$\hat{\Phi}_{1_f}$ estimated value of $\Phi_{1_f}$

$\hat{\Theta}_{1_f}$ estimated value of $\Theta_{1_f}$

## Acknowledgements

I would like to acknowledge several individuals and organizations without whose help this thesis would not have been written. First, I would like to thank my co-supervisors, Drs. Maness and Kozak. I enjoyed their warm camaraderie, as well as their prudent academic advice. The other members of my committee were equally helpful. I would like to thank Dr. Joe for providing me with quick insights and thought-provoking questions. I would also like to especially thank Dr. LeMay, who gave me seemingly endless amounts of her time and always sound advice.

This research was funded by the Natural Science and Engineering Research Council of Canada, the British Columbia Science Council, and the Izaak Walton Killam Foundation. Their generous support gave me the ability to have access to the latest technology, present my work at international conferences, and avoid student loans. I am grateful to these organizations and the people who make their continued research support possible.

Both Weyerhaeuser Company and Canadian Forest Products were extremely helpful in providing a real environment in which to test the technology used in this research. I would like to thank them for their patience in letting a very green student use their facilities and learn from them.

On a personal note, there were many friends, colleagues, and relatives whose support was constant. Despite repeated inquiries of "when will you be done", my family was unwavering in their econcouragement. They could not fathom that I would not finish, and that was sometimes all I needed. I am also thankful for Tony Kozak and Guillaume Thérien, who gave me guidance and sound SAS advice, as well as dear friendship. Finally, I want to especially thank Jay Shapka for being not just a sounding board, but a real contributor to my work. He has a place in my heart and a direct line to my brain. Thank-you Jay for pushing me, encouraging me, making me think critically, and giving me what I needed at the right time.

# Chapter 1 Introduction

In modern commodity sawmills, the efficient production of dimensionally accurate lumber with minimal waste is the major manufacturing objective. Variation in width and thickness of sawn boards occurs as a result of inaccurate sawing. This variation is usually caused by movement in the saws or the log hold down mechanisms during the cut, or by movement in the saw or log positioning just prior to the cut. Reducing the amount of sawing variation maximizes log recovery and can substantially increase profits (Wang 1983; Maness and Lin 1995; Lister 1997).

In a typical mill, logs travel at 100 metres/minute on a specialized conveyor (infeed), passing through computerized scanning stations that determine the optimal sawing pattern. The computer controls the saw position mechanism, or setworks, to move the saws in the proper position for the optimal cut. Common sawing defects (also called shape defects) can occur because of movement in the setworks, or because of worn parts, poor alignment, or uneven pressure during the cut (Rasmussen *et al.* 2004). Quality problems such as these result in high within-board and between-board variability.

Lumber size control systems based on the concepts of statistical process control (SPC) have been developed that help sawmills monitor the quality of sawing, thereby reducing sawing variation. A basic size control system may involve manually taking 1 or 2 samples per shift from each sawing machine, with a sample consisting of 5-10 boards. Thickness and/or width are measured on each board in 6-10 places using digital calipers. Statistics, such as the sample average and the within- and between-board variance are then calculated and plotted on Shewhart control charts (Shewhart 1931).

Real-time size control systems using non-contact laser measuring systems have recently become available to sawmills. At normal mill operating speeds, these systems have the capability to take

more than ten measurements per centimetre of lumber sawn. Thousands of measurements can be taken on each and every board processed, and when multiple lasers are used, a three-dimensional profile of each board could potentially be produced. Although many mills have already installed these systems, functional methods for utilizing the wealth of data they generate have yet to be developed. The statistical procedures associated with these new measuring systems[1] are largely based on "traditional" procedures developed in the 1970's and 1980's; however, using these procedures can lead to false indications of an out of control process. Laser data must first be filtered, as it may contain measurement errors due to inadequate reflectance and other anomalous measurements from natural non-sawing defects, such as loose knots. Moreover, there is a greater chance of Type I and Type II errors because of the increased sampling intensity involved in real-time data collection and variance under-estimation due to autocorrelation (Wheeler 1995). Therefore, a new method of statistical process control must be developed to perform SPC for real-time laser data.

## 1.1 Context

This thesis is part of a larger research project underway at the Faculty of Forestry and Department of Mechanical Engineering at The University of British Columbia. The goals of this three-year project are to develop improved data acquisition and analysis techniques, which will form the basis of an automated expert system to diagnose and correct problems in real time. The goal of this thesis is to develop a statistical model for this expert system and demonstrate its use.

## 1.2 Background

SPC in wood products manufacturing is unique in the types of variability present and the sampling schemes used to monitor it. The within- and between-board components of variation

---

[1] Real-time LRS-based systems are available from SiCam, SizeCheck, SeeCon, and others.

are important to identify and monitor because of their connection to identifying machining problems that result in common sawing defects. Furthermore, real-time sampling is not yet commonplace, making the gap between the SPC methods used in the industry and those required for real-time application large. There is a great need for additional work that follows from the summary below.

### 1.2.1 Common Sawing Defects

Lumber shape defects occur frequently in the sawing process, and have a variety of causes. Figure 1-1 shows a normal board versus five common defects that are identifiable with laser scanning technology (Rasmussen *et al.* 2004). All of these defects have obvious consequences for production mills, where wood products are made to meet specific customer demands.

Machine positioning or setworks problems occur when saw guides are not set to the correct distance, causing sawn boards to be too thick or too thin along the entire length of the board. This defect can occur because of software problems, worn parts, or improper pressure applied to saw guides (Maness *et al.* 2003). Whereas machine positioning problems tend to cause a uniform change across the width and length of the board, wedge is characterized by an unevenly sawn surface. Wedge often occurs when the saws are misaligned, causing a thickening (or thinning) from the bottom to the top of the board which is consistent along the length of the board (Rasmussen *et al.* 2004). Like wedge, taper occurs when there is machine misalignment. In the case of taper, there is a gradual thickening or thinning along the length of the board.

Flare or snipe can occur when the hold-down rolls do not engage at the proper time. If the feedroll engages too soon, the log (or cant) will be misaligned when entering the saws (Rasmussen *et al.* 2004). Snipe occurs when this misalignment causes a triangular-shaped section to be removed from the end of the sawn lumber. Flare occurs when a triangular shaped section is added to the end of the sawn lumber.

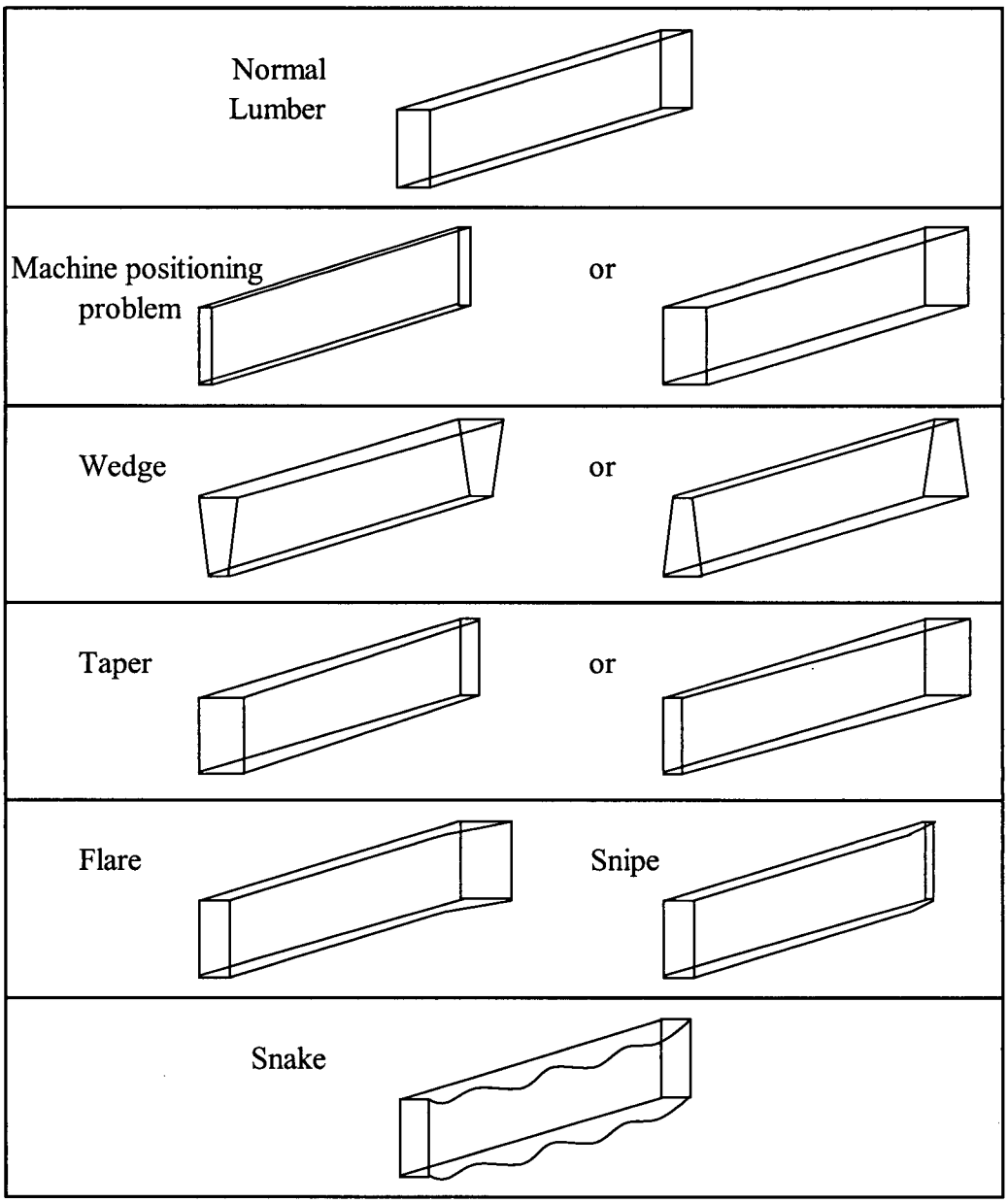



**Figure 1-1. Normal sawing versus five types of sawing defects.**

Snake is a term used to describe a variety of sawing problems that result in high within board variability. For instance, snake can occur when the saw operating speed is near the "critical speed" (Schajer 1989). In this case, the vibration of the unstable saw causes an uneven wave pattern on the surface of the lumber. Snake can also occur due to incorrect tensioning or other saw maintenance issues.

### 1.2.2 Current Methods for SPC in Automated Lumber Manufacturing

SPC was not introduced to the softwood lumber industry until the 1970's (Brown 1982). Current methods rely heavily on the basic methods of sampling and monitoring that were derived by Warren (1973) and Whitehead (1978). Typically, Shewhart control charts are used to monitor the average board thickness, as well as some measure of dispersion such as the range or the variance. Some innovations have been suggested. Brown (1979) partitioned the variation so that it was related to its location along the length of a board; Wang (1984) accounted for multiple sources of variation by analyzing each surface of the board separately. However, the lumber industry has experienced little change in operational SPC practices since their introduction. A brief discussion and summary of current SPC methods follows; a full review is given in Maness *et al.* (2002).

Under current methods, SPC personnel periodically sample a subgroup of $b$ boards from a sawing machine and, using digital calipers, measure board thickness in $n$ places. The measurements are typically entered into an SPC software package that plots the mean, range, and within- and between-board variance from the subgroup on control charts. One important issue is, that while industrial SPC software packages have given adequate results using manual sampling at long-spaced intervals of time, Maness *et al.* (2002) indicated that incorrect methods are used to estimate the basic components of variance in the sawing process. This has resulted in out of control signals being generated even though the process is in control. Although this is a serious model flaw, it was not detectable until the recent introduction of measuring devices that provide more frequent sample points. The theoretical impact of adopting the correct methodology was found to be large (Maness *et al.* 2003), and was verified by a designed experiment with simulated lumber data (Maness *et al.* 2004).

Using the correct components of variance approach, measurements essentially form a one-way analysis of variance (ANOVA) (Maness *et al.* 2002), and the X-bar chart for subgroup averages is constructed using the following control limits (Maness *et al.* 2003):

$$\begin{aligned} CL &= \overline{\overline{X}} \\ LCL &= \overline{\overline{X}} - 3\hat{\sigma}_{\overline{\overline{X}}} / c_4 \\ UCL &= \overline{\overline{X}} + 3\hat{\sigma}_{\overline{\overline{X}}} / c_4 \end{aligned} \qquad [1\text{-}1]$$

where: $CL$ is the centreline;

$LCL$ is the lower control limit;

$UCL$ is the upper control limit;

$\overline{\overline{X}}$ is the long-term estimate of the average thickness over all boards and measurement locations;

$$\hat{\sigma}_{\overline{\overline{X}}} = \sqrt{\frac{\hat{\sigma}_b^2}{b} + \frac{\hat{\sigma}_w^2}{nb}};$$

$\hat{\sigma}_w^2$ and $\hat{\sigma}_b^2$ are long-term estimates of the within- and between-board variances, respectively; and

$c_4$ is a control chart constant that corrects the standard error for bias[2].

These so-called "3-sigma limits" were developed by Shewhart not based on any particular probabilistic model (Nelson 1999); however, if a normal universe is assumed, the arbitrary distance of 3 times the standard error of the mean ($\hat{\sigma}_{\overline{\overline{X}}}$) corresponds to a Type I error (the probability of a "false alarm") of 0.27%.

---

[2] $c_4$, $d_2$, $D_3$, and $D_4$ are control chart constants. For detailed derivation of these constants, see, for example, Montgomery (2001).

To monitor process dispersion, charts are made for the subgroup range $R$ (R chart), and/or the within- and between-board standard deviations ($S_w$ and $S_b$ charts, respectively). Control limits for the R chart are computed as (Wheeler 1995):

$$\begin{aligned} CL &= \overline{R} \\ LCL &= D_3\overline{R} \\ UCL &= D_4\overline{R} \end{aligned} \qquad [1\text{-}2]$$

where: $\overline{R}$ is the long term average range of board thickness values for groups of boards; and

$D_3$ and $D_4$ are control chart constants.

Control limits for the $S_w$ chart are computed with the long-term estimate of the within-board standard deviation ($\hat{\sigma}_w$) (Wheeler 1995):

$$\begin{aligned} CL &= \hat{\sigma}_w \\ LCL &= B_3\hat{\sigma}_w \\ UCL &= B_4\hat{\sigma}_w \end{aligned} \qquad [1\text{-}3]$$

where: $B_3$ and $B_4$ are control chart constants.

Note that the limits given for in [1-2] and [1-3] are 3-sigma limits, which assume a normal (Gaussian) distribution. The normal distribution is symmetric, whereas the distributions of the range and standard deviation are known to be highly asymmetric (Ryan 1989). Therefore, the average run lengths (ARL) for these charts are not the same as that of the X-bar chart. That is, the expected numbers of samples before an out of control is signaled, given that the process is in control, are quite different. Assuming a normally distributed population, the ARL for the X-bar chart is 1/0.0027 $\approx$370; the corresponding ARL for the R chart is 1/0.0092 $\approx$109 (Nelson 1999). In order to have comparatively similar ARLs for these charts, control limits can be constructed with "probability limits". Since 0.27% limits are not readily available for the range, 0.2% limits are commonly used:

$$CL = \overline{R}$$
$$LCL = \overline{R}\left( D_{0.001} / d_2 \right) \qquad [1\text{-}4]$$
$$UCL = \overline{R}\left( D_{0.999} / d_2 \right)$$

where: $D_{0.001}$ and $D_{0.999}$ are cumulative probability values for the range (Harter 1960); and

$d_2$ is a control chart constant that corrects the range for bias.

The 0.2% control limits for the $S_w$ chart are computed as (Ryan 1989):

$$CL = \hat{\sigma}_w / c_4$$
$$LCL = \hat{\sigma}_w / c_4 \left( \chi^2_{(0.001;n-1)} / n-1 \right) \qquad [1\text{-}5]$$
$$UCL = \hat{\sigma}_w / c_4 \left( \chi^2_{(0.999;n-1)} / n-1 \right)$$

where: $\chi^2_{(0.001;n-1)}$ and $\chi^2_{(0.999;n-1)}$ are cumulative probability values for a chi-square distribution with $n$-1 degrees of freedom.

Control limits for the $S_b$ chart cannot be found exactly, as the long-term estimate of the between-board standard deviation, $\hat{\sigma}_b$, is a linear combination of mean squares. However, they can be estimated using the Satterthwaite procedure (Gaylor and Hopper 1969; Maness *et al.* 2004):

$$CL = \hat{\sigma}_b$$
$$LCL = \hat{\sigma}_b \sqrt{\chi^2_{(\alpha/2;df)} / (df)} \qquad [1\text{-}6]$$
$$UCL = \hat{\sigma}_b \sqrt{\chi^2_{(1-\alpha/2;df)} / (df)}$$

The degrees of freedom (*df*) are estimated using the between- and within-board mean squares from the one-way ANOVA, $MS_b$ and $MS_w$[3]:

---

[3] For derivation and calculation of one-way ANOVA mean squares, see Maness *et al.* (2002).

$$df = \frac{(n\hat{\sigma}_b^2)^2}{\frac{MS_b^2}{b-1} + \frac{MS_w^2}{b(n-1)}} \qquad [1\text{-}7]$$

Little has changed in wood products SPC in the last ten years. Three published papers (Cook 1992; Young and Winistorfer 2001; Noffsinger and Anderson 2002) specifically addressed the changing conditions in which SPC is applied in the area of wood composites processing. Preliminary research from this project (Maness *et al.* 2002; Maness *et al.* 2003; Maness *et al.* 2004) was published to quantify long-existing errors in the SPC methods commonly used in lumber mills. As processes move toward more frequent sampling and continuous process monitoring, wood products SPC methods have not kept pace with important issues, such as the increased volume of data, additional sources of variation, autocorrelation, and modeling of multiple attributes. There is a need to update wood products SPC methods for new technologies.

### 1.2.3 Real-time SPC in Automated Lumber Manufacturing

#### 1.2.3.1 Laser Devices for Real-time Size Measurement

In the demanding environment of an industrial sawmill, accurate measurement of lumber surfaces has long been a difficult task. Wood surfaces are complex because of variation in density, moisture content, fiber direction, and the quality of cutting tools. In-line systems have been limited by the need for accurate measurements and by the speed of automated lumber production (Sandak *et al.* 2003). Recent advances in laser technology have gone far in overcoming these difficulties and made accurate and affordable laser range sensors (LRS) available to sawmills. Typically, LRSs use optical triangulation to measure distance to an object[4]. Laser beams are projected from a sensor onto the object, and the laser spot is reflected

[4] LRSs of this type are manufactured by Hermary Opto Electronics, Dynavision/LMI, Keyence, Turck, and others.

from the object on an electronic camera. The distance to the object is computed from the position of the reflected image on the camera (Hermary Opto Electronics Inc. 2002).

Real-time systems can be set up to measure sawn wood as it leaves a sawing machine (Figure 1-2). In the schematic, four LRSs are mounted to allow two streams of measurements on each side of a cant. Typically, side 1-laser 1 and side 2-laser 1 are mounted at least 2.5 cm (1 inch) from the bottom of the cant, and side 1-laser 2 and side 2-laser 2 are mounted at least 2.5 cm (1 inch) from the top of the cant, so that areas of wane[5] are avoided.

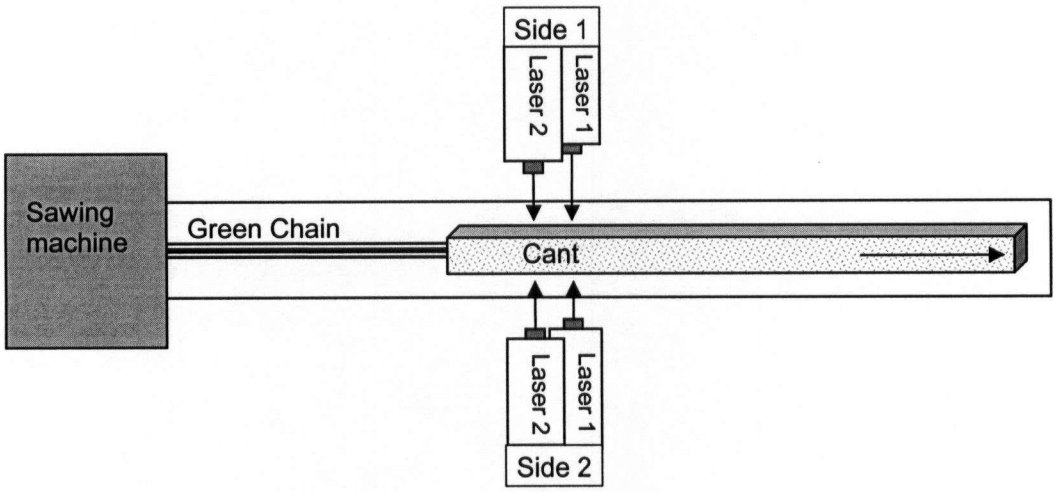



**Figure 1-2. Schematic for a real-time system using four laser range sensors.**

LRS accuracy can be affected by extreme changes in temperature and light (Kraus and Pfeifer 1998) and by the distance between the LRS and its target. The distance-dependent precision of the LRS is referred to by its manufacturers as its "resolution". For instance, the Hermary LRS-50 has a resolution of 25 $\mu$m (0.001 inch) at a distance of 14 ½ cm (5 ¾ inches). Increasing the distance to 90 cm (35 ½ inches) decreases resolution to 0.51 mm (0.020 inch) (Figure 1-3). To obtain fast response time from an SPC system, laser scanning devices should be installed

[5] Wane is the natural curvature of the edge of a board sawn from a log. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

immediately after the outfeed of the sawing machine being monitored. However, these systems should also be designed with considerations for laser accuracy.

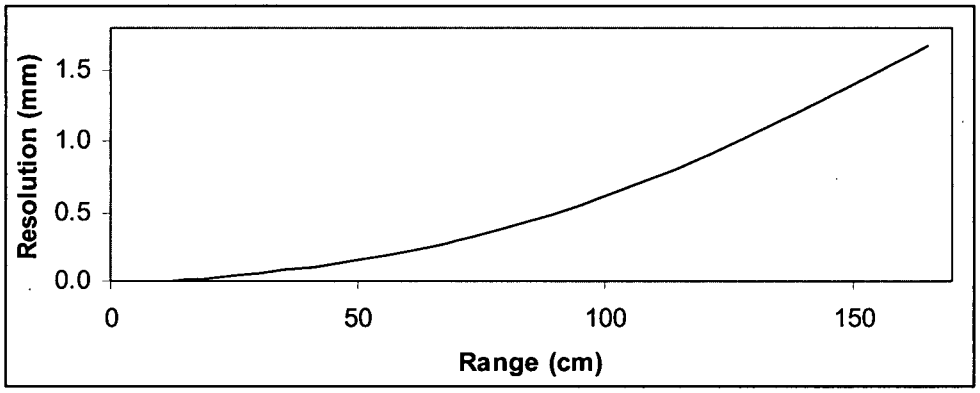



**Figure 1-3. Approximate relationship between range of LRS and resolution.**

Because this technology is very new to wood products processing, there is little information available on the operational performance of LRSs in sawmills. A recent study evaluated the accuracy of laser measurement devices with router-sawn blocks of wood from 15 different tree species (Sandak *et al.* 2003). The authors were able to accurately differentiate surface anomalies of 0.7mm (0.027 inch), except when wood was at its extreme values of density or colour. This could have important implications for some mills, where both light- and dark-coloured wood (e.g., from western hemlock (*Tsuga heterophylla* Sarg.) and western red cedar (*Thuja plicata* Donn ex D. Don)) are used to make lumber.

With rough-sawn wood, there is additional surface complexity that results in measurement error and other data anomalies. Before data are input into an SPC system, they must be filtered to eliminate machine vibration, reduce measurement error, and remove gross non-sawing defects. Machine vibration can be removed using an algorithm that was recently developed in the Department of Mechanical Engineering at The University of British Columbia (Gazzarri 2003); however, measurement error and non-sawing defects remain important issues.

Measurement error occurs in LRS data because of inadequate surface reflectance from the rough sawn wood (Wehr and Lohr 1999; Burman 2002). Removal of measurement error can be

accomplished via filtering methods, such as those found in digital terrain mapping or as suggested by Funck *et al.* (1992). This includes median filtering, Lee's sigma filter, and other types of spatial domain filtering.

Defects in the LRS data that are not the result of sawing, such as loose knots and wane, can appear like sawing defects, such as saw tear-out[6] or taper, if data are examined in isolation. To ensure that data are representative of the sawing process and not the quality of the log sawn, non-sawing defects need to be removed from the LRS data prior to performing SPC. The delineation of defects has been well-researched in conjunction with wood inspection systems for defect detection and automated grading. An array of technologies has been investigated, including ultrasound, x-ray radiation, infared, and visible light (Szymani 1985). For instance, a sheet-of-light profile imaging system can be included in a size control measurement apparatus by adding a digital x-y camera and laser line (Figure 1-4). A plane of light is projected at a 45 degree angle onto the board or cant surface using a laser line, and when viewed from an angle perpendicular to the board, the light reflects from the surface as a two-dimensional curve. The $(x,y)$ coordinates of this curve are captured via the digital camera.

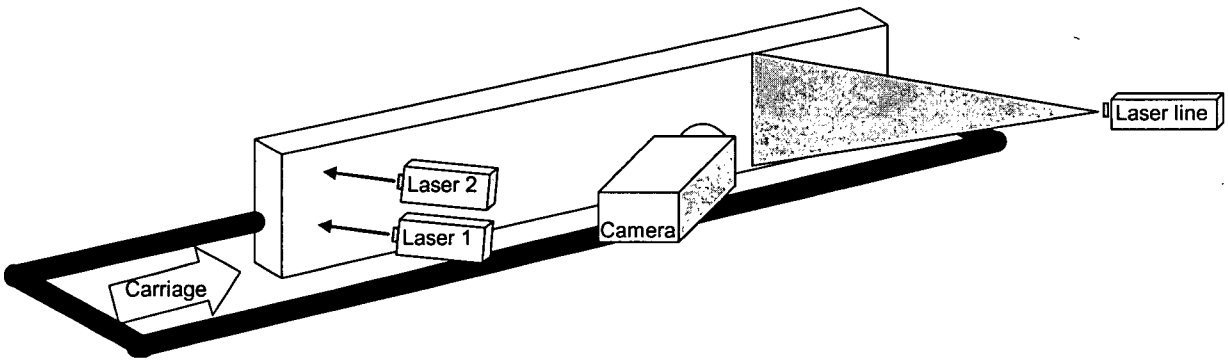



**Figure 1-4. Measurement apparatus with two LRSs and sheet-of-light profiling system (side view).**

[6] Saw tear-out is a condition in which the saw blade rips the grain on the surface of a workpiece. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

Sheet-of-light profiling systems have been investigated for use in wood products processing as part of complex machine grading systems (e.g., Kline *et al.* 2001; Lee *et al.* 2001). Other more expensive technologies, such as colour cameras (Butler *et al.* 1989; Butler *et al.* 2002) and magnetic resonance imaging (Coates *et al.* 1998) are also available.

Algorithms for feature extraction and classification of defects range from simple statistics and time series modeling (Koivo and Kim 1989) to neural networks and fuzzy logic (Conners *et al.* 1992). As part of an LRS-based SPC system, wane and other defects are only of concern when they occur in the same place as the laser measurements. Whereas grading systems must precisely delineate defect boundaries over an entire board area, a system for removing non-sawing defects from LRS data can be much simpler, relying on point estimation at the exact location of LRS measurements.

### 1.2.3.2 Methods for Real-time SPC

The assumptions underlying traditional SPC methods can be summarized as follows (Montgomery and Friedman 1989):

1. Data are obtained in periodic samples;
2. Observations between and within samples are independent;
3. Samples are grouped in rational subgroups, with sample size greater than one; and
4. The data follow a particular probability distribution.

When periodic samples are taken at irregular time intervals, these assumptions are roughly met. However, using LRSs, thousands of measurements on each and every board or cant are taken. If a single line of measurements is taken down the length of a board, a virtually continuous description of the wood surface can be obtained. Using four LRSs, data are also available from

both sides of each board or cant at two locations (e.g., 2.5 cm above the bottom of the board/cant and 2.5 cm from the top of the board/cant).

Since measurements are very close together, data from a single LRS are serially and auto-correlated; data from the two LRSs on a single side are likely also correlated. Furthermore, the lumber is essentially censused. As a consequence, assumptions 1, 2, and 3 are no longer valid. Although control charts are reasonably robust to moderate departures from some of the above assumptions, e.g. moderate non-normality, the assumption of independent data is critical to the use of inferential statistics and the proper function of a control chart (Montgomery and Friedman 1989; Wheeler 1995). In real-time systems, SPC techniques must be modified for the change in sampling scheme, the addition of a correlative structure, and the multivariate nature of the problem.

The first step in designing the SPC system using LRS data is deriving a statistical model. The mathematical description of the sawing process is complex because each sawn surface is machined by a separate process, each subject to several sources of variation. Present SPC systems are much simpler since only the thickness of sawn pieces of lumber is examined. Taking advantage of the LRS technology, thickness information, as well as surface profile information can be used for a more advanced system of SPC that examines each surface. A model based on the LRS sampling scheme will partition the variance of the LRS data to account for the complex sources of variation generated by each profile.

Once the model is identified, an SPC system needs to be designed for data from this model. That is, a system will be designed that explicitly accounts for all sources of variability and autocorrelation. Moreover, the SPC system needs to be designed to take advantage of the wealth of data captured. The increase in data represents an opportunity to monitor more than simple board thicknesses. Each and every board can be analyzed to look for specific defects and

diagnose specific sawing problems. For instance, machine positioning problems can be targeted by monitoring board thicknesses, whereas saw maintenance issues such as tensioning can be targeted by monitoring for wave patterns along each board surface. A group of control charts can be developed as part of a system for SPC, each with a specific purpose.

In order to make the system attractive to mill personnel, it is desirable that the SPC system uses Shewhart-type charts. These charts are easy to use and understand, and are more likely to be accepted by mill staff (Wheeler 1995). On the other hand, because of the form of the statistical model, traditional 3-sigma Shewhart-type charts, such as X-bar and S charts, are not appropriate for monitoring all process parameters and alternative control charting techniques need to be investigated.

Alternative Shewhart-type charts have been developed for non-normal data using quantile function values (Padgett and Spurrier 1990; Grimshaw and Alt 1997). In Levinson (1997), for example, a standard gamma distribution was fit to non-normal data, and the upper and lower quantiles of this distribution were used for control limits. In situations where a large amount of data are available, distribution-free methods can be utilized to develop empirical charts. Using bootstrapping methods, several researchers (Liu and Tang 1996; Willemain and Runger 1996; Jones and Woodall 1998) have developed and tested control limits for dependent and non-normal measurements. Under non-normal conditions, both methods have been found to out-perform traditional 3-sigma based Shewhart charts.

Other alternatives to Shewhart charts include cumulative sum (CUSUM) charts and exponentially weighted moving average (EWMA) charts; however, these charts are sensitive to non-normal and autocorrelated data (Montgomery 2001). Moreover, they are notoriously difficult to develop, maintain, and interpret, and numerous studies have shown that there is little, if any, improvement in using these charts over Shewhart charts (Wheeler 1995).

SPC methods have also been developed to explicitly account for autocorrelation in SPC data. Two main methods have been used: (1) adjusting control limits and estimates of the process variance to account for autocorrelation in the data, and (2) modeling the data with an appropriate time series model and applying control charts to residuals. Most researchers recommend the second method (e.g., Lu and Reynolds 1999a; Lu and Reynolds 1999b; Lu and Reynolds 2001; Montgomery 2001), although fitting, maintaining, and interpreting an appropriate model can be cumbersome and difficult (Faltin *et al.* 1997).

Several researchers have applied these techniques to account for autocorrelation in paper and fiberboard SPC systems (Cook 1992; Young and Winistorfer 2001; Noffsinger and Anderson 2002). In Young and Winistofer (2001), for example, the moisture content of medium-density fiberboard (MDF) samples taken at one-hour intervals were highly autocorrelated. In automated lumber processing, the biggest source of autocorrelation is between measurements within each board; however, it is not feasible to monitor individual measurements within board given the large amount of data collected by the LRSs. Autocorrelation is only a concern if individual observations within a single piece of lumber are to be monitored

SPC methods have also been explicitly developed for situations where multiple quality characteristics are monitored simultaneously. The Hotelling's $T^2$ control chart is a multivariate Shewhart type chart that tests for a general shift in the mean vector of a single observation. However, multivariate charts have several disadvantages. The data must follow a multivariate normal distribution, the effects of shifts in the mean, variance, or covariance structure are confounded, and these methods are sensitive to a shift in only one variable (Hawkins 1991; Mastrangelo *et al.* 2001; Montgomery 2001). Moreover, out of control signals still must be investigated via univariate control charts in order to determine the cause of the signal (Does *et al.* 1999). In the area of wood composites, Young *et al* (1999) investigated multivariate methods for

application to independent streams of several density and weight measurements using the Hotelling's $T^2$ chart. They concluded that multivariate control charting procedures were better at detecting special cause variation than univariate charts, but recommended the simultaneous use of both univariate and multivariate charts for ease of interpretation.

## 1.3 Challenges for Real-time Systems

The goal of this research is to develop a mathematical system that is capable of modeling lumber size control and will reliably monitor out-of-control conditions in real time. Specifically, the wood surface texture on two faces will be scanned and analyzed so that machines producing defective material can be identified, allowing prompt repairs to be made.

The primary research problems for real-time SPC systems can be divided into two broad areas. The first area is the application of LRS data to the lumber manufacturing environment. Raw data coming from the laser range sensors must first be filtered before applying SPC techniques. Also, real lumber data consist of boards that have anomalous surfaces due to circumstances outside of sawing. These anomalous surfaces, such as wane, must be identified prior to applying SPC techniques. The second broad area is the development of a real-time SPC system based on LRS data. A new statistical model must be derived because the mathematical principles on which the current automated lumber manufacturing SPC methods are based do not adequately describe the real-time process. Further, using this new model, an SPC system must be designed that takes advantage of the wealth of data being acquired by the laser range sensors.

## 1.4 Statement of Research Objectives

The overall research objective for this thesis is to develop a real-time SPC system that (1) continuously and accurately monitors the quality of wood products being manufactured; and (2)

provides reliable information to improve the performance of the process in real-time. This overall objective can be broken down into four sub-objectives:

1. To evaluate filtering algorithms for removing measurement error and non-sawing defects from automated SPC data taken from rough green lumber (Chapter 2);
2. To develop a multi-sensor system for identifying sawing and non-sawing defects in automated SPC data taken from rough green lumber (Chapter 3);
3. To describe a statistical model for online LRS profile data taken on sawn lumber, which will enable an SPC system to be created (Chapter 4); and
4. To develop a SPC system based on a statistical model of the sequence of real-time LRS measurements (Chapter 5).

Chapter 6 presents overall conclusions and directions for further research.

## *1.5 Literature Cited*


Brown, T. D. 1979. Determining lumber target sizes and monitoring sawing accuracy. Forest Products Journal 29(4): 48-54.

Brown, T. D. 1982. Quality Control in Lumber Manufacturing. San Francisco, CA, Miller Freeman Publications.

Burman, H. 2002. Laser strip adjustment for data calibration and verification. Presented at Symposium 2002: ISPRS Commission III. September 9-13, Graz, Austria.

Butler, D. A., C. C. Brunner and J. W. Funck. 1989. A dual-threshold image sweep-and-mark algorithm for defect detection in veneer. Forest Products Journal 39(5): 25-28.

Butler, D. A., C. C. Brunner and J. W. Funck. 2002. Wood surface feature classification via extended-color imagery. Forest Products Journal 52(6): 80-84.

Coates, E. R., S. J. Chang and T. W. Liao. 1998. A quick defect detection algorithm for magnetic resonance images of hardwood logs. Forest Products Journal 48(10): 68-74.

Conners, R. W., T. H. Cho, C. T. Ng, T. H. Dryer, P. A. Araman and R. L. Brisbin. 1992. A machine vision system for automatically grading hardwood lumber. Industrial Metrology 2(3/4): 317-342.

Cook, D. 1992. Statistical process control for continuous forest producers manufacturing operations. Forest Products Journal 42(7/8): 47-53.

Does, R. J. M. M., K. C. B. Roes and A. Trip. 1999. Handling multivariate problems with univariate control charts. Journal of Chemometrics 13: 353-369.

Faltin, F. W., C. M. Mastrangelo, G. C. Runger and T. P. Ryan. 1997. Considerations in the monitoring of autocorrelated and independent data. Journal of Quality Technology 29(2): 131-133.

Funck, J. W., F. A. Gabner and F. Aslam. 1992. Using image analysis for lumber size control measurements. Industrial Metrology 2: 237-252.

Gaylor, D. W. and F. N. Hopper. 1969. Estimating the degrees of freedom for linear combinations of mean squares by Satterthwaite's formula. Technometrics 11(4): 691-706.

Gazzarri, J. 2003. Sequential measurements method for moving surfaces profiling. MSc Thesis. Department of Mechanical Engineering, The University of British Columbia, Vancouver, Canada.

Grimshaw, S. D. and F. B. Alt. 1997. Control charts for quantile function values. Journal of Quality Technology 29(1): 1-7.

Harter, H. L. 1960. Tables of range and studentized range. The Annals of Mathematical Statistics 31(4): 1122-1147.

Hawkins, D. M. 1991. Multivariate quality control based on regression adjusted variables. Technometrics 33(1): 61-75.

Hermary Opto Electronics Inc. 2002. What is scanner technology [online]. Available from http://www.hermaryopto.com/scantech.html [cited January 14, 2004].

Jones, L. A. and W. H. Woodall. 1998. The performance of bootstrap control charts. Journal of Quality Technology 30: 362-375.

Kline, D. E., C. Surak and P. A. Araman. 2001. Evaluation of a multi-sensor machine vision system for automated hardwood lumber grading. *In* Proceedings of 4th International Conference on Image Processing and Scanning of Wood, August 21-23, 2000, Mountain Lake, VA. pp. 75-87.

Koivo, A. J. and C. W. Kim. 1989. Automatic classification of surface defects on red oak boards. Forest Products Journal 39(9): 22-30.

Kraus, K. and N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53: 193-203.

Lee, S. M., A. L. Abbott and D. L. Schmoldt. 2001. Wane detection of rough lumber using surface approximation. *In* Proceedings of 4th International Conference on Image Processing and Scanning of Wood, August 21-23, 2000, Mountain Lake, VA. pp. 115-126.

Levinson, W. A. 1997. Watch out for non-normal distributions of impurities. Chemical Engineering Progress 93(5): 70-76.

Lister, P. 1997. Understanding the relationship between circular saw kerf width, sawing variation, and feed speed. *In* Proceedings of 5th International Conference on Sawing Technology, Seattle, WA. Wood Machining Institute, Berkeley, CA. pp. 17-22.

Liu, R. Y. and J. Tang. 1996. Control charts for dependent and independent measurements based on bootstrap methods. Journal of the American Statistical Association 91: 1694-1700.

Lu, C. and M. R. Reynolds. 1999a. Control charts for monitoring the mean and variance of autocorrelated processes. Journal of Quality Technology 31(3): 259-274.

Lu, C. and M. R. Reynolds. 1999b. EWMA control charts for monitoring the mean of autocorrelated processes. Journal of Quality Technology 31(2): 166-188.

Lu, C. and M. R. Reynolds. 2001. Cusum charts for monitoring an autocorrelated process. Journal of Quality Technology 33(3): 316-334.

Maness, T. C., R. A. Kozak and C. L. Staudhammer. 2003. Applying real-time statistical process control to manufacturing processes exhibiting between and within part size variability in the wood products industry. Quality Engineering 16(1): 113-125.

Maness, T. C., R. A. Kozak and C. L. Staudhammer. 2004. Reliability testing of statistical process control procedures for manufacturing with multiple sources of variation. Wood and Fiber Science 36(3): 443-458.

Maness, T. C. and Y. Lin. 1995. The influence of sawkerf and target size reductions on sawmill revenue and volume recovery. Forest Products Journal 45(11/12): 43-50.

Maness, T. C., C. L. Staudhammer and R. A. Kozak. 2002. Statistical considerations for real-time size control systems in wood products manufacturing. Wood and Fiber Science 34(3): 476-484.

Mastrangelo, C. M., J. M. Porter and R. V. Baxley. 2001. Multivariate process monitoring for nylon fiber production. *In* Frontiers in Statistical Quality Control. T. Wilrich (Ed.). Springer-Verlag: pp. 228-246.

Montgomery, D. C. 2001. Introduction to Statistical Quality Control. 4th ed. New York, John Wiley & Sons.

Montgomery, D. C. and D. J. Friedman. 1989. Statistical process control in a computer-integrated manufacturing environment. *In* Statistical process control in automated manufacturing. N. F. Hubele (Ed.). M. Dekker, New York: pp. 67-87.

Nelson, L. S. 1999. Notes on the Shewhart control chart. Journal of Quality Technology 31(1): 124-126.

Noffsinger, J. R. and R. B. Anderson. 2002. Effect of autocorrelation on composite panel production monitoring and control: a comparison of SPC techniques. Forest Products Journal 52(3): 60-67.

Padgett, W. J. and J. D. Spurrier. 1990. Shewhart-type charts for percentiles of strength distributions. Journal of Quality Technology 22(4): 283-290.

Rasmussen, H. K., R. A. Kozak and T. C. Maness. 2004. An analysis of machine caused lumber shape defects in British Columbia sawmills. Forest Products Journal 54(6): 47-56.

Ryan, T. P. 1989. Statistical methods for quality improvement. New York, John Wiley & Sons.

Sandak, J., C. Tanaka and T. Ohtani. 2003. Evaluation of surface smoothness by a laser displacement sensor 1: effect of wood species. Journal of Wood Science 49: 305-311.

Schajer, G. S. 1989. The importance of circular saw tensioning. Forest Industries/World Wood Technical Report (Sawing Technology): T14-T16.

Shewhart, W. A. 1931. Economic control of quality of manufactured product. New York, D. Van Nostrand Company Inc.

Szymani, R. 1985. An overview of scanning technology in sawmilling. *In* Proceedings of 1st International Conference on Scanning Technology in Sawmilling, San Francisco, CA. Miller Freedman Publications. pp. 1-33.

Wang, S. J. 1983. An analytic approach to estimating the increase in lumber recovery due to reduced target sizes and saw kerfs. Forest Products Journal 33(11/12): 29-32.

Wang, S. J. 1984. A new approach to calculating target sizes. Forest Products Journal 34(9): 53-60.

Warren, W. G. 1973. How to calculate target thickness for green lumber. Department of the Environment, Canadian Forestry Service, Western Forest Products Laboratory, Vancouver, B.C.

Wehr, A. and U. Lohr. 1999. Airborne laser scanning - an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing 54: 68-82.

Wheeler, D. J. 1995. Advanced Topics in Statistical Process Control. Knoxville, TN, SPC Press.

Whitehead, J. C. 1978. Procedures for developing a lumber-size control system. Canada Department of the Environment, Forestry Directorate, Western Forest Products Laboratory, Vancouver, B.C.

Willemain, T. R. and G. C. Runger. 1996. Designing control charts using an empiricial reference distribution. Journal of Quality Technology 28: 31-38.

Young, T. M. and P. M. Winistorfer. 2001. The effects of autocorrelation on real-time statistical process control with solutions for forest products manufacturers. Forest Products Journal 51(11/12): 70-77.

Young, T. M., P. M. Winistorfer and S. Wang. 1999. Multivariate control charts of MDF and OSB vertical density profile attributes. Forest Products Journal 49(5): 79-86.

# Chapter 2 Filtering Methods for Laser Generated Data in Real-time Statistical Process Control for Lumber Manufacturing

## 2.1 Introduction

Sawmills aim to produce quality lumber of a consistent size with minimum variation. As raw logs are a significant portion of sawmill costs, careful monitoring of lumber sizes with a statistical process control (SPC) system is extremely important. In the green stage, lumber is sawn to a target size that allows for variability in the drying and sawing process (Maness 1996). Reducing variability in lumber sizes enables sawmills to reduce target size, which results in higher product recovery, and hence, higher profitability (Wang 1983; Maness and Lin 1995; Lister 1997).

For more than two decades, non-contact laser sensors have been employed in automated lumber manufacturing, for example, to signal the presence of lumber in a particular machine, or to determine the geometry of a flitch for edging and trimming. More recently, laser range sensors (LRS) have been introduced to measure tactile roughness (Sandak *et al.* 2004) and to measure cant and board thicknesses for SPC. Typically, LRSs use optical triangulation to measure distance to an object. As shown in Figure 2-1, the LRS projects a laser beam from a sensor onto an object. The laser spot is reflected from the object on an electronic camera, and the distance to the object is computed from the position of the reflected image on the camera.

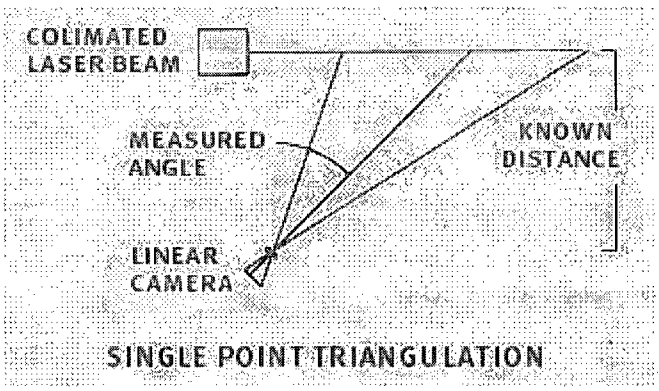



**Figure 2-1. Optical triangulation in LRSs (source: http://www.hermaryopto.com/scantech.html).**

In an automated SPC system for lumber size control, an LRS is coupled with a position detection system, such as an encoder, producing a data acquisition system referred to as a "laser scanner". In laser scanning, the ranging beam is deflected in a specific pattern so that an object surface is sampled with a high point density (Wehr and Lohr 1999). This point density can be up to 1000 measurements per second and the LRS can be calibrated to make extremely accurate measurements in the range of 10 to 75 cm (4 to 30 inches). This makes LRSs very attractive in automated SPC systems.

The purpose of SPC in lumber manufacturing is to control sawing performance by monitoring lumber sizes. In real-time automated SPC systems, pairs of LRSs measure lumber or cant thickness directly after a particular sawing machine. For instance, LRSs have been installed in lumber mills to measure cants as they leave the canterline. In addition, some mills measure boards by singulating them as they leave the canterline, leave the gangsaw, or enter the edger. The optimal time to make measurements for SPC is when lumber is in its "rough green" stage (not planed nor dried), as decisions made early in the production chain have the greatest financial impact (Maness 1993; Abbott *et al.* 2001).

Real-time LRS data can improve the effectiveness of SPC programs because they can provide early and immediate feedback; however, it is subject to measurement errors. Because of the industrial atmosphere, data contain machine-caused vibrations. Sawn lumber is not a perfectly flat surface, and so LRS data invariably contain erroneous measurements due to inadequate reflectance of the laser beam. In a typical sawmill, anomalous measurements can also occur because of non-sawing defects in the lumber (e.g., wane[7] and holes), or can be a result of the sawing environment (e.g., sawdust). In a related research project, researchers at The University of British Columbia, Mechanical Engineering Department developed an algorithm to identify

[7] Wane is the natural curvature of the edge of a board sawn from a log. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

and remove machinery vibration from the LRS data (Gazzarri 2003). Measurement errors due to non-sawing defects and inadequate reflectance still need to be addressed. In order to accurately represent sawing performance in the data, erroneous measurements need to be filtered out of the LRS data before SPC algorithms are applied.

## 2.2 Research Objective

Data filtering is an important first step in obtaining accurate data from LRSs for subsequent SPC applications. The objective of this chapter is to develop an algorithm for removing measurement error and non-sawing defects, such as wane, from laser scanner data taken from rough green lumber.

## 2.3 Review of Literature

There are two main purposes for filtering laser data in sawmilling applications: (1) removing erroneous data that result from inaccurate measurement, and (2) removing anomalous data that result from the data collection environment. Data filtered out for either purpose are often referred to as "noise". Laser-specific filtering algorithms for (1) largely come from the area of digital terrain mapping (DTM), where laser range data are used to create maps with methods from digital image processing. Optimally, algorithms for (2) would originate from the automated lumber manufacturing area; however, filtering methods for lumber data have appeared in the literature only as applied to image processing.

### 2.3.1 Sources of Errors in Laser Scanning

Wehr and Lohr (1999) and Baltasavias (1999) gave overviews concerning laser scanners in DTM, and outline factors affecting accuracy of laser measurements, including the ranging signal and the signal-to-noise ratio (S/N) (Wehr and Lohr 1999). The ranging signal is affected by the reflectivity of a target over a given wavelength. The S/N is dependent on several underlying factors, including the measurement rate, the power of the received signal, and the amount of

background radiation. Burman (2002) gave three sources of errors in laser scanning: inadequate reflection of the laser beam off the given object, erroneous laser length, and erroneous laser orientation.

These details are particularly important in wood products applications. While errors in the laser length and orientation may be corrected with careful calibration, inadequate reflection and background radiation are more complicated problems. According to Baltasavias (1999), the minimum detectable object is more dependent on its reflectivity than its size, and for aircraft-based laser applications, the worst results are obtained in bright sunlight. Background radiation due to sunlight may not be easily controlled in some industrial situations, and the reflectivity of wood varies appreciably with moisture content and species (Sandak *et al.* 2003).

Inadequate reflection is also related to the physical attributes of the lumber surface. Wehr and Lohr (1999) noted that clean dry pine dimension lumber has a reflectivity of 94%, whereas clean rough sawn pallets have a reflectivity of only 25%. Reflectivity of rough lumber is dependent on many factors, such as colour, species, surface irregularities, and edges (e.g., from knots or saw tear-out[8]). Regardless of the resolution of a laser measurement device, the laser "footprint" is not a point, but an area (Burman 2002). Therefore, a laser beam hitting a rough knot or tear-out can return an erroneous range. Compared with highly accurate laboratory stylus profilometry, Sandak *et al.* (2003) found that laser based methods tended to round edges and "flatten" the natural variation in the surface of the lumber (e.g., fibers). When using laser profilometers with smaller spot sizes, researchers found that this effect was greatly lessened (D.C. Wong, personal communication 2004[9]).

---

[8] Saw tear-out is a condition in which the saw blade rips the grain on the surface of a workpiece. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

[9] Wood Machining Scientist, Forintek Canada Corporation, 2665 East Mall, Vancouver, B.C. Canada V6T 1W5

Because of the difficulties in obtaining accurate images of rough wood, research has been focused on planed lumber (Abbott *et al.* 2001). One of these difficulties is that there is a large variation in what is considered a normal, acceptable image. Dark knots can be mistaken for holes; edges can be broken, sharp or gradual, or occluded by sawdust or wood slivers (Funck *et al.* 1992). Scanning of un-planed wood surfaces presents additional challenges, because data are often contaminated with residual bark, debris, and dust. In addition, the surface reflectance of wood varies, because of the "tracheid effect"; this differential reflectance of laser light in response to grain angle and wood density (Soest and Matthews 1985) can cause erroneous measurements on the side of the most reflectance (Lee *et al.* 2001). Despite these difficulties, Funck *et al.* (1992) reported that problems with background radiation, surface reflection, and surface roughness can usually be minimized by data filtering.

### 2.3.2 Filtering Methods

Data filtering methods are used to enhance contrast, smooth images, and minimize spurious data resulting from sampling or transmission functions. If the signal degradation process is known *a priori*, filters can be used to minimize effects of "noisy" data using a mathematical description of the degradation process via methods such as constrained least-squares, the Wiener filter, or the Kalman filter; however, these methods are computationally intensive and may produce blurred images that conceal subtle details (Lee 1983). Moreover, in a sawmill making lumber of varying species and quality, the signal degradation cannot be known.

In forestry and mapping applications, filtering is often combined with interpolation because of incomplete penetration of the laser signal (e.g., Kraus and Pfeifer 1998). Unwanted measurements, such as those from trees, houses, and power lines, are removed with less computationally expensive methods derived from image processing. Filters are applied with local operators, which involve only a small number of pixels in computations. Acceptance-based

local filtering is also common. For instance, Vosselman and Maas (2001) filtered laser data based on neighborhood values of the median and slope, removing points where the absolute value of the slope exceeded a certain value, or the median exceeded the neighborhood median by a certain value. Methods are often based on arbitrary limits and, in some situations, removal of erroneous points by manual editing is not uncommon (Vosselman and Maas 2001).

Some methods used for DTM, such as those based on slope, are not applicable for wood surfaces, as changes due to sawing defects may be abrupt. On the other hand, some of these DTM methods for filtering laser data may be directly transferable to wood products processing applications. However, there are no "standard methods", which has led to a proliferation of algorithms based on many concepts (Axelsson 1999; Wehr and Lohr 1999; Vosselman and Maas 2001). Further, because most algorithms are proprietary, they are difficult to investigate.

Filtering literature also exists in the medical imaging area, where complicated algorithms are used to remove noise from many different kinds of imaging systems. However, unlike wood processing, speed, cost, and harsh environments are not critical factors in medical imaging. For real-time SPC in automated lumber manufacturing, fast and simple methods for data acquisition and filtering are necessary (Kline *et al.* 2001).

Spatial domain filters are among the more simple filtering methods. Low-pass filters eliminate high-frequency values in the frequency domain, e.g., neighborhood averaging. On the other hand, high-pass filters, e.g., median filtering, eliminate low-frequency components. Whereas low-pass filters tend to blur edges and other sharp details (Gonzalez and Woods 1992), high-pass filters give sharper edges. If a noise source is known to exist at a certain value, band-pass filters are used to eliminate selected frequency regions (Gonzalez and Woods 1992). Frequency domain filtering is common in electronics applications, and is performed using the Fourier

transform. Spatial domain filtering is preferred in image processing, because Fourier transforms can be computationally inefficient and introduce low-frequency interference (Funck *et al.* 1992).

For edge detection in lumber grading, Funck *et al.* (1992) investigated average filters, median filters, and Lee's sigma filter (Lee 1983). The median filter is a simple method whereby each observation is replaced by the median of a moving window centred around it. This method has also been used by other wood products researchers in conjunction with image processing (e.g., for classification of compression wood with digital images (Coates *et al.* 1998)).

Lee's sigma filter was built on the assumption that image noise has a Gaussian distribution. First, a "sigma-based probability limit", $S_\sigma$, from an inverse normal distribution was chosen for use over the whole dataset. For each *i*th point in the dataset, Lee calculated the moving standard deviation (MSD) over a small window (neighborhood), and re-calculated the centre element of the neighborhood with respect to the sigma-based probability limits. For example, choosing $S_\sigma$=2 gave probability limits of $\pm 2 \times$ MSD, which approximated a 95% confidence interval around the *i*th point. If all the elements in the neighborhood of the *i*th point fell within the sigma-based probability limits, the centre element of the neighborhood was replaced with the average. If any elements in the neighborhood fell outside of the sigma-based probability limits, they were judged to be "from a statistically different population" (Lee 1983), and the centre element was replaced using an average calculated without those elements.

According to Lee (1983), advantages of this method included efficiency in computation, effective noise smoothing, and the ability to preserve subtle details and retain edges. Lee's algorithm also included a procedure to remove high-contrast spot noise; if there were less than some minimum number of points available for the re-calculation of the centre point, the centre point was replaced with the value of the neighboring centre point.

Lee *et al.* (2003) investigated a variety of smoothing methods for their work in wane detection for rough lumber, including neighborhood threshold limits. Simple threshold limits were constructed with ± 3 times the standard error of the mean, approximating a 99.7% confidence interval under the assumption of normally distributed data. These limits, as well as limits based on a moving standard deviation did not work well in eliminating noise from residual bark and other debris (e.g., sawdust). Their chosen method was a 3-dimensional approach, using the curvature and orientation of the surface.

Many filtering algorithms have been used in connection to LRS data for non-wood products, or with wood products data obtained with other image processing applications; no algorithm was found that specifically targeted LRS data taken on wood products. Choosing a filtering algorithm based on the existing literature is not straightforward, as one algorithm can be effective for some types of images or signals, but ineffective for others (Lee 1983).

## 2.4 Methods and Materials

Three filtering methods were chosen for evaluation: two are existing methods and one is a new method. Because the LRS measurements are made in a fast-moving and harsh environment, simple and computationally efficient methods were desirable. Moreover, simplicity is appropriate for lumber data, as in comparison to Lee's (1983) radar data, lumber data are "not very noisy" (Funck *et al.* 1992). Also, methods previously used in the wood products literature were desirable in order to make comparisons. Thus, a simple median filter and Lee's sigma method were evaluated. Following the ideas used in Lee *et al.* (2003), a new rule-based method was developed, which was tailored to the observed properties of lumber scan data. This new method was based on neighborhood thresholds and used moving standard deviations and other local statistics. The three methods chosen for evaluation were therefore:

1. Method 1: the median filter;

2. Method 2: Lee's sigma filter; and

3. Method 3: the "MSD filter".

### 2.4.1 Materials

One hundred and ten pieces of rough green western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) lumber measuring 51.5 mm x 135 mm ($2\frac{1}{32}$ inches× $5\frac{5}{16}$ inches) were obtained from Weyerhaeuser's New Westminster sawmill (British Columbia, Canada). The lumber was processed by several different sawing machines: approximately half the lumber was bandsawn, one-quarter was circular-sawn, and one-quarter was chipped (i.e., cut by a chipper-head). This produced sawn surfaces of varying quality and numerous sawing and non-sawing defects. Lumber samples were judgmentally selected for this study so that they represented a range of sawing quality. Six samples (Boards 001 – 006) were selected to replicate defect-free lumber, as well as lumber with non-sawing defects (e.g., wane) and sawing defects (e.g., tear-out).

After sawing, the lumber was immediately taken to the Q-Lab of the Department of Wood Science in the Forest Sciences Centre at The University of British Columbia, Vancouver, Canada. Before scanning, the lumber was cut to 2.44-metre (8-foot) lengths. The laser scanning apparatus used a high quality motor-driven carriage to move lumber at constant speed through a scan zone. The scan zone could be configured to use up to four Hermary LRS-50 point laser range sensors to obtain simultaneous measurements from multiple sides of the lumber; however, only data from one LRS were used for evaluating filtering methods (Figure 2-2).

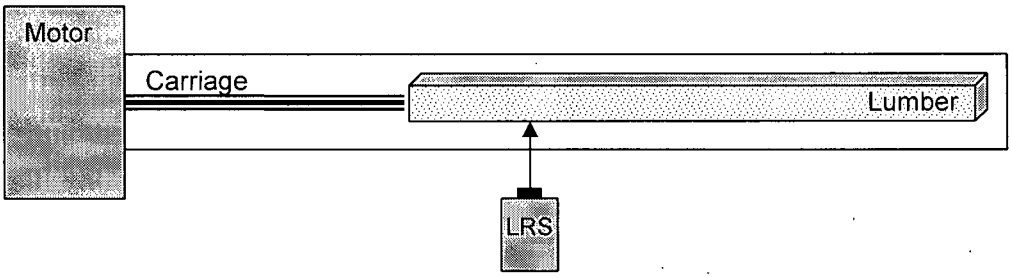



**Figure 2-2. Set up of measurement apparatus in Q-lab, as viewed from above.**

The maximum measurement rate in a typical sawmill was simulated by setting the LRS to take 500 measurements per second (half the maximum measurement speed) and the carriage speed to 500 rpm. The LRS took approximately 2800 measurements over the 2.4-metre boards, giving a sampling frequency of approximately 12 measurements per centimeter (30 per inch). In order to eliminate the possibility of the lumber moving in the apparatus during the scan, the lumber was secured by three clamps at 15, 122, and 129 cm (6 inches, 4 feet, and 7 ½ feet) along the board's length[10], two at each end and one at the midpoint.

The data captured by the LRS were distances from the laser to the wood surface; $l_{km}$ is the $m$th measurement from the $k$th sample board. It is known that the precision of the LRS is dependent on $l_{km}$. This distance-dependent precision is referred to by the LRS manufacturer as its "resolution". A maximum resolution of 25 $\mu$m (0.001 inch) is obtained at a distance of 14 ½ cm (5 ¾ inches). In the Q-lab, the average distance, $\bar{l}_{..}$, was approximately 40 cm (15 ¾ inches), giving a precision of 0.08 mm (0.003 inch). Since precision is effectively constant for the measurement apparatus over all scans, this resolution represented a uniform decrease in the precision of the LRS measurements. As defined in this research, resolution was separate from the "measurement error" previously described in this chapter and therefore, adjustment for resolution was not considered in the filtering process.

The data were collected from March 27 to April 4, 2003. Although the measurement precision of the laser was assumed constant, the accuracy was not. Even a slight movement in the laser mounting frame could cause laser measurements to deviate. For example, if frame movement changed the angle of the laser by 1°, LRS measurements for a target 50 mm (19 ¾ inches) away would change by 75 $\mu$m (0.003 inch). To ensure consistent accuracy over the collection period, a calibration block made of molded Teflon® and machined to 25 $\mu$m (0.001 inch) accuracy was

[10] Data collected from the clamped areas were excluded from the analysis.

measured three times per day. The deviation between the laser readings and the known dimensions of the block were used to calibrate measurements. Inconsistencies between calibrations were noted and investigated. Although the reflectance properties of Teflon® and wood are known to be different, it was not possible to machine a block of wood to this accuracy. It was assumed that inaccuracy introduced by material properties was consistent over all data and would therefore not affect the evaluation of the filtering methods.

After scanning, information about the sawn surface was recorded. Surface anomalies located in the scanning zone were mapped in detail. These data included:

1. Saw type for each side (Bandsaw, Circular Saw, or Chipper-head);
2. The direction of sawing;
3. Obvious saw mark patterns;
4. Discolourations, such as smooth knots, pitch pockets, and stains;
5. Sawing defects, such as tear-out, skip, or step; and
6. Non-sawing defects, such as jagged knots, holes, and wane.

### 2.4.2 Algorithms

The six boards were filtered using the three proposed methods. Each of these methods used the same initial two steps:

1. Points from the scan that were obviously out of range were removed:

   If $\left| l_{km} - \bar{l}_{k\cdot} \right| \geq 2$ cm, then $l_{km}$ was set to missing;

   where: $\bar{l}_{k\cdot}$ is the average measurement from the $k$th board; and

2. Rough board edges created by the chop saw were eliminated by removing the first and last 30 observations:

If $m \leq 30$ or $m \leq n_k - 30$, then $l_{km}$ was set to missing;

where: $n_k$ is the total number of measurements taken on the $k$th board.

### 2.4.2.1 Method 1

Method 1 was a median method using a window of $w_1$ points. After the initial two steps, the steps for filtering under Method 1 were as follows for each board:

1. The $w_1$-point moving median by measurement and laser was calculated; and
2. Each centre point was replaced with its median.

Variations on this method resulted from changing:

1. The number of points in the moving median calculation ($w_1$ =3, 5); and/or
2. The number of passes through the median filter (1 or 2).

### 2.4.2.2 Method 2

Method 2 used Lee's sigma filtering algorithm with a window size of $w_2$. Using Lee's (1983) recommendation, the sigma-based probability limit, $S_\sigma$, was set at 2; the minimum number of points, *MIN*, was set at 3. After the initial two steps, the steps for filtering under Method 2 were as follows for each board:

1. The $w_2$-point moving average and moving standard deviation for the $m$th point on the $k$th board ($\tilde{l}_{km}$ and $\tilde{\sigma}_{km}$, respectively) were calculated. For example, with $w_2$=5:

$$\tilde{l}_{km} = \frac{\sum_{t=m-2}^{m+2} l_{kt}}{5} \qquad [2\text{-}1]$$

$$\tilde{\sigma}_{km} = \sqrt{\frac{\sum_{t=m-2}^{m+2} (l_{kt} - \tilde{l}_{km})^2}{5-1}} \qquad [2\text{-}2]$$

2. Sigma-probability limits were calculated:

$$\text{Lower limit} = \tilde{l}_{km} - S_{\sigma}\tilde{\sigma}_{km} \qquad [2\text{-}3]$$

$$\text{Upper limit} = \tilde{l}_{km} + S_{\sigma}\tilde{\sigma}_{km} \qquad [2\text{-}4]$$

3. The number of points in the $w_2$-point window that fell within the sigma-probability limits (*min*) was calculated; and

4. The data were filtered with the following rule: if $min \geq MIN$, the *m*th point was replaced with the average of those points in the window that fell in the sigma-probability limits; otherwise, the *m*th point was replaced with the previous ($m$-1$th$) point.

Variations on this method resulted from changing:

1. The number of points in the moving window calculation ($w_2$ = 7, 9); and/or

2. The number of passes through the filter (1 or 2).

### 2.4.2.3 Method 3

While the Methods 1 and 2 are existing methods designed to remove random noise, Method 3 is a new method designed to target specific non-sawing defects for removal from the laser data. Following techniques used in DTM filtering (e.g., Vosselman and Maas 2001), defects were described in terms of local statistics. For instance, random noise from splinters can cause one or two observations to be substantially higher than expected. On the other hand, wane can be indicated by an increase or decrease in the short-term variation of the observations. Holes can be characterized by large deviations from the average value of the observations.

Method 3 was based on the idea of having smooth local statistics and was developed with a series of steps, each targeting a certain type of defect. The last step was a smoothing step, which targeted areas that were only partially filtered in the previous steps. Each step in the algorithm

represents one loop through the dataset. After the initial two steps, the steps for filtering under this method were as follows for each board:

1. For the $m$th measurement from the $k$th board, the raw $w_3$-point moving average and moving standard deviation were calculated using [2-1] and [2-2].

2. Using half the target thickness value (T=25.75 mm, or $1\frac{1}{64}$ inch) and a maximum allowable deviation ($\Delta = \pm$ 6.4 mm or $\pm$ ¼ inch), the data were filtered for sporadic anomalies (such as splinters) with the following rule: if a measurement was outside T $\pm$ $\Delta$, then the measurement was removed, i.e., if $|T - l_{km}| > \Delta$, then $l_{km}$ was set to missing;

3. The MSDs were re-calculated, using the filtered points from step two: $\tilde{\sigma}'_{km}$;

4. The median MSD of each board using the filtered points from Step 3 was calculated: $\overline{\tilde{\sigma}'_{k\cdot}}$;

5. The data were filtered where there was a steep increase or decrease in the short-term variation (such as would be caused by wane) with the following rule: if the raw MSD centred at the measurement was more than $S_\sigma$ times the median filtered MSD for the board, then the measurement was removed, i.e., if $\tilde{\sigma}_{km} > S_\sigma \times \overline{\tilde{\sigma}'_{k\cdot}}$ then $l_{km}$ was set to missing;

6. The average measurement by board, $\bar{l}_{k\cdot}$, was calculated;

7. The "hole depth" for each measurement by board was calculated as:

$$h_{km} = l_{km} - \bar{l}_{k\cdot} \qquad [2\text{-}5]$$

8. The standard deviation for the board, $\hat{\sigma}_{k\cdot}$, was calculated;

9. The data were filtered when the measurements deviated substantially from the average (such as would be caused by holes) by:

   a) If $h_{km} > 2\,\hat{\sigma}_{k\cdot}$, then the measurement was suspected to be part of a hole; and

   b) If at least four of five consecutive points were suspected to be part of a hole, then the five consecutive points were set to missing;

10. The smoothing rule was: if at least 3 of 5 consecutive points were missing, then all 5 points were set to missing.

Variations on this method resulted from changing:

1. the number of points used in the MSD's in Steps 1 and 3 ($w_3$=3, 5); and/or

2. the number of standard deviations used in hole cleaning (Step 7, $S_\sigma$= 2, 3).

### 2.4.3 Evaluation

Since there is no standard methodology for evaluating data filtering algorithms, the three filtering algorithms were evaluated using a combination of subjective measures and descriptive statistics. For this application, an ideal filtering method was defined as one that eliminated anomalous data and non-sawing defects while preserving the basic structure and pattern of the data.

Parameters (e.g., $w_1$, $w_2$, and $w_3$) for the three methods were selected in a purely subjective manner in comparison to this ideal. Scans of defect-free (clear) wood were graphed and examined with respect to how the filtering methods changed the data. Parameters were chosen from a subjective visual perspective. Using the selected parameters, the three methods were then evaluated using clear lumber and lumber that contained defects. First, graphs of filtered scans were visually compared to data descriptions. Next, descriptive statistics, such as the mean, range, and standard deviation, were examined to assess how the methods altered the data.

The semivariogram is a measure of the dissimilarity between spatially separate measurements as the distance between them, $t$, increases. The semivariogram of the measurements from the $k$th board is given by:

$$\gamma_k(t) = \tfrac{1}{2}\,\mathrm{var}[l_{k,m} - l_{k,m+t}] \qquad [2\text{-}6]$$

The semivariogram has been used in remote sensing to objectively compare pixel data (Curran and Atkinson 1998) and measure improvements to image data (Pedit 2003). Whereas descriptive statistics, such as the standard deviation, give the magnitude of the noise, the spatial structure gives a more visual perception of noise (Pedit 2003). Semivariograms were estimated using the SAS procedure PROC VARIOGRAM[11], and examined to quantify changes in the spatial structure of data introduced by filtering.

Sections with gross defects were examined in detail. For each defect under each filtering method, the start and end points and approximate defect areas were compared. Approximate areas were found by numerical integration over the region of the defect. Numerical integration was performed using the SAS procedure PROC EXPAND with cubic spline interpolation.

## 2.5 Results

### 2.5.1 Method Parameters

Sample board 001 was used to choose the parameters for each of the three methods. This circular-sawn board was specifically selected because it was free of obvious saw marks and defects. Figure 2-3 compares the unfiltered scan with four versions of the Method 1 (median) filter. The first two filters used three points with one and two passes, respectively. The last two filters used five points with one and two passes, respectively. A 30 cm (one-foot) section of the board (from 60 to 90 cm, or 12 to 24 inches) is shown for better detail. The sharp dip at 48 cm

[11] All SAS procedures referred to in this chapter were run with Version 8.2 (SAS Institute 2002).

corresponds to a section where two points spike 0.6mm (0.025 inch) in comparison to their neighbors. (Two points translate as a distance along the board of less than 1 mm, or $\frac{1}{16}$ inch). Sharp peaks were smoothed by the median filters, especially during the second pass of the filter. The 3-point 1-pass median filter preserved the most detail, but did not filter the dip at 48 cm.

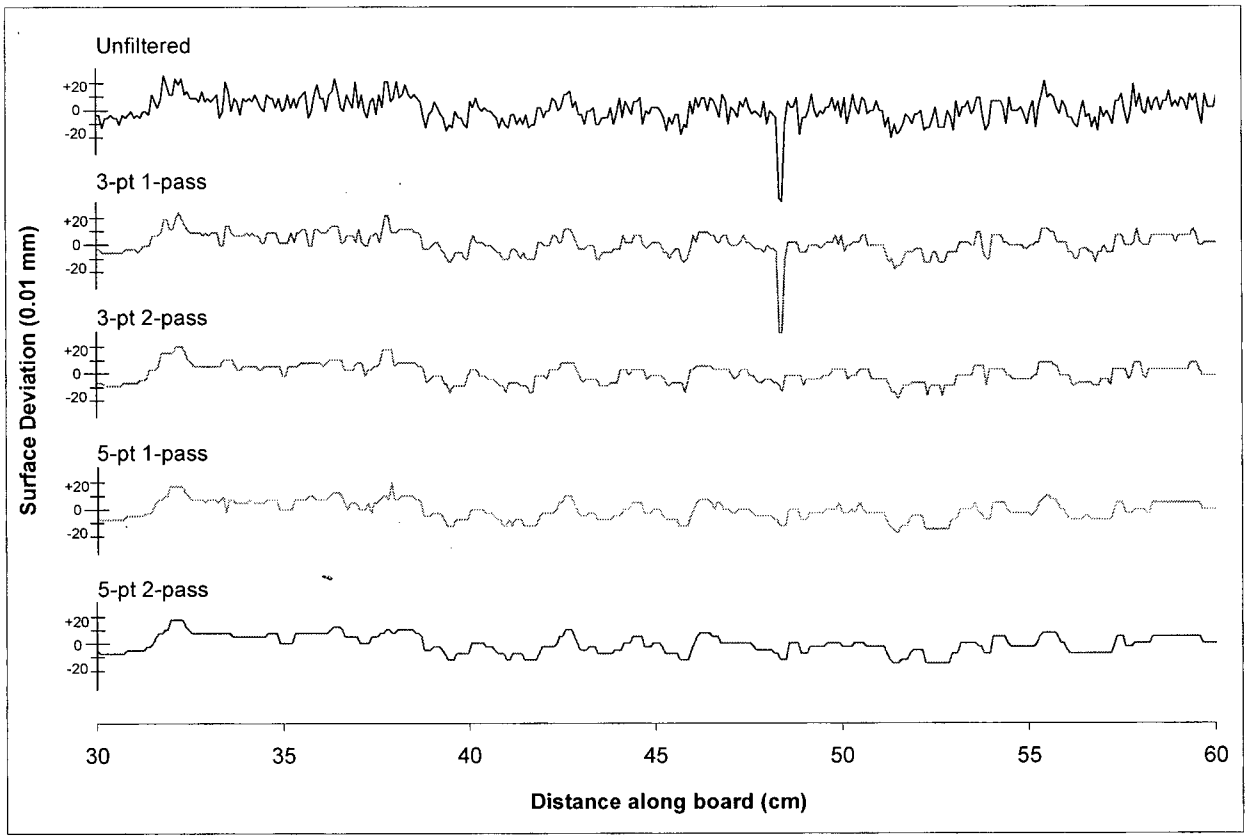



**Figure 2-3. Board 001 unfiltered scan versus four versions of the Method 1 (median) filter[12].**

Table 2-1 shows the simple statistics for the 30-cm section shown above. Because of the large volume of data (~400 points), there were only minor differences between methods. Using any of the four versions of Method 1 had little impact on the mean and maximum value. Under the 3-point 1-pass version, the minimum was slightly lower because it preserved the dip. With each of the four versions, the standard deviation decreased, indicating additional smoothing.

[12] The y-axis for each of Figure 2-3 through Figure 2-11 is the deviation from the average distance from the laser to the lumber surface (in 0.01 mm).

**Table 2-1. Descriptive statistics (cm) for Board 001 versus four versions of Method 1 filter (30-60 cm only).**

| Method | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|
| Unfiltered | 39.13 | 39.06 | 39.15 | 0.0102 |
| 3-point 1-pass | 39.13 | 39.06 | 39.15 | 0.0089 |
| 3-point 2-pass | 39.13 | 39.11 | 39.15 | 0.0072 |
| 5-point 1-pass | 39.13 | 39.11 | 39.15 | 0.0072 |
| 5-point 2-pass | 39.13 | 39.11 | 39.14 | 0.0070 |

Figure 2-4 compares the same unfiltered scan as shown in Figure 2-3 with four versions of the Method 2 (Lee's sigma) filter. The first two filters used seven points with one and two passes, respectively. The last two filters used nine points with one and two passes, respectively. As above, the 1-pass versions preserved the most detail. The sharp dip at 48 cm along the board was smoothed rather than passed in all but the 9-point 2-pass method, and all versions reduced the size of the dip substantially.

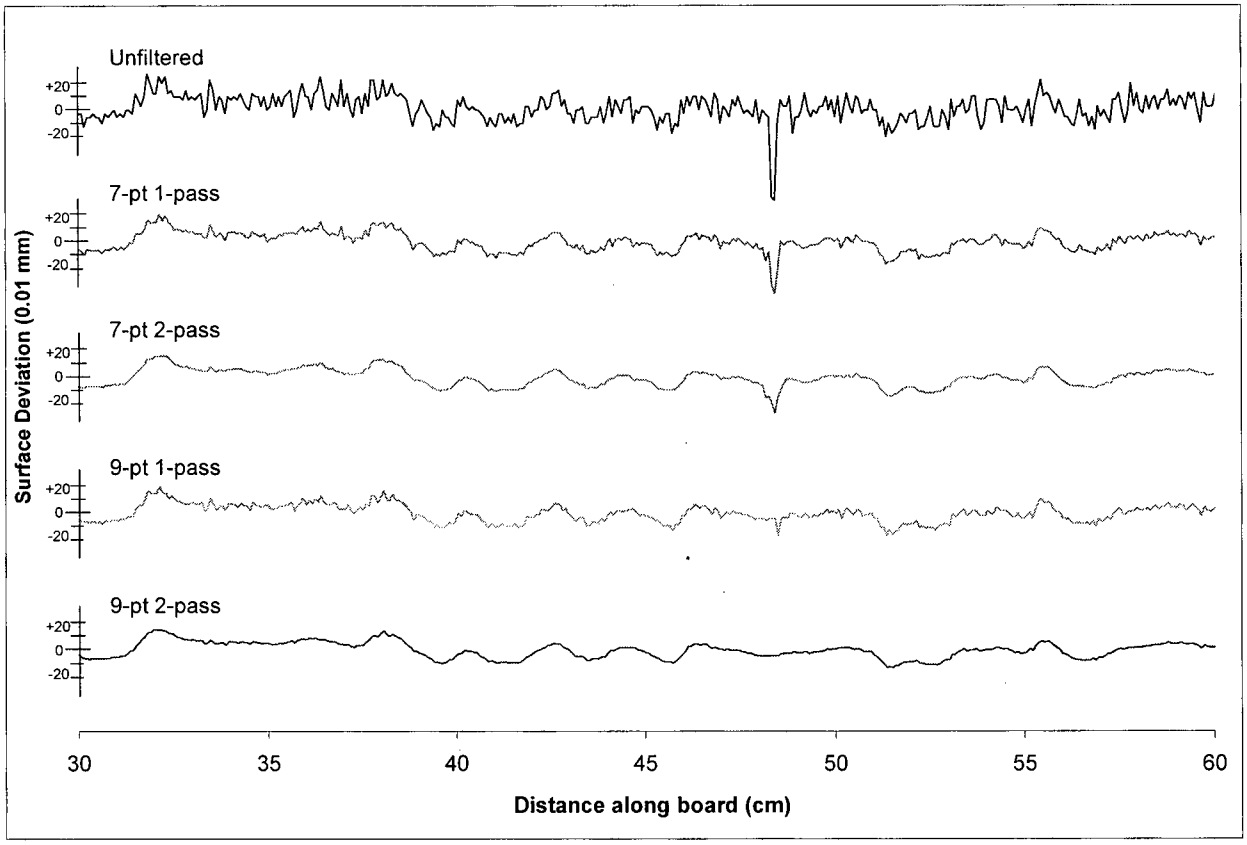



**Figure 2-4. Board 001 unfiltered scan versus four versions of the Method 2 (Lee's sigma) filter.**

As in the median method, using any of the four versions of Method 2 had little impact on the mean and maximum value (Table 2-2). Each of the four versions gave marginal increases in the minimum value, and each decreased the standard deviation, indicating additional smoothing.

**Table 2-2. Descriptive statistics (cm) for Board 001 versus four versions of Method 2 filter (30-60 cm only).**

| Method | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|
| Unfiltered | 39.13 | 39.06 | 39.15 | 0.0102 |
| 7-point 1-pass | 39.13 | 39.09 | 39.15 | 0.0073 |
| 7-point 2-pass | 39.13 | 39.10 | 39.14 | 0.0066 |
| 9-point 1-pass | 39.13 | 39.11 | 39.15 | 0.0065 |
| 9-point 2-pass | 39.13 | 39.11 | 39.14 | 0.0059 |

Figure 2-5 compares the same unfiltered scan with four versions of the Method 3 (MSD) filter. The first two filters used a 3-point MSD with 3-sigma and 2-sigma thresholds for the MSD, respectively. The last two filters used a 5-point MSD with 3-sigma and 2-sigma thresholds, respectively. In contrast to the previous methods, Method 3 preserved the detail and variation inherent in the data, but removed anomalous points. The dip at 48 cm was completely removed in all versions. There was very little difference between the filters using 3-point and 5-point MSDs, whereas there was a large difference between 2- and 3-sigma thresholds; the 2-sigma threshold tended to unnecessarily remove many points that were not anomalous measures.

As in the previous methods, using any of the four versions of Method 3 had little impact on the mean and maximum value (Table 2-3). Each of the four versions increased the minimum value, and each decreased the standard deviation. However, this decrease was much less than in previous methods. The standard deviations under Method 3 were about 25% higher on average than those of Methods 1 and 2.

**Table 2-3. Descriptive statistics (cm) for Board 001 versus four versions of Method 3 filter (30-60 cm only).**

| Method | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|
| Unfiltered | 39.13 | 39.06 | 39.15 | 0.0102 |
| 3-point 3-sigma | 39.13 | 39.10 | 39.15 | 0.0089 |
| 3-point 2-sigma | 39.13 | 39.10 | 39.15 | 0.0086 |
| 5-point 3-sigma | 39.13 | 39.10 | 39.15 | 0.0089 |
| 5-point 2-sigma | 39.13 | 39.10 | 39.15 | 0.0089 |

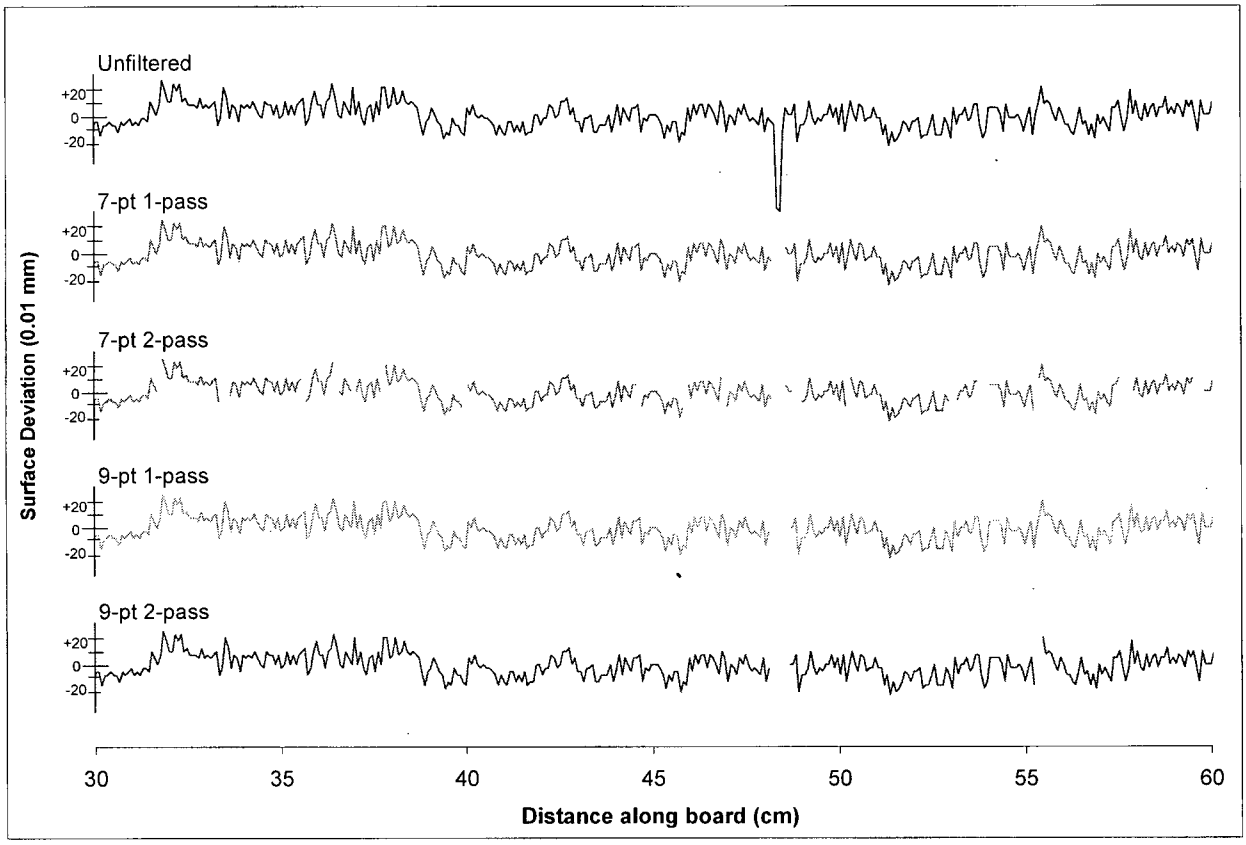



**Figure 2-5. Board 001 unfiltered scan versus four versions of the Method 3 (MSD) filter.**

For the remainder of this section, only one version of each filtering method was considered. For Method 1, the 5-point 1-pass version was chosen; for Method 2, the 9-point 1-pass version was chosen. The Method 3 version chosen used the 5-point window and 3-sigma MSD threshold. These choices were made subjectively. For Methods 1 and 2, the decision was based, in part, on filtering of the dip at 48 cm. How such a dip should be handled in the filtering methodology depends on the precision of the data and the use of the data post-filtering. Since these data are to be used for real-time size control, this degree of detail was deemed unnecessary and methods that passed the dip were preferred. On the other hand, some small-scale variation was desired. The parameters chosen represent a balance between these preferences.

## 2.5.2 Comparison of Methods

### 2.5.2.1 Descriptive Evaluation

Boards were compared with graphs and simple statistics using their entire lengths[13]. In Figures 2–6 through 2–11, scanned boards were compared to a "manual filter", created by manually removing non-sawing defects and anomalous measurements from the scans.

Figure 2-6 compares the Board 001 unfiltered scan with the manual filter and Methods 1-3. Although no visible defects were noted, the unfiltered scan contained many anomalous measurements, appearing as sharp spikes. These observations could have been raised wood fibers or errors resulting from inadequate reflectance of the rough surface. Methods 1 and 2 removed more of these anomalous measurements than Method 3, but also smoothed the data.

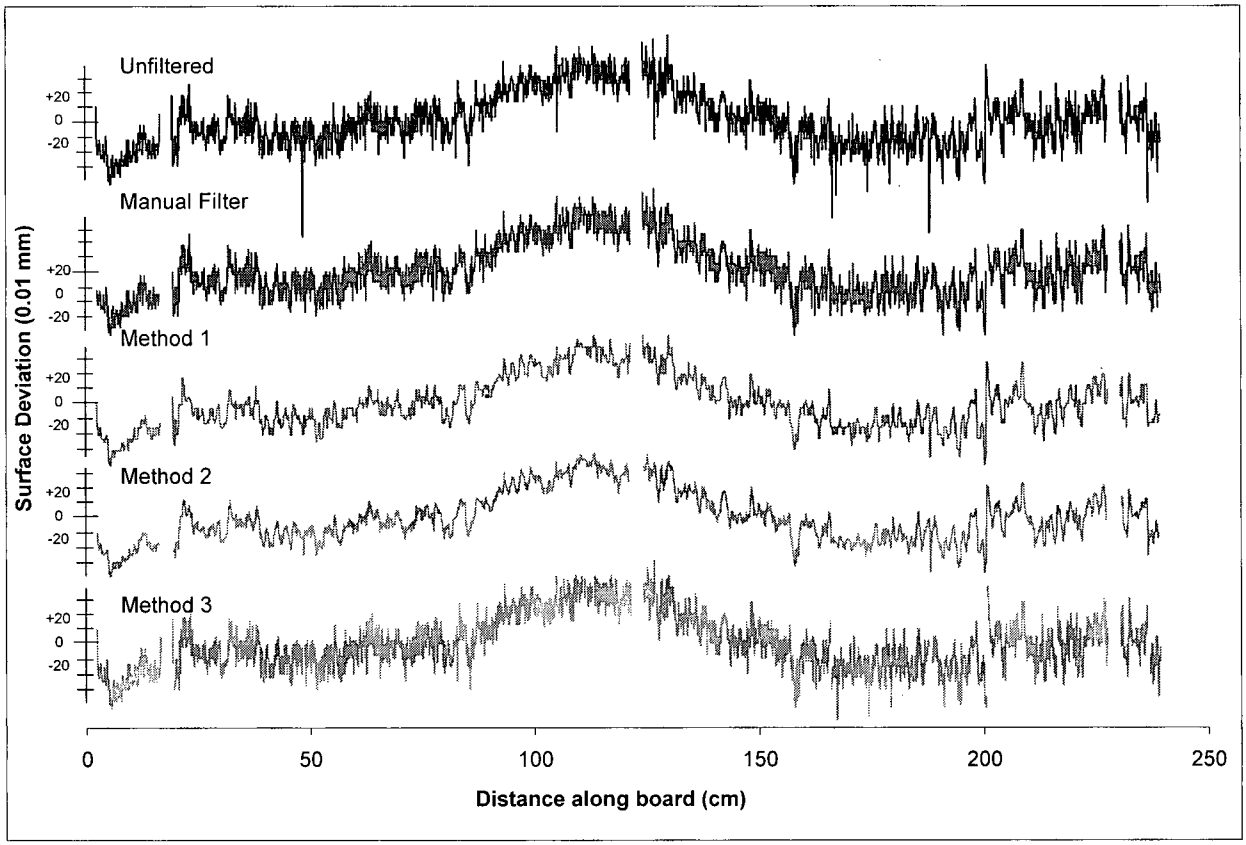



**Figure 2-6. Board 001 unfiltered scan versus manual filter and three filtering methods.**

[13] Note that missing values occurred at 15 cm, 122 cm, and 229 cm due to the clamping design.

Table 2-4 shows the simple statistics for Board 001 versus the manual filter and three filtering methods. Using any of the methods raised the minimum value, but had little impact on the mean and maximum value. Under Methods 1 and 2, the standard deviation decreased marginally.

**Table 2-4. Descriptive statistics (cm) for Board 001 versus manual filter and three filtering methods.**

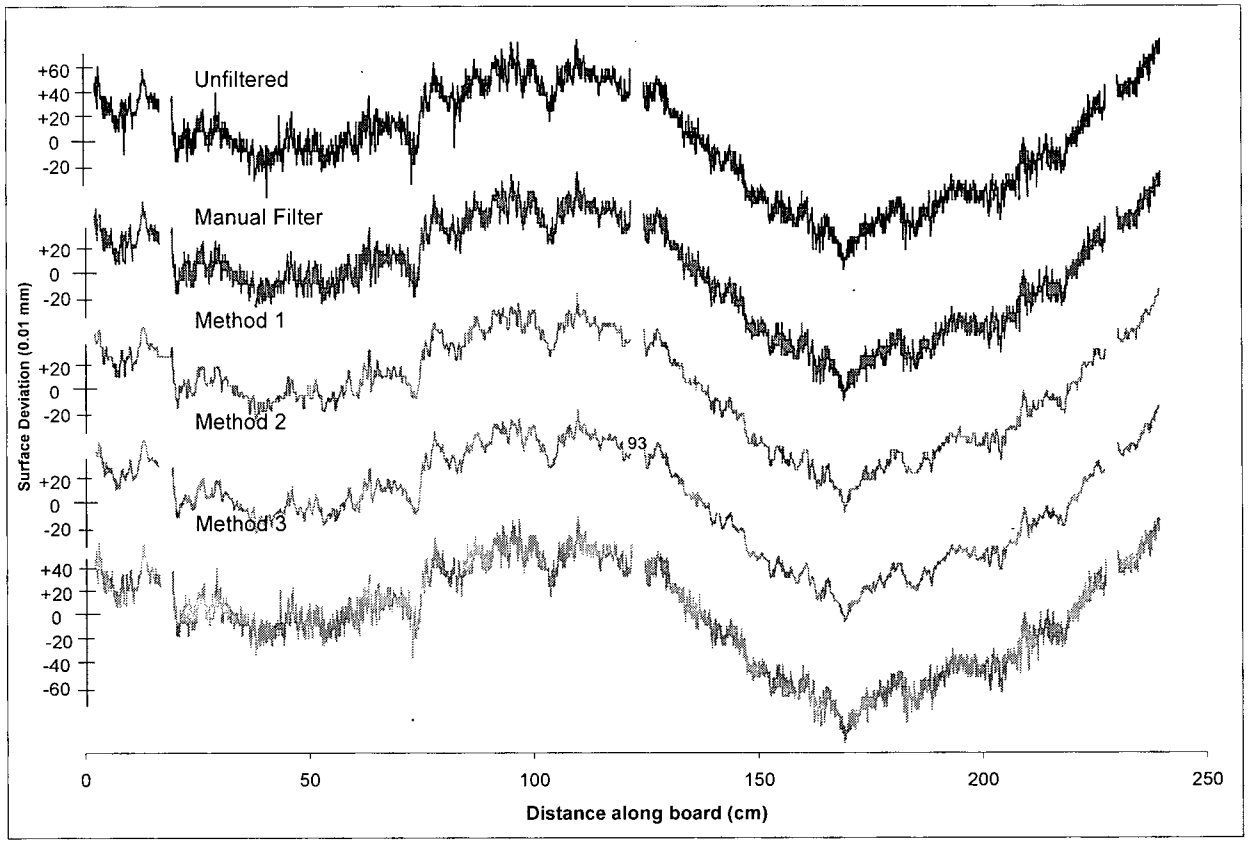
| Method | N | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| Unfiltered | 2901 | 39.13 | 39.06 | 39.19 | 0.0181 |
| Manual filter | 2870 | 39.13 | 39.09 | 39.19 | 0.0177 |
| Method 1 | 2899 | 39.13 | 39.09 | 39.18 | 0.0168 |
| Method 2 | 2869 | 39.13 | 39.09 | 39.18 | 0.0165 |
| Method 3 | 2828 | 39.13 | 39.08 | 39.19 | 0.0173 |

Sample board 002 is shown in Figure 2-7. This board was bandsawn, and like Board 001, it was noted to be free of saw marks and other defects, with anomalous measurements present in the unfiltered scan. Two short dips in the laser signal at approximately 74 and 102 cm along the scan were caused by small sections of rough fibers. As in the previous figures, the unfiltered scan is shown versus the manual filter and three filtering methods. Methods 1 and 2 eliminated more of the anomalous measurements, whereas Method 3 only eliminated extreme points.

Table 2-5 shows the simple statistics for Board 002. As in Board 001, the methods had almost no impact on the mean, minimum, or maximum values. Methods 1 and 2 reduced the standard deviation. Because Method 3 resulted in the deletion of observations that were close to the mean value, filtering with Method 3 increased the standard deviation slightly.

**Table 2-5. Descriptive statistics (cm) for Board 002 versus manual filter and three filtering methods.**

| Method | N | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| Unfiltered | 2906 | 39.01 | 38.91 | 39.09 | 0.0399 |
| Manual filter | 2886 | 39.01 | 38.91 | 39.09 | 0.0399 |
| Method 1 | 2902 | 39.01 | 38.91 | 39.08 | 0.0396 |
| Method 2 | 2883 | 39.01 | 38.91 | 39.08 | 0.0395 |
| Method 3 | 2888 | 39.01 | 38.91 | 39.09 | 0.0400 |




S

**Figure 2-7. Board 002 unfiltered scan versus manual filter and three filtering methods.**

Board 003 was processed by the chipper-head and was noted to have several non-sawing and sawing defects along the path of the laser (Figure 2-8). Wane occurred from the beginning of the board to 15 cm, saw tear-out occurred at 107-117 cm, 130 cm, and 142-150 cm. The manual filter preserved the sawing defect (tear-out) in the scan, while the non-sawing defect (wane) was removed from the dataset. The defect regions were preserved in Methods 1 and 2, whereas Method 3 partially removed both the wane and tear-out areas.

Table 2-6 shows simple statistics for Board 003. No method affected the mean value, but all methods raised the minimum value. Only Method 3 substantially changed the maximum and standard deviation values, bringing them to the level of the manual filter.

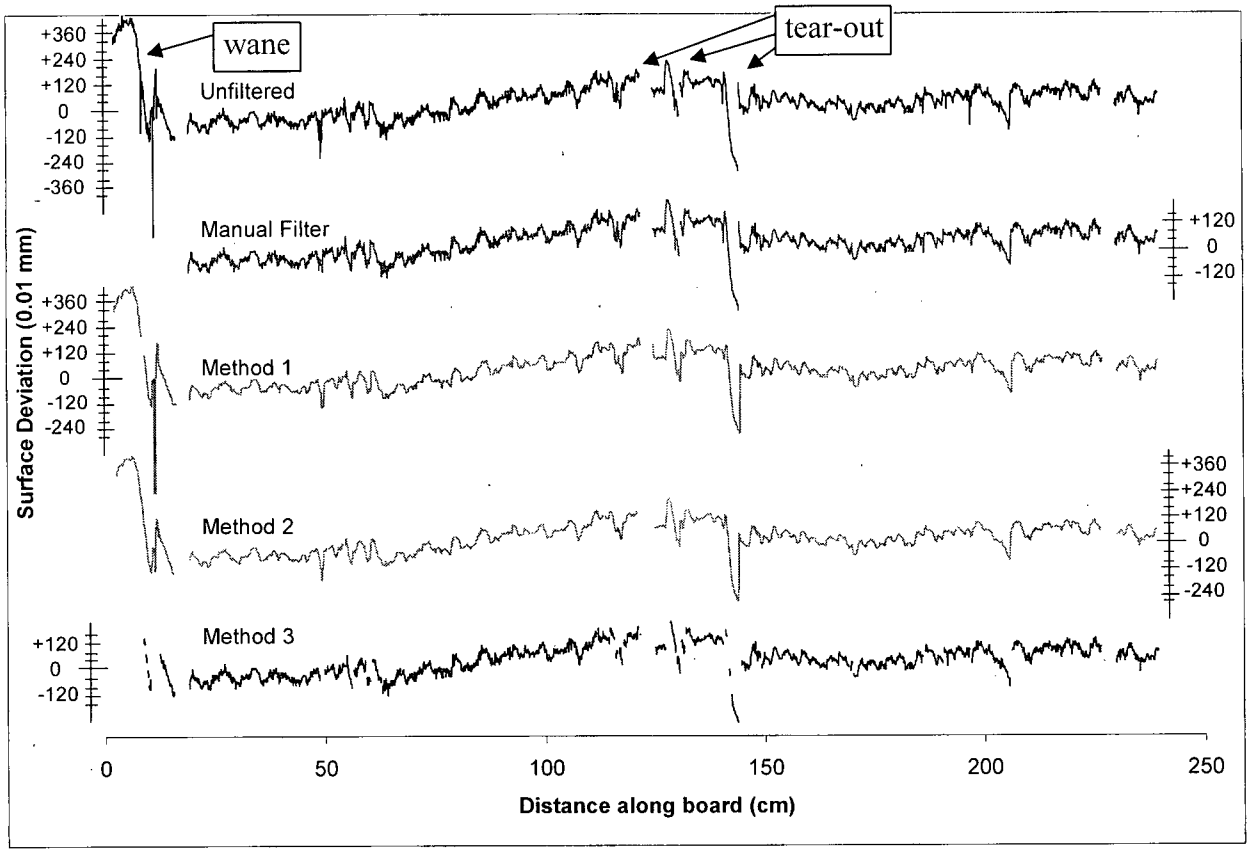



**Figure 2-8. Board 003 unfiltered scan versus manual filter and three filtering methods.**

**Table 2-6. Descriptive statistics (cm) for Board 003 versus manual filter and three filtering methods.**

| Method | N | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| Unfiltered | 2892 | 38.98 | 38.34 | 39.42 | 0.091 |
| Manual filter | 2704 | 38.97 | 38.65 | 39.17 | 0.068 |
| Method 1 | 2887 | 38.98 | 38.38 | 39.38 | 0.091 |
| Method 2 | 2851 | 38.98 | 38.67 | 39.38 | 0.087 |
| Method 3 | 2649 | 38.97 | 38.67 | 39.16 | 0.068 |

A graph of the unfiltered and filtered scans for Board 004 is shown in Figure 2-9. This chipped board had a large section of wane (from the beginning of the board to 38 cm), two sections of tear-out (64-71 cm and 147-152 cm), and a prominent splinter at 183 cm. The roughness of the wane section resulted in some missing values in the original unfiltered laser data. As in the previous sample, the wane section was preserved by the first two methods, but was mostly removed by the third method. No method removed the second section of tear-out, and only Method 3 eliminated the first section. All methods removed the splinter.

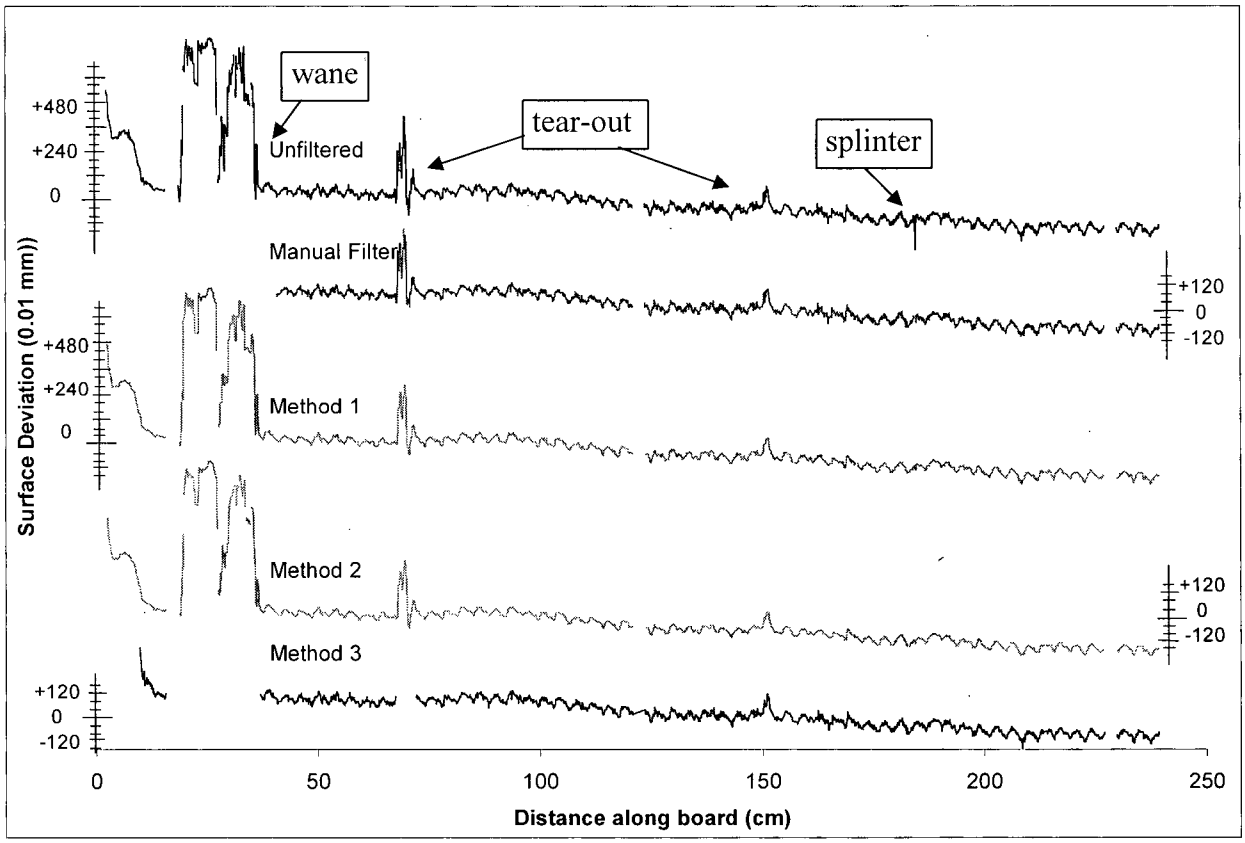



**Figure 2-9. Board 004 unfiltered scan versus manual filter and three filtering methods.**

Table 2-7 shows the simple statistics for Board 004. Method 3 noticeably changed the mean value, standard deviation, and maximum values. While it would appear visually that Method 1 reduced the variability of the data, the standard deviation under Method 1 was slightly greater than that of the unfiltered data. This was a result of missing values in the wane section of the original unfiltered scan being replaced by median values with the median filter.

**Table 2-7. Descriptive statistics (cm) for Board 004 versus manual filter and three filtering methods.**

| Method | N | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| Unfiltered | 2891 | 39.05 | 38.72 | 40.57 | 0.351 |
| Manual filter | 2431 | 38.43 | 38.28 | 38.83 | 0.076 |
| Method 1 | 2886 | 39.05 | 38.79 | 40.57 | 0.352 |
| Method 2 | 2868 | 39.05 | 38.80 | 40.57 | 0.349 |
| Method 3 | 2495 | 38.94 | 38.77 | 39.29 | 0.075 |

Graphs of the unfiltered and filtered scans for Board 005 are shown in Figure 2-10. This bandsawn sample had one sawing defect: a deep tear at 69-81 cm. This tear was abrupt and

rough, resulting in missing values in the original unfiltered laser data. Methods 1 and 2 retained the tear, while Method 3 partially eliminated it.

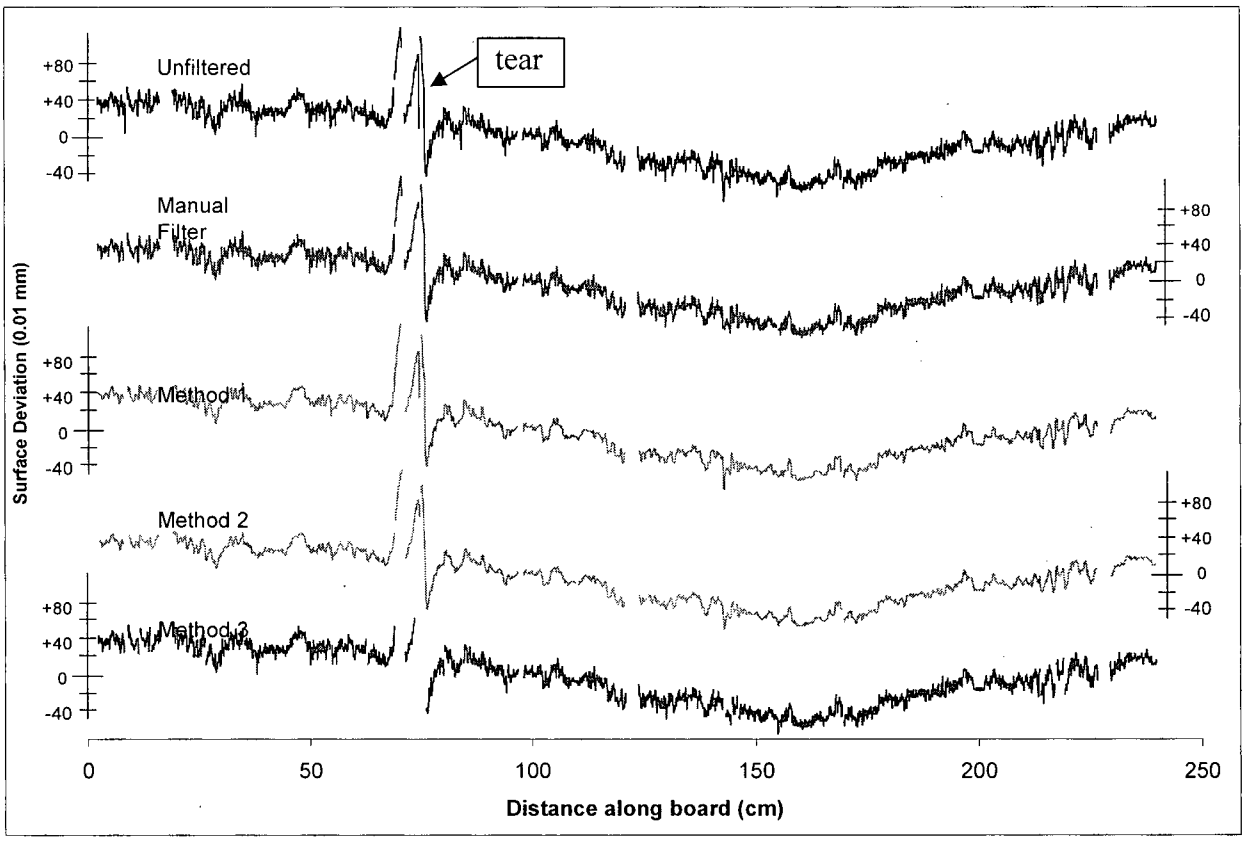



**Figure 2-10. Board 005 unfiltered scan versus manual filter and three filtering methods.**

Table 2-8 shows the simple statistics for Board 005. There was little difference between methods for the mean, minimum, and standard deviation values. Only Method 3 produced a lower maximum value.

**Table 2-8. Descriptive statistics (cm) for Board 005 versus manual filter and three filtering methods.**

| Method | N | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| Unfiltered | 2867 | 38.98 | 38.84 | 39.22 | 0.061 |
| Manual filter | 2845 | 38.98 | 38.86 | 39.22 | 0.060 |
| Method 1 | 2848 | 38.98 | 38.85 | 39.21 | 0.060 |
| Method 2 | 2833 | 38.98 | 38.87 | 39.21 | 0.060 |
| Method 3 | 2764 | 38.98 | 38.85 | 39.11 | 0.056 |

Graphs of the unfiltered and filtered scans for Board 006 are shown in Figure 2-11. This sample was chipped and had one non-sawing defect: wane from 221 cm to the end of the board. A

splintery region at 69 cm was also apparent on the scan. Method 1 retained both of these features, Method 2 partially removed them, and Method 3 mostly eliminated them.

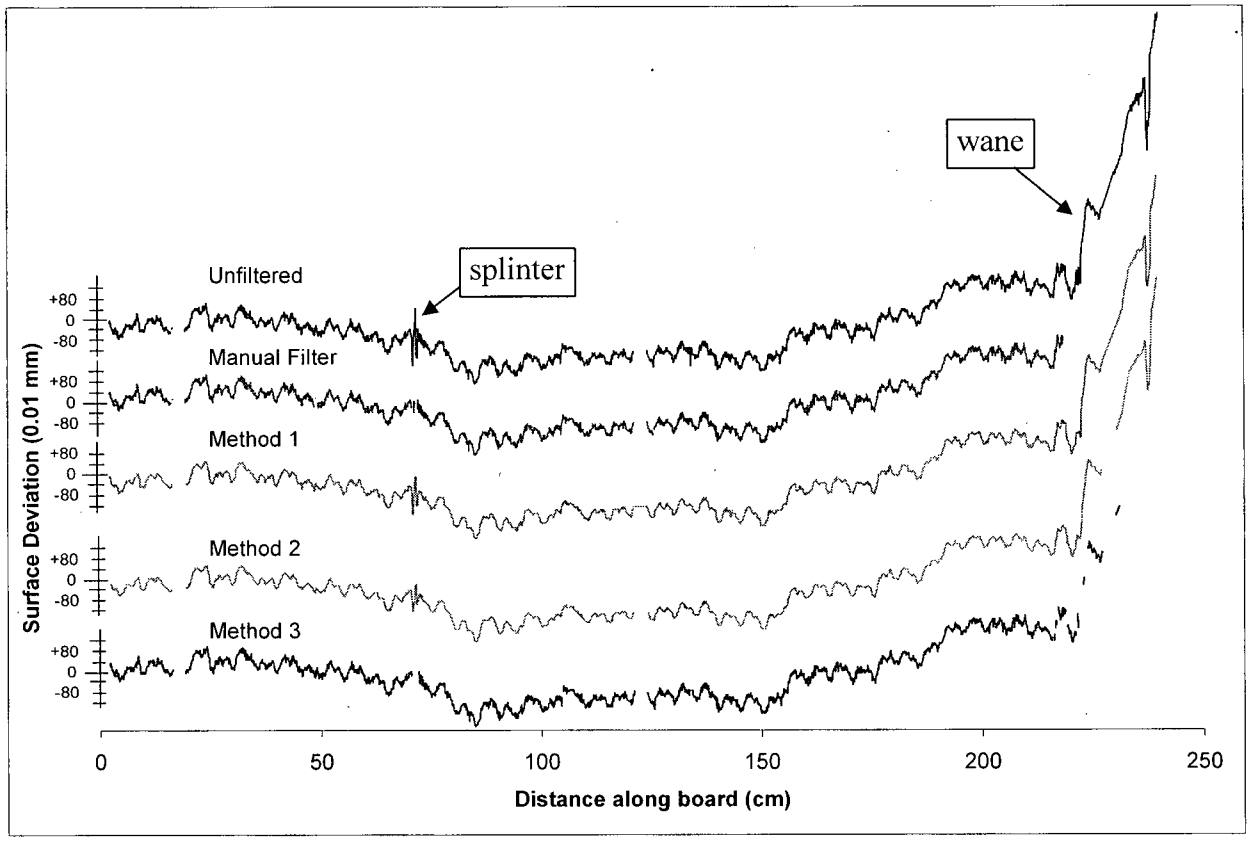



**Figure 2-11. Board 006 unfiltered scan versus manual filter and three filtering methods.**

Table 2-9 shows the simple statistics for Board 006. There was little difference between Methods 1 and 2 for the mean, minimum, maximum, and standard deviation values. With Method 3, the mean, maximum, and standard deviation values were lower.

**Table 2-9. Descriptive statistics (cm) for Board 006 versus manual filter and three filtering methods.**

| Method | N | Mean | Minimum | Maximum | Standard Deviation |
|---|---|---|---|---|---|
| Unfiltered | 2899 | 38.98 | 38.72 | 40.17 | 0.216 |
| Manual filter | 2643 | 38.93 | 38.72 | 39.19 | 0.101 |
| Method 1 | 2902 | 38.98 | 38.73 | 40.15 | 0.214 |
| Method 2 | 2876 | 38.97 | 38.73 | 40.15 | 0.211 |
| Method 3 | 2738 | 38.94 | 38.72 | 39.73 | 0.127 |

## 2.5.2.2 Performance by Defect Type

Table 2-10 compares the filtering of defect areas for each sample. The percentage refers to the proportion of the board area preserved after filtering. The percentage shown under the manual

filtering method is the desired result of filtering. For instance, the wane in Board 003 should be eliminated, whereas the tear-out should remain after filtering. Methods 1 and 2 gave similar results, regardless of the type of defect; defects were smoothed, and therefore the areas retained as defect-free were close to 100% in almost all cases. Under Method 3, more defect areas were eliminated, with about half or both wane and tear-out areas eliminated. Overall, no method completely eliminated wane while completely preserving tear-out.

**Table 2-10. Comparison of defect areas by filtering method.**

| | | Percentage of Total Board Area | | | | |
|---|---|---|---|---|---|---|
| Board | Defect | Unfiltered | Manual | Method 1 | Method 2 | Method 3 |
| 001 | None | 100.0% | 98.9% | 99.9% | 98.9% | 97.5% |
| 002 | None | 100.0% | 99.3% | 99.9% | 99.2% | 99.4% |
| 003 | None | 83.8% | 83.4% | 83.6% | 83.5% | 81.8% |
| | Wane 0-15 cm | 6.6% | 0.0% | 5.5% | 5.6% | 1.7% |
| | Tear-out 107-117 cm | 4.9% | 4.9% | 4.9% | 4.9% | 3.3% |
| | Tear-out 130 cm | 0.7% | 0.7% | 0.6% | 0.7% | 0.3% |
| | Tear-out 142-150 cm | 3.9% | 3.9% | 3.8% | 3.9% | 2.6% |
| 004 | None | 77.3% | 77.0% | 77.0% | 77.0% | 77.2% |
| | Wane 0-38 cm | 16.0% | 0.0% | 16.0% | 16.0% | 3.7% |
| | Tear-out 64-71 cm | 3.9% | 3.9% | 3.9% | 3.9% | 2.1% |
| | Tear-out 147-152 cm | 2.9% | 2.9% | 2.9% | 2.9% | 2.7% |
| 005 | None | 94.0% | 93.6% | 93.8% | 93.8% | 93.8% |
| | Tear at 69-81 cm | 6.0% | 6.0% | 6.0% | 6.0% | 3.5% |
| 006 | None | 90.9% | 90.5% | 99.0% | 97.3% | 90.5% |
| | Wane 221 cm to end | 9.1% | 0.0% | 8.1% | 6.4% | 3.6% |

### 2.5.2.3 Spatial Variability

The affect of filtering on spatial variability was illustrated using the semivariogram. Figures 2-12 – 2–17 show the semivariograms of the unfiltered scans versus the different filtering methods. For Boards 001, 002, and 005, the semivariograms for all of the methods were close to that of the unfiltered. On the other hand, for Boards 003, 004, and 006, the semivariograms for Methods 1 and 2 tracked much closer to the unfiltered scan than those of Method 3 or the manual filter. For these samples, the general shape of the semivariograms was variable for Methods 1 and 2, but was closer to bell-shaped under Method 3 and the manual filter.

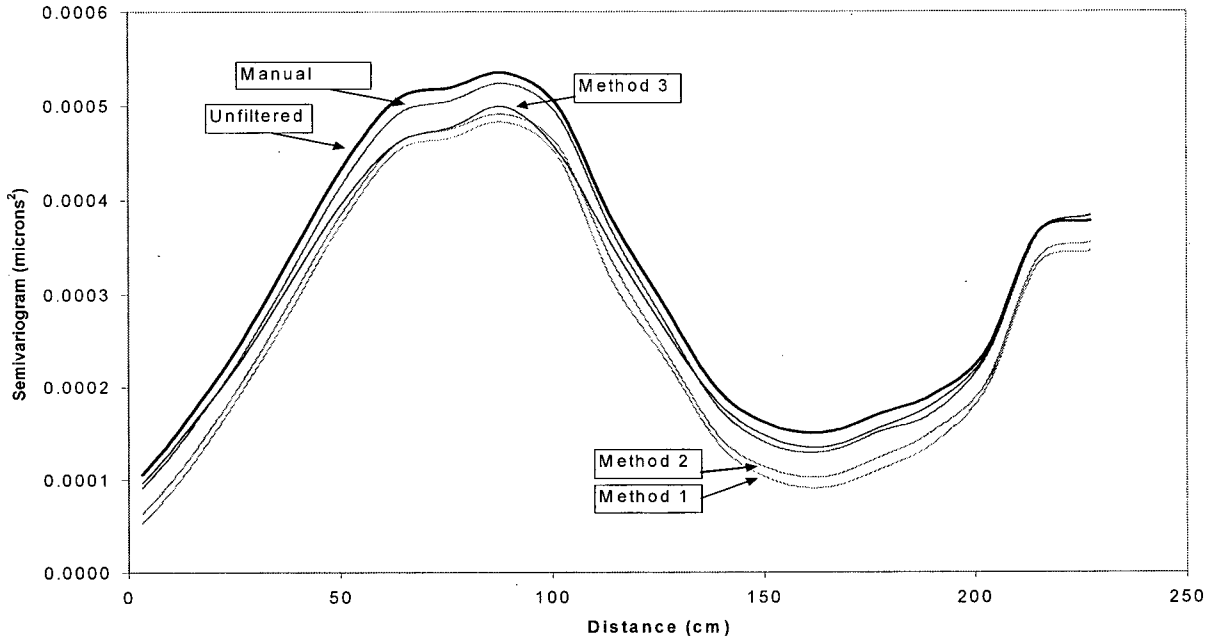



**Figure 2-12. Board 001 semivariogram.**

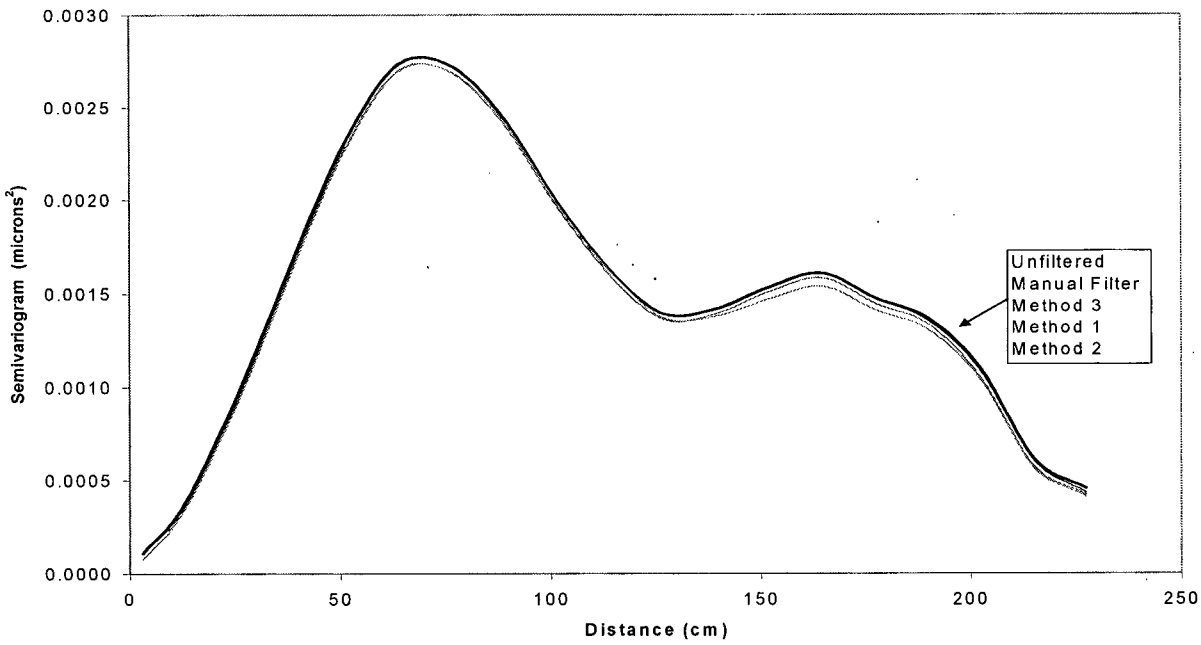



**Figure 2-13. Board 002 semivariogram.**

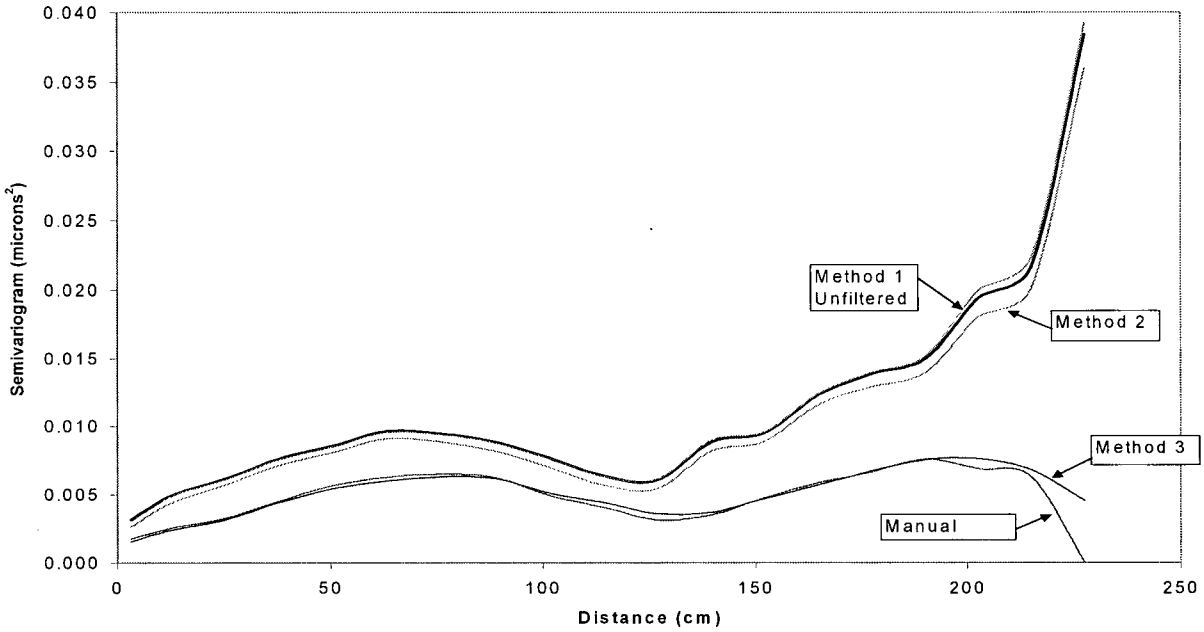



**Figure 2-14. Board 003 semivariogram.**

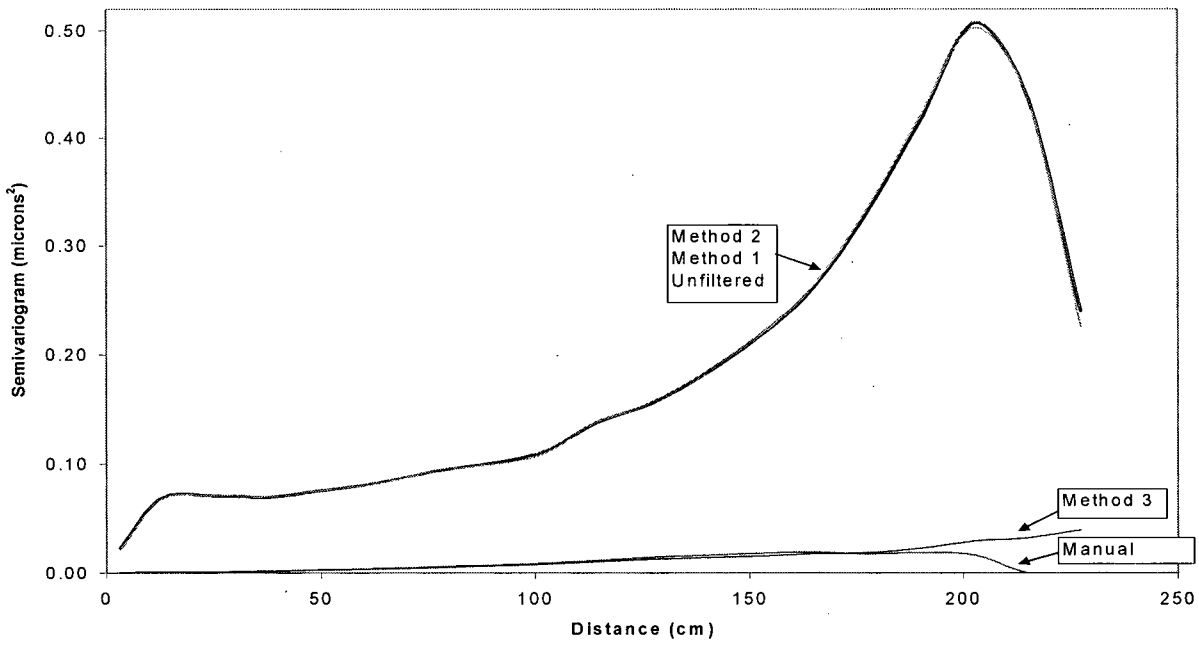



**Figure 2-15. Board 004 semivariogram.**

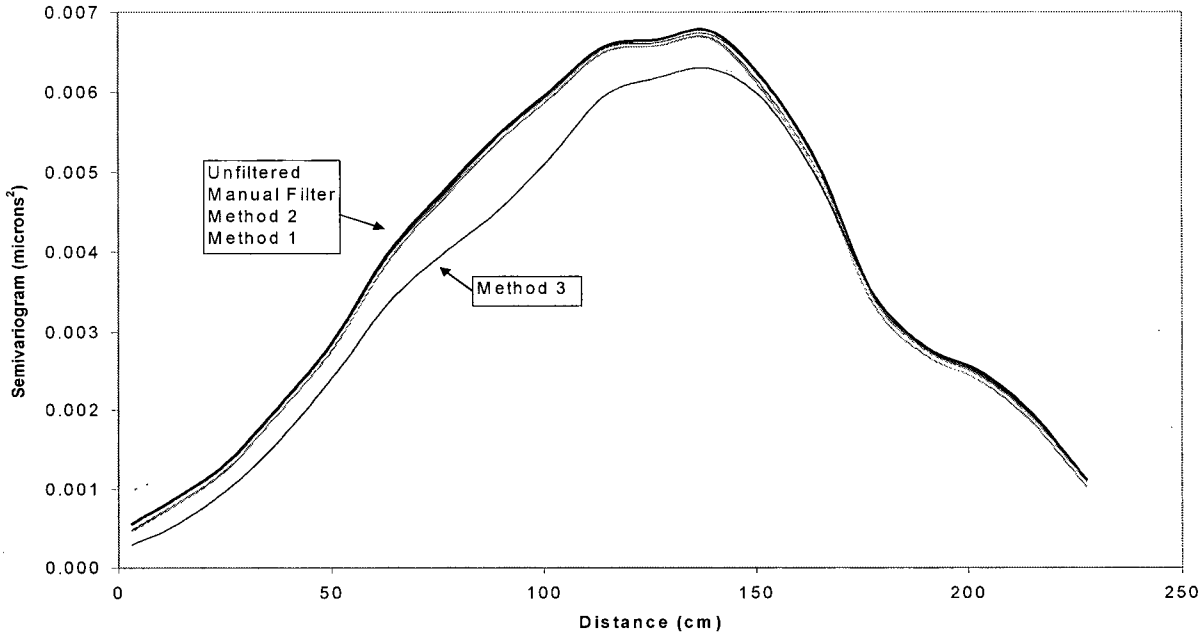



**Figure 2-16. Board 005 semivariogram.**

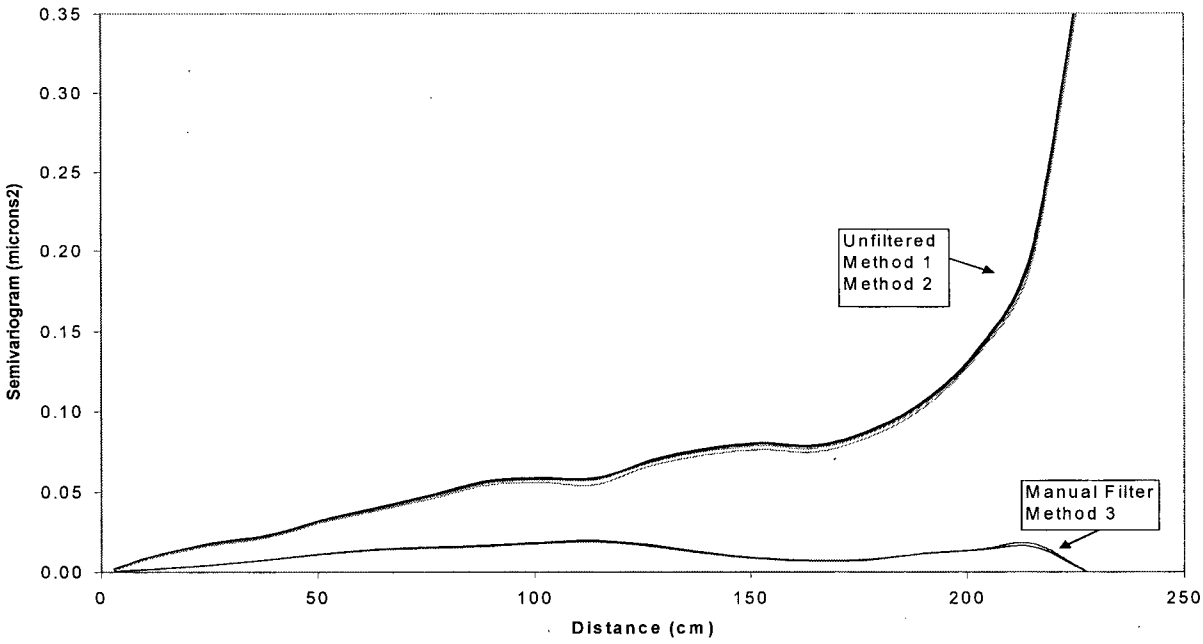



**Figure 2-17. Board 006 semivariogram.**

## 2.6 Discussion

For all the sample boards, Methods 1 and 2 tended to eliminate small-scale variation in the measurements, while Method 3 retained small scale variation and eliminated only gross errors.

For example, the small-scale variation from the chipper-head in Board 003 was very much smoothed using Methods 1 and 2, but was preserved in Method 3.

None of the filtering methods appreciably changed the descriptive statistics of the data, except in terms of the variance. When defects were present, Method 3 partially eliminated these areas, reducing variation. Similar to the findings of Funck *et al.* (1992), Methods 1 and 2 tended to slightly reduce variation in defect-free areas, as these methods eliminated anomalous measurements, which tended to be outliers. However, these methods also eliminated the finer details in the scans.

A small amount of variation in the data is expected, because the LRSs have a small amount of inherent error due to their precision. Moreover, lumber is not a perfectly smooth surface. Saw marks are expected, as saws are comprised of thousands of teeth, each making a bite into the lumber. Since the average bite per tooth of a primary bandsaw in a typical mill is 0.4 to 1 mm (Schajer 1990) and observations from the LRS were taken at approximately 0.8 mm intervals, some variation in the data is expected; however, Methods 1 and 2 smoothed this variation.

When filtering defect areas, Method 3 completely or partially eliminated wane as well as tear-out. Methods 1 and 2, on the other hand, tended to preserve these features. For example, in Board 005, Methods 1 and 2 smoothed the deep tear, while Method 3 retained only about 60% of the tear area. This is consistent with the findings of Lee *et al.* (2003), who found a local statistics method unsuitable for their needs, as areas with large defects were important for later processing.

In the case of SPC data collection, the automatic processing of data containing wane and other non-sawing defects could lead to false signals of an out of control process. Thus, these defects should be detected and eliminated from data prior to processing, and a method that filters these defects is desirable. Conversely, sawing defects, such as tear-out, should be retained. Although no method adequately distinguished non-sawing from sawing defects, it is important to note that

sawing defects, such as tear-out, tend to be short in length and therefore, the negative impact of eliminating these defects on the dataset as a whole is minimal. Non-sawing defects, on the other hand, are a much larger part of the dataset and therefore the positive impact of eliminating these defects has a far larger impact on the dataset.

Retaining patterns of spatial variability from unfiltered scans may also be important for post-processing, especially if spatial or serial autocorrelation is incorporated into modeling efforts. When defects were present, Method 3 semivariograms were much different from that of the other methods and the unfiltered scan, and more similar to that of the manual filter. The Method 3 semivariogram tended to be lower than that of the unfiltered scans at larger distances, indicating that observations farther apart were more highly correlated when filtered with Method 3. When no defects were present, the semivariograms from Methods 1 and 2 tended to be parallel and lower than that of the unfiltered data, indicating that spatial correlation was introduced uniformly by filtering with Methods 1 and 2.

The semivariogram was decreased by Methods 1 and 2, indicating an increase in the information contained in each observation and additional redundancy in subsequent observations. It is therefore possible that when methods one or two are used, data may be compressed for later analysis stages without loss of meaningful information in the scan.

## 2.7 Conclusion

For each way laser and image data are used, there are different filtering needs. If the data are used for wane detection and elimination, for instance, the primary goals of filtering are likely to be edge preservation and definition. For SPC, the primary goal of filtering is to ensure that the laser data are representative of the sawing process. Therefore, an SPC data filter should eliminate erroneous measurements and non-sawing defects from the data.

Both the median methods and Lee's sigma method greatly smoothed the scan while eliminating spurious noise. Lee's sigma method, based on averages, tended to smooth defect edges in all but the most abrupt cases. The median method tended to keep defect edges. The MSD method, based on local statistics, preserved the variation inherent in the scan data while eliminating most spurious noise and defects. If the goal of filtering is to eliminate short-term noise, such as saw marks, the median method is recommended. On the other hand, if defect elimination is of primary concern, the MSD method is more appropriate.

The MSD method is recommended for use with SPC data; however, future research should include improvements. Specifically, modifications to this method should improve defect type recognition; however, differentiating between different kinds of defects is difficult with laser data alone, as holes and wane can appear very much the same if laser data are taken in isolation. This difficulty has led researchers to investigate multi-sensor approaches involving additional types of sensors. Therefore, suggested improvements to filtering include adding hardware that can be used in conjunction with LRSs, such as colour or black and white cameras (e.g., Funck *et al.* 1992; Kline *et al.* 2001). Although filtering with LRSs alone is adequate, improving defect recognition would mean that the objective of filtering for measurement errors and non-sawing defects was fully accomplished.

## 2.8 Literature Cited

Abbott, A. L., D. L. Schmoldt, P. A. Araman and S. M. Lee. 2001. Automatic scanning of rough hardwood lumber for edging and trimming. *In* Proceedings of ScanTech 2001, the 9th International Conference on Scanning Technology and Process Optimization for the Wood Industry, November 4-6, Seattle, WA. pp. 101-110.

Axelsson, P. 1999. Processing of laser scanner data - algorithms and applications. ISPRS Journal of Photogrammetry and Remote Sensing 54: 138-147.

Baltasavias, E. P. 1999. Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing 54: 199-214.

Burman, H. 2002. Laser strip adjustment for data calibration and verification. Presented at Symposium 2002: ISPRS Commission III. September 9-13, Graz, Austria.

Coates, E. R., S. J. Chang and T. W. Liao. 1998. A quick defect detection algorithm for magnetic resonance images of hardwood logs. Forest Products Journal 48(10): 68-74.

Curran, P. J. and P. M. Atkinson. 1998. Geostatistics and remote sensing. Progress in Physical Geography 22(1): 61-78.

Funck, J. W., F. A. Gabner and F. Aslam. 1992. Using image analysis for lumber size control measurements. Industrial Metrology 2: 237-252.

Gazzarri, J. 2003. Sequential measurements method for moving surfaces profiling. MSc Thesis. Department of Mechanical Engineering, The University of British Columbia, Vancouver, Canada.

Gonzalez, R. C. and R. C. Woods. 1992. Digital Image Processing. Reading, MA, Addison-Wesley.

Kline, D. E., C. Surak and P. A. Araman. 2001. Evaluation of a multi-sensor machine vision system for automated hardwood lumber grading. *In* Proceedings of 4th International Conference on Image Processing and Scanning of Wood, August 21-23, 2000, Mountain Lake, VA. pp. 75-87.

Kraus, K. and N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 53: 193-203.

Lee, J. S. 1983. Digital image smoothing and the sigma filter. Computer Vision, Graphics, and Image Processing 24: 255-269.

Lee, S. M., A. L. Abbott and D. L. Schmoldt. 2001. Wane detection of rough lumber using surface approximation. *In* Proceedings of 4th International Conference on Image Processing and Scanning of Wood, August 21-23, 2000, Mountain Lake, VA. pp. 115-126.

Lee, S. M., A. L. Abbott and D. L. Schmoldt. 2003. A system for optimal edging and trimming of rough hardwood lumber. *In* Proceedings of 5th International Conference on Image Processing and Scanning of Wood, March 23-26, 2002, Bad Waltersdorf, Austria. pp. 25-34.

Lister, P. 1997. Understanding the relationship between circular saw kerf width, sawing variation, and feed speed. *In* Proceedings of 5th International Conference on Sawing Technology, Seattle, WA. Wood Machining Institute, Berkeley, CA. pp. 17-22.

Maness, T. C. 1993. Real-time quality control system for automated lumbermills. Forest Products Journal 43(7/8): 17-22.

Maness, T. C. 1996. Study shows sawmill quality control programs can be a good investment. Branchlines 7(1): 3.

Maness, T. C. and Y. Lin. 1995. The influence of sawkerf and target size reductions on sawmill revenue and volume recovery. Forest Products Journal 45(11/12): 43-50.

Pedit, J. 2003. Spatial structure of noise in scans from photographic film [online]. Available from http://www.unc.edu/~pedit/noise3/noise3a.html [cited September 20, 1993].

Sandak, J., C. Tanaka and T. Ohtani. 2003. Evaluation of surface smoothness by a laser displacement sensor 1: effect of wood species. Journal of Wood Science 49: 305-311.

Sandak, J., C. Tanaka and T. Ohtani. 2004. Evaluation of surface smoothness by a laser displacement sensor II: comparisons of lateral effect photodiode and multi-element array. Journal of Wood Science 50: 22-27.

SAS Institute. 2002. SAS/STAT User's Guide, Version 8. Cary, NC, SAS Publishing.

Schajer, G. S. 1990. Designing a saw that works involves many factors. Forest Industries 117(9): 17-20.

Soest, J. F. and P. C. Matthews. 1985. Laser scanning technique for defect detection in Sawmilling. *In* Proceedings of 1st International Conference on Scanning Technology, October 10-11, San Francisco, CA. Forest Industries/World Wood. pp. 1-4.

Vosselman, G. and H. Maas 2001. Adjustment and filtering of raw laser altimetry data. Presented at OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models. March 1-3, Stockholm, Sweden.

Wang, S. J. 1983. An analytic approach to estimating the increase in lumber recovery due to reduced target sizes and saw kerfs. Forest Products Journal 33(11/12): 29-32.

Wehr, A. and U. Lohr. 1999. Airborne laser scanning - and introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing 54: 68-82.

# Chapter 3 Surface Defect Recognition in Real-Time Automated Lumber Manufacturing

## 3.1 Introduction

Recent advances in technology have brought about the introduction of laser-based real-time statistical process control (SPC) systems in automated lumber manufacturing. Non-contact laser range sensors (LRS) have been installed to collect size measurements immediately after a particular sawing machine, allowing for online real-time assessment of the quality of the sawing process. Commercially available systems use two or more LRSs to collect data at a rate of up to 1000 measurements per second. At typical primary mill feedspeeds, thickness is measured at <0.8 mm (<0.03 inch) intervals on each and every board and/or cant processed.

Laser data collected in this manner are subject to several sources of measurement error and may contain anomalies. Some errors and anomalies are caused by the inaccuracy of the lasers themselves; for example, the lasers may be used for an application outside of their recommended range. Others result from being in the sawmill environment. For instance, machinery vibration, dust particles, and saw marks or fiber strands left after the sawing process are often in LRS data. Anomalous measurements also result if the scan zone contains lumber with non-sawing defects, such as wane[14], holes, or loose knots.

Before laser data are used in SPC algorithms, anomalous measurements from non-sawing defects should be removed. In a related research project, researchers at The University of British Columbia Mechanical Engineering Department have developed an algorithm to identify and remove machinery vibration from the LRS data (Gazzarri 2003). Removal of measurement error via filtering was investigated in the Chapter 2 using techniques from image processing and digital terrain mapping. These data filtering techniques were adequate for removing sporadic

[14] Wane is the natural curvature of the edge of a board sawn from a log. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

noise from the laser data; however, they were not sufficient to remove non-sawing defects while retaining sawing-defects. Using a method based on moving averages and other local statistics, the best-performing filter removed 60-75% of wane areas, but also removed sawing defects, such as saw tear-out[15]. While it is important to remove non-sawing defects from SPC data, it is equally important to retain data containing sawing defects, as the defects may provide important information for diagnosing size control problems.

The delineation of surface defects, such as wane, decay, and loose knots has been extensively studied in other areas of the lumber manufacturing process. As boards flow through the mill, the most profitable cut is determined based on complex grading rules and quickly changing current market prices (Abbott *et al.* 2001), and grading rules, in turn, are dependent on the amount and position of surface defects. While most lumber mills have introduced automation into edging and trimming operations via machine vision, fully automated grading systems are not yet operationally common in primary lumber manufacturing. Using scanning devices to automate the process of defect detection in lumber grading could improve lumber production, as it could improve the quality of cutting decisions and increase grading accuracy.

Much recent research into automated defect detection in lumber has focused on hardwood edging and trimming optimization, where piece size varies substantially and the rules for lumber grading are complex. This work has targeted rough, un-planed lumber, as decisions made early in the production chain yield the greatest economic gain (Abbott *et al.* 2001). Many of the systems currently under development rely on multiple sensors and an array of techniques, ranging from simple thresholds to neural networks. Although defect detection methods for use with softwoods are also being developed, research has focused on veneers and other more valuable and highly processed wood products (Butler *et al.* 1989; Butler *et al.* 2002).

---

[15] Saw tear-out is a condition in which the saw blade rips the grain on the surface of a workpiece. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

When automated measurement devices, such as non-contact laser range sensors are used to obtain data for real-time SPC, it is necessary to know the location of defects; non-sawing defects need to be removed from SPC data, while sawing defects should be retained and noted. A defect detection system for this purpose can be much simpler than that of lumber grading, as it need not involve delineation of defect boundaries in two dimensions. Instead, it can rely on point estimation of wane and other non-sawing defects at the location of measurements taken for SPC.

## *3.2 Research Objective*

The objective of this chapter is to develop a multi-sensor system for identifying sawing and non-sawing defects in the specific region where automated SPC data are collected. Non-sawing defects of interest include anomalies in the board surface profile that could be misinterpreted as sawing defects in automated SPC, such as wane. Sawing defects of interest include saw-caused holes and tear-out, which are indicative of saw quality problems. This research extends methods previously investigated in conjunction with complex grading systems, giving new methods to apply to data from point laser range sensors.

## *3.3 Review of Literature*

Automated defect classification first appeared in the wood products literature in the 1970's. Many techniques were investigated for both data acquisition and data processing. These techniques were developed and evaluated with the goal of improved accuracy and efficiency over human machine operators and graders.

### 3.3.1 Data Acquisition

An array of hardware options, including ultrasound, x-ray radiation, infared, and visible light have been investigated for use in defect detection systems (Szymani 1985). In the 1980's, camera technology dominated in the area of surface defect detection. In the 1990's, logs were scanned for interior defects using computed tomography (CT) and magnetic resonance imaging

(MRI) technologies. Much recent research has focused on multi-sensor systems, incorporating two or more hardware technologies.

Solid-state cameras have been used for measuring surface roughness, for defect detection, and as part of proposed automated grading systems for lumber and veneers. Forrer *et al.* (1988a; 1988b) and Butler *et al.* (1989) used a three-tube colour camera to detect surface defects in Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) veneer. As part of a system for identifying compression wood in Norway spruce (*Picea abies* (L.) Karst.) logs, Andersson and Walter (1995) used images from a digital RGB (red green blue) CCD (charge-coupled device) camera. Also using Norway spruce, Silvén and Kauppinen (1996) described a defect detection system using spectrophotometric measurements from colour imaging. Because colour images were designed with human viewing in mind, they have not proven satisfactory for every application (Brunner *et al.* 1992). Conners *et al.* (1992) used colour cameras as part of a proposed grading system, but recommended including X-ray scanners to overcome knot detection and other classification difficulties, and laser-based ranging cameras to detect changes in thickness.

Black and white camera technologies have also been used for surface defect detection. Koivo and Kim (1989) developed a defect detection system for planed oak (*Quercus rubra* L.) boards using grey level computer images. Defect identification for cork (*Quercus suber* L.) planks was investigated by Gonzalez-Andrados *et al.* (2000) using digitized black and white images. A review by Masi (2003) recommended a black and white smart camera, as colour camera data were shown to be overly affected by natural colour variation in wood.

Schmoldt *et al.* (1997) and Coates *et al.* (1998) developed systems for identifying internal defects in oak (*Quercus* spp.) logs using a CT scanner and MRI, respectively. Conners *et al.* (1997) included X-ray scanners as part of a multi-sensor automatic grading system for hardwood lumber. In subsequent research, X-ray scanners were rejected in favour of RGB line scan colour

cameras for identifying compression wood in green yellow pine (*Pinus spp.*) (Nyström and Kline 2000).

Many automatic hardwood lumber grading systems that are currently under development rely on a multi-sensor approach. For example, a system was developed where thickness irregularities, holes, and cracks were identified using a structured light profiling system and triangulation algorithm, colour cameras were used for defect classification related to discoloration, and X-rays were used to identify defects related to density differences (Conners *et al.* 1997; Abbott *et al.* 2001; Kline *et al.* 2001; Lee *et al.* 2001).

### 3.3.2 Algorithms

Developing a robust computer vision algorithm was a difficult stage of early research. First, regardless of the technology used, large amounts of data result from scanning images, leading to data storage difficulties. The computational complexity of the algorithms and the spatial resolution of the images had to be reduced to achieve real-time processing. Second, the hardware is sensitive to lighting, vibration, and airborne dust. Third, the wood material itself is highly variable in terms of colour, weathering, the amount of dirt present, moisture content, sap stain, blue stain, and other marks. These issues led Conners *et al.* (1992) to describe the development of these algorithms as "an art and not a science".

Many methods have been used for computer vision algorithm development. These algorithms have evolved to include several fairly consistent steps (Pham and Alcock 1998):

1. Image subdivision: classifying parts of the image into different types, which could involve determining the regions of the image containing wood or simply dividing the image into smaller, more manageable parts;

2. Image enhancement: smoothing and other noise reducing strategies;

3. Feature extraction: classifying the regions of wood (e.g., clear versus suspicious wood); and

4. Classification: categorizing suspicious areas into various kinds of defects.

Depending on the method, corrections to the classifications are then made manually or automatically as an additional step.

Image subdivision has been accomplished in a variety of ways. Forrer *et al.* (1988a; 1988b) and Butler *et al.* (1989) investigated several types of image sweep-and-mark algorithms for surface defect detection, while Koivo and Kim (1989) used simple statistics and time series modeling. Simple thresholding of pixel values was used as part of a system for identifying compression wood by Andersson and Walter (1995) and in conjunction with MRI images by Coates *et al.* (1998). Several studies investigated an image subdivision algorithm that quantified the surface orientation and estimated surface curvature (e.g., Conners *et al.* 1997; Abbott *et al.* 2001; Kline *et al.* 2001; Lee *et al.* 2001).

Image enhancement has been accomplished with simple methods, such as the median filter (Andersson and Walter 1995) and mean filter (Coates *et al.* 1998). As part of their system for identifying compression wood, Nyström and Kline (2000) used shade correction.

Some authors described a "segmentation system", which combined the image subdivision and feature extraction steps. For example, a segmentation system for several species of kiln-dried lumber (e.g., oak (*Quercus rubra* spp.), cherry (*Prunus virginiana* L.), and white pine (*Pinus strobus* L.)) was developed using a simple histogram-based thresholding method (Conners *et al.* 1992). Other researchers used the probability distributions derived from spectra curves of known defects to discriminate sound wood from questionable wood (Silvén and Kauppinen 1996). Image segmentation has also been performed with weighted thresholds assuming that the pixels in gray-scale images follow a bi-modal distribution (Otsu 1979, as cited in Schmoldt *et al.* 1997).

Feature extraction and classification have also been performed as one step. Koivo and Kim (1989) used tree classifiers and discriminant analysis. Andersson and Walter (1995) investigated a "supervised" maximum-likelihood classification system, which required operator intervention. Artificial neural networks within a small 3-D window were used for extraction and classification by Schmoldt *et al.* (1997), with a morphological post-processing step to re-classify anomalies.

The classification step has also been performed in a number of ways. Conners *et al.* (1992) developed a rule-based neural networks system that used fuzzy logic for defect classification. Silvén and Kauppinen (1996) used a stepwise system of rules. Defect identification for cork planks was investigated using a stepwise discriminant analysis (Gonzalez-Andrados *et al.* 2000). Multivariate image projections to latent structures was used to transform the RGB colour space into six non-linear effects that were then used as inputs to a Bayesian classifier function (Nyström and Kline 2000).

### 3.3.3 System Evaluation

A primary goal of many of the research studies cited was to improve the efficiency of lumber grading. With the current state of computer technology, grading rules are easily translated into computer programs that use the locations and sizes of particular defects to produce the correct grade. Thus, improved grading efficiency results from being able to correctly identify certain types of defects and to accurately locate and size them in real time. "Confusion matrices" are used to describe a method's success rate, showing the predicted and actual classifications for each type of defect.

One limitation of many of the cited studies is that confusion matrices were not always shown, and reported error rates were often based on anecdotal information or from single datasets without cross-validation (Schmoldt *et al.* 2001). Error rates between studies are also not comparable, in that successful detection can be defined in terms of a simple binary choice (defect

detected or defect not detected) or in terms of the proportion of the board correctly classified. Different levels of resolution can further complicate this issue and make comparisons difficult.

Estimates of the accuracy of human graders also vary by product and geographic area. Huber *et al.* (1985) reported an accuracy rate of 68% for graders in a red oak rough mill. Silvén *et al.* (2003) asserted that grading accuracy rates were rarely above 70% regardless of product.

Silvén and Kauppinen (1996) reported an error rate of 3% for wane and 12% for other types of defects, but results were based on a single dataset. Schmoldt *et al.* (1997) reported a 5% confusion rate using cross-validation. Their later work used additional species and methods, and gave error rates of 3-10% (Schmoldt *et al.* 2001). Using a single sample as a training board and 16 new sample boards, Nyström and Kline (2000) reported an error rate of 10%. Using a validation dataset, Gonzalez-Andrados *et al.* (2000) reported a confusion rate of 33%; however, error rates for human graders were approximately 50% for the product under study.

Although efficiency gains are reported by many researchers, automated grading systems are not yet common in commercial sawmills.

## 3.4 Methods and Materials

### 3.4.1 Materials

Weyerhaeuser's New Westminster sawmill (British Columbia, Canada) supplied 110 pieces of rough green western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) lumber for sampling. The lumber was 51.5 mm × 135 mm ($2\frac{1}{32} \times 5\frac{5}{16}$ inches) and processed by several different sawing machines: approximately half the lumber was bandsawn, one quarter was circular-sawn, and one quarter was chipped (sawn by a chipper-head). This produced lumber of varying quality with many sawing and non-sawing defects; however, most pieces were defect-free. In order to adequately represent areas of wane and tear-out, the 110 sample boards were stratified by the

amount of each defect present, and 30 boards were selected to represent a range of quality. To follow on previous results, the six boards analyzed in Chapter 2 were included in the 30 boards.

The lumber was taken immediately after sawing to the Q-Lab at the Department of Wood Science in the Forest Sciences Centre at The University of British Columbia (Vancouver, Canada), and cut into 2.44-metre (8-foot) lengths. The lumber was measured using a scanning apparatus consisting of a high quality motor-driven carriage, a black and white digital camera fitted with an infrared filtering lens, a laser line, and four Hermary LRS-50 point laser range sensors (Figure 3-1). Two LRSs were mounted vertically on each side of the apparatus so that measurements were taken at 22 mm and 106 mm (⅞ and 4 3/16 inches) above the bottom edge of the lumber. The data acquisition components of this system were chosen because of their relative low cost and record of success in previous research.

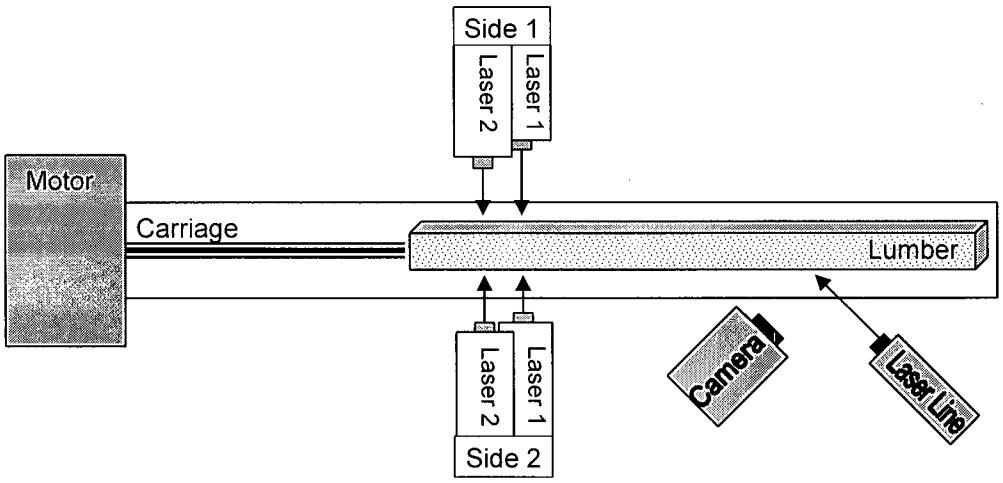



**Figure 3-1. Set up of measurement apparatus in Q-lab, as viewed from above.**

The combination of the camera and laser line is known as a sheet-of-light profile imaging system, and is similar to the profiling system used in Lee *et al.* (2001). With this commonly used system, a plane of light is projected at a 45 degree angle onto the board or cant surface using a laser line. When viewed from an angle perpendicular to the board, the light reflects from the surface as a two-dimensional curve, sometimes called a "laser stripe". The $(x,y)$ coordinates of this curve are captured via the black and white digital camera, which is mounted at an angle of

45 degrees to the board and 90 degrees to the laser line in order to capture the maximum light reflectance from the laser stripe. Figure 3-2 shows the side view of the system.

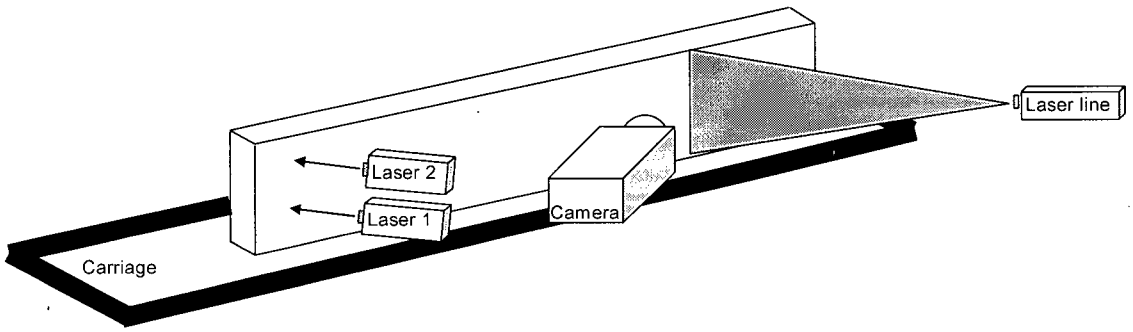



**Figure 3-2. Set-up of two LRSs and sheet-of-light profiling system (side view).**

The motor speed of the carriage was set to approximate the processing speed in a typical lumber mill. As the carriage moved the lumber past the camera at constant speed, pictures were taken at a rate of 47 pictures per second (approximately one picture every 8 mm, or three pictures per inch). An encoder connected to the carriage motor recorded the position of each LRS measurement taken along the length of the board. In order to eliminate the possibility of the lumber moving in the apparatus during the scan, the lumber was secured to the carriage by three clamps at 15, 122, and 129 cm (6 inches, 4 feet, and 7 ½ feet) along the board's length.

Lumber was scanned twice. After scanning, information about the sawn surface was recorded. Surface anomalies that were in the path of the LRSs (i.e., at 22 and 106 mm above the bottom of the board) were mapped in detail. These data included:

1. saw type for each side (Bandsaw, Circular Saw, or Chipper-head);
2. the direction of sawing;
3. obvious saw mark patterns;
4. discolourations, such as smooth knots, pitch pockets, and stains;
5. sawing defects, such as tear-out, skip, or step; and

6. non-sawing defects, such as jagged knots, holes, and wane.

A detailed record was then made for each side of each piece of lumber at each laser location. For each 6 mm (¼ inch) section of the length, the board was classified as defect-free, sawing defect (saw-caused hole/tear), or non-sawing defect (natural hole or wane).

### 3.4.2 Data

The camera data consisted of an intensity value at each pixel on a 640 × 480 grid. At a rate of 47 pictures per second, the ASCII files generated from this system were enormous. Using an infrared filter reduced the file sizes considerably; however, the volume of data was still too large to accommodate real-time processing. Thus, the camera data were filtered using a threshold intensity value of 160 at each pixel. This value was selected based on trial and error and ensured that the datasets generated were adequate for image processing, but of manageable size.

Sample filtered data from a single frame of defect-free lumber and from a board with wane are shown in Figure 3-3 . Before plotting, the vertical coordinate of each data point was transformed to correspond to a distance (in cm) along the width of the board. Because the camera makes a two-dimensional image from the three-dimensional laser line projection on the lumber, the pixel locations are not an exact representation of the laser line. This effect, which is known as perspective projection or parallax, is symmetric with respect to the middle of the image. Therefore, pixel locations may be slightly biased, but consistently so, and thus, no adjustments to the data were made to account for parallax.

Although LRS data were collected simultaneously with the camera data, the LRS data were not incorporated into the proposed algorithms for defect recognition. There were two main reasons for this. First of all, there was a large difference in data acquisition speeds between the LRSs and the digital camera; while the camera took approximately 1.2 pictures per cm, the LRSs took approximately 12 measurements per cm. This made matching the frame to the LRS

measurements inaccurate at best. Moreover, given the large datasets generated by the camera, using a single sensor led to simpler, less computationally expensive methods.

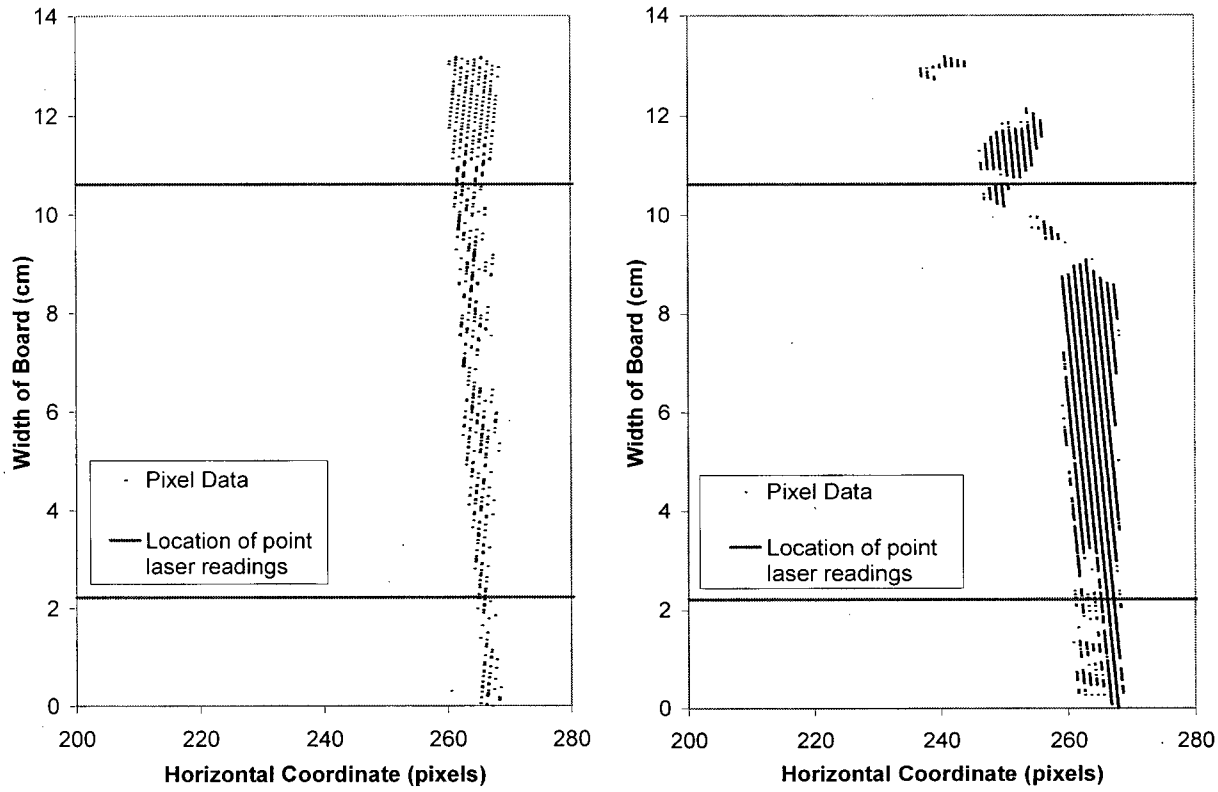



**Figure 3-3. Pixel coordinate locations from digital camera for single frame of defect-free lumber (left) and lumber with wane (right).**

### 3.4.3 Algorithms

Two methods for identifying defects in the scan region were developed and evaluated. The first method was a rule-based method; the second used discriminant analysis. Both methods met the goal of computing simplicity and were supported by previous research (e.g., Koivo and Kim 1989; Conners *et al.* 1992; Silvén and Kauppinen 1996; Gonzalez-Andrados *et al.* 2000). Computational simplicity was particularly important, as the algorithm must be capable of performing as part of a real-time system. For both methods, image enhancement of the filtered camera data was performed by frame via a sequential pre-processing algorithm.

### 3.4.3.1 Image Enhancement Pre-processing Algorithm

The image enhancement pre-processing algorithm consisted of several steps, involving data cleaning, curve fitting, and region delineation. Each frame was analyzed independently and in sequence.

#### *3.4.3.1.1 Data Cleaning*

A data cleaning procedure was necessary to eliminate frames that corresponded to pictures taken when the clamps (at 15, 122, and 129 cm) were in frame or partially in frame. The steps in the data cleaning procedure were as follows:

1. If the frame was within ± 25 mm (± 1 inch) of the clamp, then the frame was labeled as "no information".
2. Data were filtered to a set range of $(x,y)$ coordinates. For the $y$-coordinates, the range corresponded to the board width dimensions ± 5 pixels (~1.7 mm or 0.065 inch). For the $x$-coordinates, a range covering +10 pixels and -70 pixels (~3.3 and 18 mm, or 0.13 and 0.91 inch) of the average was chosen to ensure coverage of wane areas.
3. The number of pixels remaining in the frame after Step 2 was counted, and the range of the $y$-coordinates was measured. If the range was less than 100 pixels (~33 mm or 1.28 inches) or the number of pixels in the frame was less than 300, the frame was discarded.

These values were judgmentally chosen based on pilot tests of boards with wane, tear-out, and other defects, and ensured that all defect areas were included in the sample data.

#### *3.4.3.1.2 Curve Fitting*

Curve fitting was used to translate the pixel "cloud" into a single line of data. Several methods of curve fitting were investigated, including higher order polynomial regressions and cubic splines. Since higher order polynomials can be unstable, the latter technique was chosen.

In order to perform meaningful spline fitting of data oriented along the $y$-axis, the data must be one-to-one with respect to the $y$-axis; for every $y$ coordinate value, there must be a unique $x$ coordinate value. However, since the laser line was not a perfectly thin line, the pixel data from the camera were spread along the $x$-axis, covering five or more pixel values. Thus, it was necessary to find the leading edge of the $x$ pixel values. Furthermore, in order to facilitate later processing using subsequent frames, the fitted lines were made to correspond to uniform pixel values by delineating them into groups.

The steps for pixel grouping and leading edge determination were as follows:

1. For each possible $y$ pixel value, the greatest $x$ pixel (the rightmost) value was found.
2. The $y$ pixels were assigned to 8-pixel groups. This group size was chosen to correspond to a distance along the board of approximately 2.5mm (0.10 inch), which was arbitrarily determined to give an adequate number of points for subsequent curve fitting.
3. For each 8-pixel $y$ group, the average of the greatest $x$ pixel value from Step 1 was determined, and called the "leading edge".

The steps for curve fitting were as follows:

1. Cubic splines were fit to the leading edge data determined in Step 3 using the SAS procedure PROC EXPAND[16]. Splines were fit using the horizontal coordinates ($x$) as the dependent variable. This procedure estimated the coefficients of the fitted splines and predicted values of the splines where data were missing.
2. Using the coefficients from the splines, first derivatives along the horizontal axis, $x'$, were calculated with respect to the vertical axis ($y$). The first derivative at the exact location of the two lasers was interpolated using the two nearest points.

[16] All SAS procedures referred to in this chapter were run with Version 8.2 (SAS Institute 2002).

### *3.4.3.1.3 Region Delineation*

For subsequent calculations, each frame was delineated into several regions along the width of the board. The area above the top laser (at 106 mm) was labeled as ROI1, the region between the two lasers was labeled as ROI2, and the region below the bottom laser (at 22 mm) was labeled as ROI3. Some calculations were made only using points within a small "sensitivity" amount of the lasers. The area within 6.4mm (¼ inch) of the top laser (100-112 mm, or 3 15/16 - 4 7/16 inches) was labeled S1; the area within 6.4mm (¼ inch) of the bottom laser (16-28 mm, or ⅝-1⅛ inches) was labeled S2. The area within 12.8 mm (½ inch) of the top laser (93-119 mm, or 3 11/16 - 4 11/16 inches) was labeled $S1^2$; the area within 12.8 mm (½ inch) of the bottom laser (9-35 mm, or ⅜- 1⅜ inches) was labeled $S2^2$.

Figure 3-4 shows a frame that has undergone the image enhancement pre-processing steps. The pixel values are shown as dots, the leading edge determined for the board is shown as a fitted curve, and the regions are labeled.

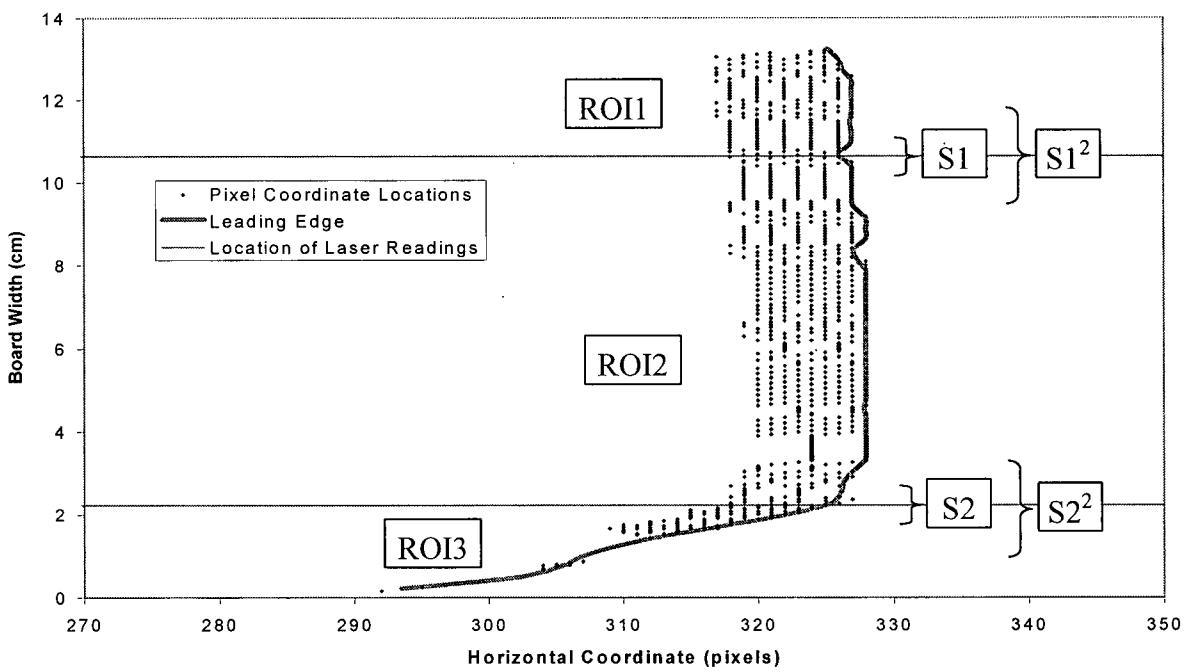



**Figure 3-4. Example of leading edge calculation, curve fitting, and region delineation for a board with wane (ROI1=region of board above top laser, ROI2=region between lasers, ROI3=region below bottom laser, S1=region within 6.4 mm of top laser, S2=region within 6.4 mm of bottom laser, $S1^2$=region within 12.8 mm of top laser, $S2^2$=region within 12.8 mm of bottom laser).**

### 3.4.3.2 Rule-based Method

The rule-based method consisted of an image subdivision step, a feature extraction step, and a defect classification step.

#### *3.4.3.2.1 Image Subdivision*

Image subdivision involved a series of sequential steps performed independently. For each frame, statistics were computed that best differentiated defective areas from non-defective. First, missing pixels indicated that the laser stripe was fully absorbed by the lumber, or that the wood was missing. Thus, pixels were counted by region and in the sensitivity areas by:

1. The number of points missing in each of ROI1, ROI2, and ROI3 was counted: $num\_miss_1$, $num\_miss_2$, and $num\_miss_3$.
2. The number of points missing in each of S1 and S2 was counted: $num\_miss_{S1}$ and $num\_miss_{S2}$.

The curvature indicated the shape of the curve made by the laser stripe. Thus, the angle of the stripe and the first derivative of the curve fitted to the stripe were found in the sensitivity areas. Using trial and error, an angle of greater than ten degrees was found to indicate unusual surface patterns. The steps were therefore:

3. The number of points in each of S1 and S2 where $\arctan(x') > 10$ degrees was counted: $num\_angled_{S1}$ and $num\_angled_{S2}$.
4. The average first derivative along the horizontal axis with respect to the vertical axis (from fitted splines: $x'$) in each of S1 and S2 was computed: $avg\_xprime_{S1}$ and $avg\_xprime_{S2}$.

If a large number of pixels was missing, then full absorption of the laser light could have occurred and classification was not possible. These frames were labeled as follows:

5. If the number of missing points in ROI2 was large, then the frame was ignored, i.e.:

   If $num\_miss_2 > 30$, then the top and bottom laser were labeled with "no information".

The number of missing pixels and the number of angled pixels was an indication of possible defects. Frames where defects were suspected were labeled as follows:

6. If at least two points in S1 were angled and/or missing, then a defect was suspected and the frame was marked for further processing:

   a. If $num\_miss_{S1} + num\_angled_{S1} \geq 2$, then the frame was marked for further processing of defects for the top laser; and

   b. If $num\_miss_{S2} + num\_angled_{S2} \geq 2$, then the frame was marked for further processing of defects for the bottom laser.

Finally, the first derivative was used to differentiate knots, holes, and wane using a threshold of ± 0.1, which was based on trial and error. Suspected defects marked for further processing in Step 6 were labeled as follows:

7. Frames marked in step 6 were labeled as suspected wane or suspected knots/holes:

   a. If $avg\_xprime_{S1} > -0.1$, then *suspected_defect_top* = knot/hole;

   b. If $avg\_xprime_{S1} \leq -0.1$, then *suspected_defect_top* = wane;

   c. If $avg\_xprime_{S2} > +0.1$, then *suspected_defect_bottom* = knot/hole; and

   d. If $avg\_xprime_{S2} \leq +0.1$, then *suspected_defect_bottom* = wane.

#### *3.4.3.2.2 Feature Extraction*

The feature extraction steps used information from previous and subsequent frames to label frames as suspected defects. It also ensured that the first and last frames, as well as frames near the clamps were not corrupted by reflectance from the hardware located in those areas of the

carriage. Again, it was a rule-based algorithm, which cycled through each frame in order along the board. First, frames near clamps were checked to ensure that they were not incorrectly labeled as defects:

1. If a suspected defect was preceded by one frame that contained no suspected defects and was followed by two frames recorded as "no information", then the suspected defect was deleted.
2. If a suspected defect was preceded by two frames recorded as "no information" and was followed by one frame that contained no suspected defects, then the suspected defect was deleted.

Next, first and last frames for each board were checked and re-labeled to ensure that defect types were consistent:

3. If a suspected defect was recorded on the first frame and was followed by two non-suspected frames, then the suspected defect was deleted.
4. If a suspected defect was recorded on the last frame and was preceded by two non-suspected frames, then the suspected defect was deleted.
5. If the first frame was recorded as "no information" and was followed by at least two identical suspected defects, then the first frame was labeled with the same defect as the following frames.
6. If the last frame was recorded as "no information" and was preceded by at least two identical suspected defects, then the last frame was labeled with the same defect as the preceding frames.

Based on trial and error, it was found that most "singletons" were incorrectly labeled. That is, defect frames that were both preceded and followed by non-defects were not defects (and vice-versa). In the last steps, "singletons" were re-labeled as follows:

7. If a suspected defect was both preceded and followed by two non-suspected or "no information" frames, then the frame was assumed to be an area of discolouration (e.g., a closed knot) and labeled as not defective.

8. If a non-suspected frame was preceded and followed by two suspected defect frames, then the frame was labeled with the suspected defect of the frame preceding it.

#### *3.4.3.2.3 Defect Classification*

The defect classification steps for the rule-based method used information from previous and subsequent frames to classify the suspected defects into three categories: non-defective (including closed knots), holes (including sawing defects and tear-out), and wane (including all non-sawing defects). Several assumptions were made based on basic knowledge about defects. Suspected knot/holes lasting two or more frames were classified as holes. Subsequent frames that were not consistently labeled with the same defect type were labeled with the majority defect type. Again, it was a rule-based system, which cycled through each frame in order along the board. First, "doubleton" suspected knot/holes were labeled as holes:

1. If two consecutive suspected knot/hole frames were preceded and followed by at least one non-suspected or "no information" frame, then they were labeled as holes.

Next, two or more consecutive suspected defect frames were labeled with the majority defect:

2. If three consecutive frames were labeled as suspected defects, then they were all re-labeled with the majority defect. For instance, if two of three consecutive frames were labeled as "knot/hole", then all three were re-labeled as holes.

3. Step 2 was repeated to ensure that no singletons were left. Even numbers of consecutive defects were labeled using the first two out of three defect label.

At the end of these steps, there could have been suspected defects that were not classified. For instance, two consecutive frames labeled as "wane" and "knot/hole" could have been left unclassified. These were assumed to be knots or other surface anomalies not caused by sawing.

### 3.4.3.3 Discriminant Analysis Method

Both image segmentation and defect classification were accomplished via discriminant analysis. Discriminant analysis is a technique used to determine which variables discriminate between two or more groups via the development of discriminant functions, which are then used to predict group membership (Manly 1994). Linear and quadratic discriminant analyses were performed with the SAS procedure PROC DISCRIM.

The frame level input variables for discriminant analysis were chosen based on prior knowledge of how defects were reflected under the sheet-of-light profiling system. The $x$ pixel value at the point of the LRS data collection was important, as it increased/decreased in value with the thickness of the board. The first derivative of the leading edge fitted curve at the LRS was also important, as it indicated the curvature of the fitted curve at the point where data were collected.

Other variables describing the variability and paucity of the data around the LRS were also considered important. These included the variance of the $x$ pixels and the number of missing values above and below the LRS. Because only defects that occurred in the region where the LRSs collected data were of interest, these variables were calculated for each frame using only points in regions $S1^2$ and $S2^2$ (i.e., within 12.8 mm (½ inch) of each of the two LRSs).

The trend of the leading edge fitted curve around the LRS with respect to the horizontal ($x$) direction was also considered. For each frame, a linear regression using $x$ as the dependent variable was fitted for the pixel coordinates in each of $S1^2$ and $S2^2$. This generated a slope and

an intercept for each LRS area. The fit of this regression line was also considered a valuable statistic, and was measured using the root mean square error (RMSE). In total, eight variables were included from each frame for each of the top and bottom lasers:

1. $x$ pixel value at LRS;
2. $x'$ (first derivative along horizontal direction ($x$) with respect to vertical direction ($y$)) at LRS;
3. Variance of $x$ pixel values in each of sections $S1^2$ and $S2^2$;
4. Number of missing values above LRS ($num\_miss_1$ for top LRS, $num\_miss_2$ for bottom LRS);
5. Number of missing values below LRS ($num\_miss_2$ for top LRS, $num\_miss_3$ for bottom LRS);
6. Intercept of $x$ pixel versus $y$ pixel regression line, with $x$ as the dependent variable;
7. Slope of $x$ pixel versus $y$ pixel regression line, with $x$ as the dependent variable; and
8. Root mean square error (RMSE) of $x$ pixel versus $y$ pixel regression line, with $x$ as the dependent variable.

In addition, the values of each of these variables from the previous frame were also used as independent variables in the discriminant analysis. Therefore, sixteen independent variables in total were considered in the discriminant analysis.

In order to perform linear discriminant analysis, several assumptions must be met: (1) for each observation of the dependent variable, the observations from the independent variables are independent of all others, (2) the data are a sample from a multivariate normal distribution, and (3) the data are homoscedastic (i.e., the variance/covariance matrices are homogeneous across groups) (Manly 1994). If assumption (1) is not met, the resulting discriminant functions cannot

be tested for statistical significance; however, the discriminant analysis can still be used in a descriptive sense. If assumptions (2) or (3) are not met, the correct classification rates from the discriminant functions decline. Whereas non-normality problems can be avoided with transformations, problems with heteroscedastic data can be remedied by performing a quadratic discriminant analysis. The discriminant functions derived with linear discriminant analysis use a covariance matrix pooled over all groups, while quadratic discriminant analysis is performed so that each group uses its own covariance matrix (Tabachnick and Fidell 1983).

Since the frame data were sequential, the assumption of independent observations was violated and the discriminant functions were not tested. The assumption of multivariate normality is difficult to test; therefore, univariate normality was verified by examining histograms constructed for each variable, and transformations were made as needed. The test for homoscedasticity was performed within PROC DISCRIM using a significance level of 0.10. If the assumption was not met, a quadratic discriminant analysis was performed.

Discriminant analysis results in the minimum of ($v_1$-1, $v_2$) discriminant functions, where $v_1$ is the number of groups and $v_2$ is the number of independent variables. Using three groups (holes, wane, and no defect), resulted in two discriminant functions. The group membership of a particular frame was indicated by the function that gave the highest value.

One way to ensure unbiased estimates of the proportion of correct classifications from the discriminant functions (the classification rate) is to validate the discriminant functions with data different from that which was used to derive the discriminant functions (Manly 1994); the functions are derived with a training dataset and validated with a validation dataset. Ideally, the training and validation datasets should be uncorrelated, but this would require doubling the number of sample boards scanned. Instead, each sample board was scanned twice, with two

different (arbitrary) starting places along the board. Using the two scans of each board as the training and validation datasets, respectively, may have overstated classification rates.

When all of the assumptions are met, discriminant analysis can be used to find which independent variables are significant predictors in the discriminant functions. Variables can be selected using a backwards stepwise procedure: first all variables are used as inputs to the discriminant analysis, then the variable that makes the smallest contribution to the discriminant functions is dropped until all variables left in the discriminant analysis are significant (Klecka 1980). Since the assumption of independence was violated in this analysis, statistical significance could not be correctly attached to particular variables. However, the discriminant functions using different combinations of the independent variables could be tested to find the group that gave the best validation dataset classification rate. First, all variables were used in the analysis. Then, using a backwards stepwise method, the variable with the lowest discriminant loading was dropped and the discriminant analysis was re-run. This continued until only one variable remained in the analysis. The chosen combination of variables was that which gave the lowest overall classification rate for the validation dataset.

Discriminant analysis can only be performed when all variables used in the discriminant function are measured. Missing values of the variables $x$, $x'$, *slope*, and *intercept* occurred when pixel values were missing in the sensitivity region. In this situation, an alternative discriminant analysis was run based only on the variables $num\_miss_1$ and $num\_miss_2$.

### 3.4.4 Evaluation

The two algorithms were evaluated by examining confusion matrices for the validation dataset. These matrices list the number of observations by group that are classified into each possible group. An overall classification rate was calculated as a weighted average of the classification rates by group. Confusion matrices were also broken down by saw type and LRS location.

In order to investigate areas of high and low performance, examples of boards with high classification rates and low classification rates were examined in detail. Detailed examinations were also made for the boards used in Chapter 2 in order to show improvements in defect filtering using this camera-based system.

## 3.5 Results

The distribution of actual defects for the 30 samples in the validation dataset is given in Table 3-1 by number of frames and by combined length of defects. The total of the average distance by defect type sums to less than the length of a board (244 cm) because of the removal of clamped areas and defects that occurred after the boards left the sawing machine (e.g., marks from conveyors).

**Table 3-1. Distribution of actual defects in validation dataset.**

| Classification | Number of Frames | Average Distance Per Board (cm) | Proportion of Boards |
|---|---|---|---|
| Defect-free | 13,562 | 199.4 | 87.2% |
| Hole/tear | 704 | 10.4 | 4.5% |
| Wane | 1,295 | 19.1 | 8.3% |
| Total | 15,561 | 228.9 | 100.0% |

### 3.5.1 Rule-based Method

Table 3-2 lists the results for the rule-based method. This method correctly classified the boards in the validation dataset 93% of the time, and classified non-defective regions 98% of the time; however, it was not very accurate at classifying defective wood. The method was particularly poor at classifying holes and tears; most sections of tear-out and holes were incorrectly classified as defect-free wood. This method correctly classified wane areas only 75% of the time, also incorrectly classifying them as clear wood.

**Table 3-2. Summary of rule-based method classifications for validation dataset by defect type.**

| | Number of Frames Classified into Defect | | | | % Correctly |
|---|---|---|---|---|---|
| From Defect | Defect-free | Hole/tear | Wane | Total | Classified |
| Defect-free | 13,300 | 92 | 170 | 13,562 | 98.1% |
| Hole/tear | 376 | 219 | 109 | 704 | 31.1% |
| Wane | 314 | 14 | 967 | 1,295 | 74.7% |
| Total | 13,990 | 325 | 1,246 | 15,561 | 93.1% |

Table 3-3 breaks the classifications down for the top and bottom sections of the lumber. Although the overall classification rates for both sections were very close, wane areas were more accurately classified in the bottom sections (82% versus 68%). This result was traced to a higher proportion of missing pixel values in the upper sections of the digital pictures.

**Table 3-3. Summary of rule-based method classifications for validation dataset by defect type and LRS location.**

| LRS | | Number of Frames Classified into Defect | | | | % Correctly |
|---|---|---|---|---|---|---|
| Location | From Defect | Defect-free | Hole/tear | Wane | Total | Classified |
| Top | Defect-free | 6,728 | 45 | 68 | 6,841 | 98.3% |
| | Hole/tear | 162 | 89 | 30 | 281 | 31.7% |
| | Wane | 208 | 6 | 444 | 658 | 67.5% |
| | Total | 7,098 | 140 | 542 | 7,780 | 93.3% |
| Bottom | Defect-free | 6,572 | 47 | 102 | 6,721 | 97.8% |
| | Hole/tear | 214 | 130 | 79 | 423 | 30.7% |
| | Wane | 106 | 8 | 523 | 637 | 82.1% |
| | Total | 6,892 | 185 | 704 | 7,781 | 92.9% |

Table 3-4 shows the classifications by saw type. Bandsawn boards were more accurately classified than circular-sawn boards or chipped boards, but there were comparatively few boards of this type sampled. For chipped boards versus circular-sawn boards, holes and tear-out were better classified (32% versus 19%) and wane was slightly better classified (75% versus 73%).

**Table 3-4. Summary of rule-based method classifications for validation dataset by defect type and saw type.**

| Saw | | Number of Frames Classified into Defect | | | | % Correctly |
|---|---|---|---|---|---|---|
| Type | From Defect | Defect-free | Hole/tear | Wane | Total | Classified |
| Bandsaw | Defect-free | 994 | 5 | 0 | 999 | 99.5% |
| | Hole/tear | 4 | 14 | 0 | 18 | 77.8% |
| | Wane | 0 | 0 | 22 | 22 | 100.0% |
| | Total | 998 | 19 | 22 | 1,039 | 99.1% |
| Chipper-head | Defect-free | 9,900 | 72 | 142 | 10,114 | 97.9% |
| | Hole/tear | 309 | 187 | 96 | 592 | 31.6% |
| | Wane | 226 | 12 | 721 | 959 | 75.2% |
| | Total | 10,435 | 271 | 959 | 11,665 | 92.7% |
| Circular Saw | Defect-free | 2,412 | 15 | 28 | 2,455 | 98.2% |
| | Hole/tear | 63 | 18 | 13 | 94 | 19.1% |
| | Wane | 82 | 2 | 224 | 308 | 72.7% |
| | Total | 2,557 | 35 | 265 | 2,857 | 92.9% |

Individual boards with high proportions of correct classifications and high proportions of incorrect classifications were further examined. Table 3-5 shows the top five boards in each of these categories, along with their saw type, the number of defective sections, and the average

length of defect. The boards with the most correctly classified sections tended to have few numbers of defects, and the average defect length was long. On the other hand, the boards with the most incorrectly classified sections tended to have many defects that were mostly shorter in length. Upon examining the detailed scan notes, it was also found that many of the incorrectly classified boards had additional areas of wane that were above (or below) the scan zone.

**Table 3-5. Top five correctly classified and top five incorrectly classified samples in validation dataset using rule-based method.**

| Sample | LRS Location | Saw Type | % Correctly Classified | Number of Defects | | Avg. Length of Defect (cm) | |
|---|---|---|---|---|---|---|---|
| | | | | Hole/tear | Wane | Hole/tear | Wane |
| 038 | Bottom | Chipper-head | 100.0 | 0 | 1 | | 15.2 |
| 002* | Bottom | Bandsaw | 100.0 | 0 | 0 | | |
| 093 | Top | Bandsaw | 100.0 | 0 | 1 | | 17.8 |
| 001* | Top | Circular Saw | 99.6 | 0 | 0 | | |
| 067 | Bottom | Chipper-head | 99.6 | 0 | 1 | | 47.0 |
| 003* | Bottom | Chipper-head | 82.4 | 4 | 1 | 5.8 | 15.2 |
| 102 | Top | Chipper-head | 83.1 | 0 | 2 | | 24.9 |
| 038 | Top | Chipper-head | 83.6 | 1 | 3 | 2.5 | 17.0 |
| 051 | Top | Chipper-head | 83.9 | 1 | 3 | 15.2 | 14.0 |
| 034 | Bottom | Chipper-head | 85.1 | 1 | 3 | | 55.1 |

* Denotes samples from Chapter 2

### 3.5.2 Discriminant Analysis Method

Discriminant functions were obtained separately for the top and bottom LRS areas because the curvature of the reflected laser line from the bottom of the board was not directly comparable to that of the top of the board. Univariate histograms of each independent variable did not reveal any obvious non-normality, and therefore the second assumption was assumed to be met. The test for homoscedasticity, on the other hand was rejected, and thus a quadratic discriminant analysis was performed.

The discriminant functions resulting from the backward stepwise method for both top and bottom LRS areas retained none of the variables associated with the previous frames. For the bottom LRS area, all of the variables were retained except for the number of $x$ pixel values missing above the LRS ($num\_miss_2$) and the value of $x'$ at the LRS. For the top LRS, only three variables were retained: the number of missing values above the LRS ($num\_miss_1$), and the slope

and RMSE of the regression line with $x$ as the dependent variable. Using these variables produced the lowest error classification rates for the validation dataset.

Table 3-6 shows the confusion matrix for the discriminant method. Overall, this method was slightly less accurate than the rule-based method, correctly classifying 92% of the lumber in the validation dataset. On the other hand, the discriminant analysis method was slightly better at classifying defects: areas of holes and tear-out were classified with an accuracy rate of 33% and areas of wane were correctly classified 77% of the time.

**Table 3-6. Summary of discriminant method classifications for validation dataset by defect type.**

| | Number of Frames Classified into Defect | | | | % Correctly |
|---|---|---|---|---|---|
| From Defect | Defect-free | Hole/tear | Wane | Total | Classified |
| Defect-free | 13,055 | 278 | 229 | 13,562 | 96.3% |
| Hole/tear | 391 | 235 | 78 | 704 | 33.4% |
| Wane | 192 | 102 | 1,001 | 1,295 | 77.3% |
| Total | 13,638 | 615 | 1,308 | 15,561 | 91.8% |

Table 3-7 shows classification rates broken down for the top and bottom sections. As in the first method, the discriminant analysis method performed similarly in the top and bottom LRS areas for defect-free areas, but defects were better identified in the bottom section: classification rates were 39% versus 25% for holes and tears, 80% versus 74% for wane. Again, this result was traced to a higher proportion of missing pixel values in the upper sections of the digital pictures.

**Table 3-7. Summary of discriminant method classifications for validation dataset by defect type and LRS location.**

| LRS | | Number of Frames Classified into Defect | | | | % Correctly |
|---|---|---|---|---|---|---|
| Location | From Defect | Defect-free | Hole/tear | Wane | Total | Classified |
| Top | Defect-free | 6,601 | 90 | 150 | 6,841 | 96.5% |
| | Hole/tear | 177 | 69 | 35 | 281 | 24.6% |
| | Wane | 96 | 35 | 527 | 658 | 80.1% |
| | Total | 6,874 | 194 | 712 | 7,780 | 92.5% |
| Bottom | Defect-free | 6,454 | 188 | 79 | 6,721 | 96.0% |
| | Hole/tear | 214 | 166 | 43 | 423 | 39.2% |
| | Wane | 96 | 67 | 474 | 637 | 74.4% |
| | Total | 6,764 | 421 | 596 | 7,781 | 91.2% |

Table 3-8 shows the classifications by saw type. The overall defect classification rates for each type of saw were very similar. Defect classification of holes and tear-out was better for circular-

sawn boards versus chipped boards (77% versus 53%), while defect classification of wane was slightly better for chipped boards versus circular-sawn boards (78% versus 76%).

**Table 3-8. Summary of discriminant method classifications for validation dataset by defect type and saw type.**

| Saw Type | From Defect | Number of Frames Classified into Defect | | | | % Correctly Classified |
|---|---|---|---|---|---|---|
| | | Defect-free | Hole/tear | Wane | Total | |
| Bandsaw | Defect-free | 990 | 0 | 3 | 993 | 99.7% |
| | Hole/tear | 4 | 14 | 0 | 18 | 77.8% |
| | Wane | 6 | 5 | 17 | 28 | 60.7% |
| | Total | 1,000 | 19 | 20 | 1,039 | 98.3% |
| Chipper-head | Defect-free | 9,687 | 255 | 172 | 10,114 | 95.8% |
| | Hole/tear | 315 | 206 | 71 | 592 | 53.2% |
| | Wane | 136 | 74 | 749 | 959 | 78.1% |
| | Total | 10,138 | 535 | 992 | 11,665 | 91.2% |
| Circular Saw | Defect-free | 2,378 | 23 | 54 | 2,455 | 96.9% |
| | Hole/tear | 72 | 15 | 7 | 94 | 76.6% |
| | Wane | 50 | 23 | 235 | 308 | 76.3% |
| | Total | 2,500 | 61 | 296 | 2,857 | 92.0% |

Individual boards with high proportions of correct classifications and high proportions of incorrect classifications were further examined. Table 3-9 and 3-10 show the top five boards in each of these categories, along with their saw type, the number of defective sections, and the average length of defect. The boards that were best classified tended to have few numbers of defects; the boards with a high proportion of incorrectly classified sections tended to have many defects and more holes and tear-out. Upon examining the detailed scan notes, it was also found that many of the incorrectly classified boards had wane that was above (or below) the scan zone.

**Table 3-9. Top five correctly and top five incorrectly classified samples in validation dataset using discriminant method.**

| Sample | LRS Location | Saw Type | % Correctly Classified | Number of Defects | | Avg. Length of Defect (cm) | |
|---|---|---|---|---|---|---|---|
| | | | | Hole/tear | Wane | Hole/tear | Wane |
| 002* | Bottom | Bandsaw | 100.0 | 0 | 0 | | |
| 074 | Top | Chipper-head | 100.0 | 0 | 1 | | 25.4 |
| 077 | Top | Circular Saw | 100.0 | 0 | 1 | | 16.5 |
| 001* | Top | Circular Saw | 99.6 | 0 | 0 | | |
| 038 | Bottom | Chipper-head | 99.6 | 0 | 1 | | 15.2 |
| 034 | Bottom | Chipper-head | 66.0 | 1 | 3 | 2.5 | 55.1 |
| 038 | Top | Chipper-head | 69.0 | 1 | 3 | 2.5 | 17.0 |
| 051 | Top | Chipper-head | 74.9 | 1 | 3 | 15.2 | 14.0 |
| 003* | Bottom | Chipper-head | 75.4 | 4 | 1 | 5.8 | 15.2 |
| 078 | Bottom | Chipper-head | 82.0 | 0 | 4 | | 6.1 |

* Denotes samples from Chapter 2

### 3.5.3 Samples Examined Using Filtering Methods

Of particular interest were the samples from Chapter 2, which were used to investigate filtering methods. The results from both methods for these six boards were examined closely. Under the rule-based method, all six boards except 003 were classified correctly along at least 92% of the board; under the discriminant analysis method, all boards except 003 and 004 were correctly classified along at least 92% of the board (Table 3-10). Boards 001 and 002 were among the top five correctly classified samples for both methods; Board 003 was among the top five incorrectly classified boards for both methods. Boards 003 and 004 were also the most variable samples with multiple sections of tear-out and wane.

**Table 3-10. Classification statistics for six samples from Chapter 2.**

| | | | % Correctly Classified | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | LRS | | Rule-based Method | | | | Discriminant Method | | | |
| Sample | Location | Saw Type | No Defect | Hole/tear | Wane | Total | No Defect | Hole/tear | Wane | Total |
| 001 | Top | Circular Saw | 99.6 | | | 99.6 | 99.6 | | | 99.6 |
| 002 | Bottom | Bandsaw | 100.0 | | | 100.0 | 100.0 | | | 100.0 |
| 003 | Bottom | Chipper-head | 88.7 | 25.7 | 100.0 | 82.4 | 77.4 | 65.7 | 71.4 | 75.5 |
| 004 | Bottom | Chipper-head | 96.3 | 81.8 | 94.3 | 92.0 | 94.0 | 0.0 | 95.7 | 88.9 |
| 005 | Bottom | Bandsaw | 97.9 | 22.2 | | 96.6 | 98.8 | 77.8 | | 97.3 |
| 006 | Top | Chipper-head | 95.3 | | 68.0 | 92.7 | 96.6 | | 8.0 | 92.7 |

The six boards from Chapter 2 were also used to compare the two camera methods with the original unfiltered scan, the desired filtering result (the "manual filter"), and the filtering method recommended in Chapter 2 (Method 3, the "MSD method"). Labeled frames from the two camera methods were matched to the laser scans, and for comparison purposes, areas which were identified as wane were set to missing. Areas identified as holes are shown in light grey for emphasis. In all samples, only the MSD Method from Chapter 2 removed the short-duration anomalous "spikes". In cases where there were no defects present (Boards 001 and 002, shown in Figure 3-5 and Figure 3-6), the camera methods had virtually no impact; under the rule-based method only one small region was removed (in Board 001, at approximately 230 cm) because of incorrect wane identification, and neither method identified any holes. In cases where there were areas of wane (Boards 003, 004, and 006, shown in Figure 3-7, Figure 3-8, and Figure 3-10), the

rule based method tended to more completely identify it. While both methods performed well identifying the tear in Board 005 (Figure 3-9), neither method completely identified the holes in Board 004 (Figure 3-8), and the discriminant method very much over-identified holes in Board 003 (Figure 3-7).

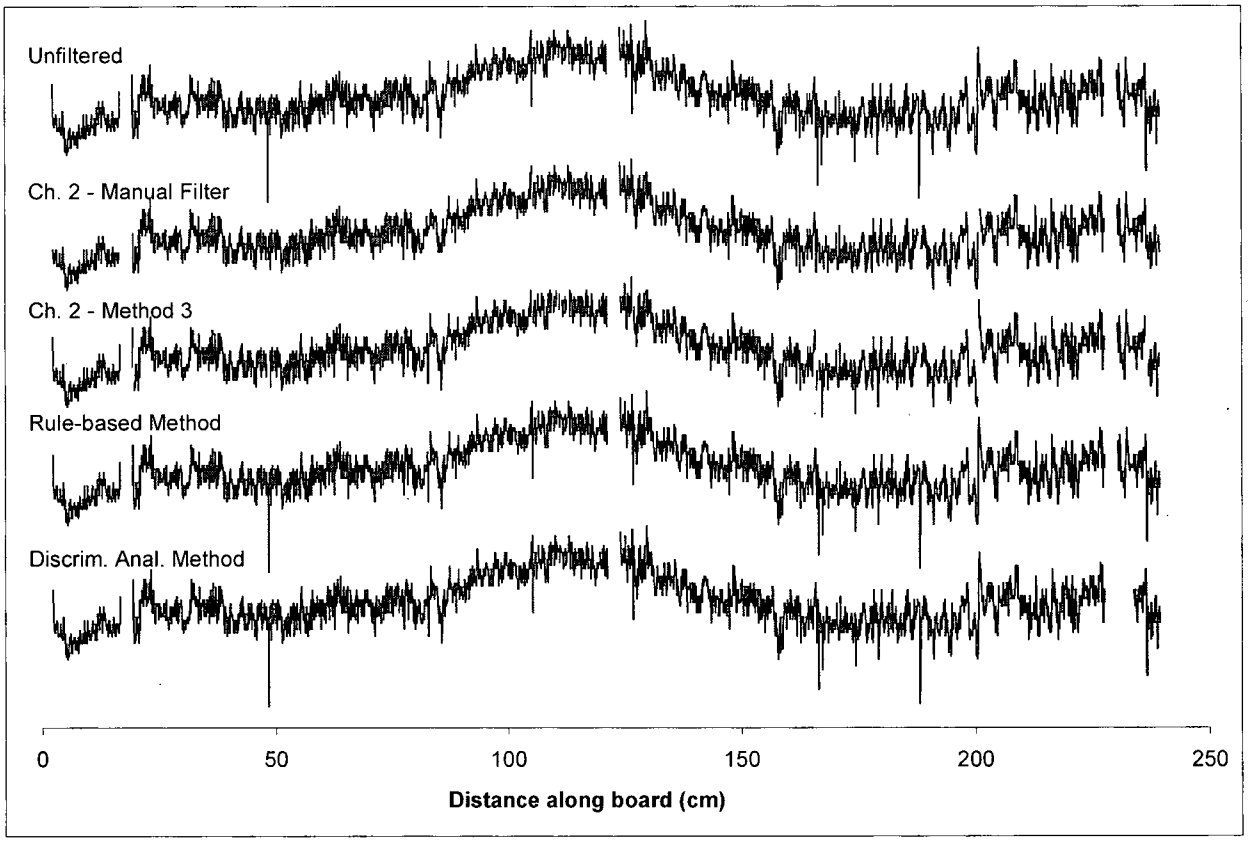



**Figure 3-5. Comparison of Chapter 2 filtering method with camera methods: Board 001[17].**

## 3.6 Discussion

The overall performance of both methods was very similar, averaging about 94% correct classification for the validation dataset. If all data were assumed non-defective, the correct classification rate would be 87%. Thus, the methods gave a 50% gain in classification accuracy over random chance.

[17] Graphs shown in Figure 3-5 through Figure 3-10 have no *y*-axis; that is, each method graphed in the figure has a unique axis (not shown).

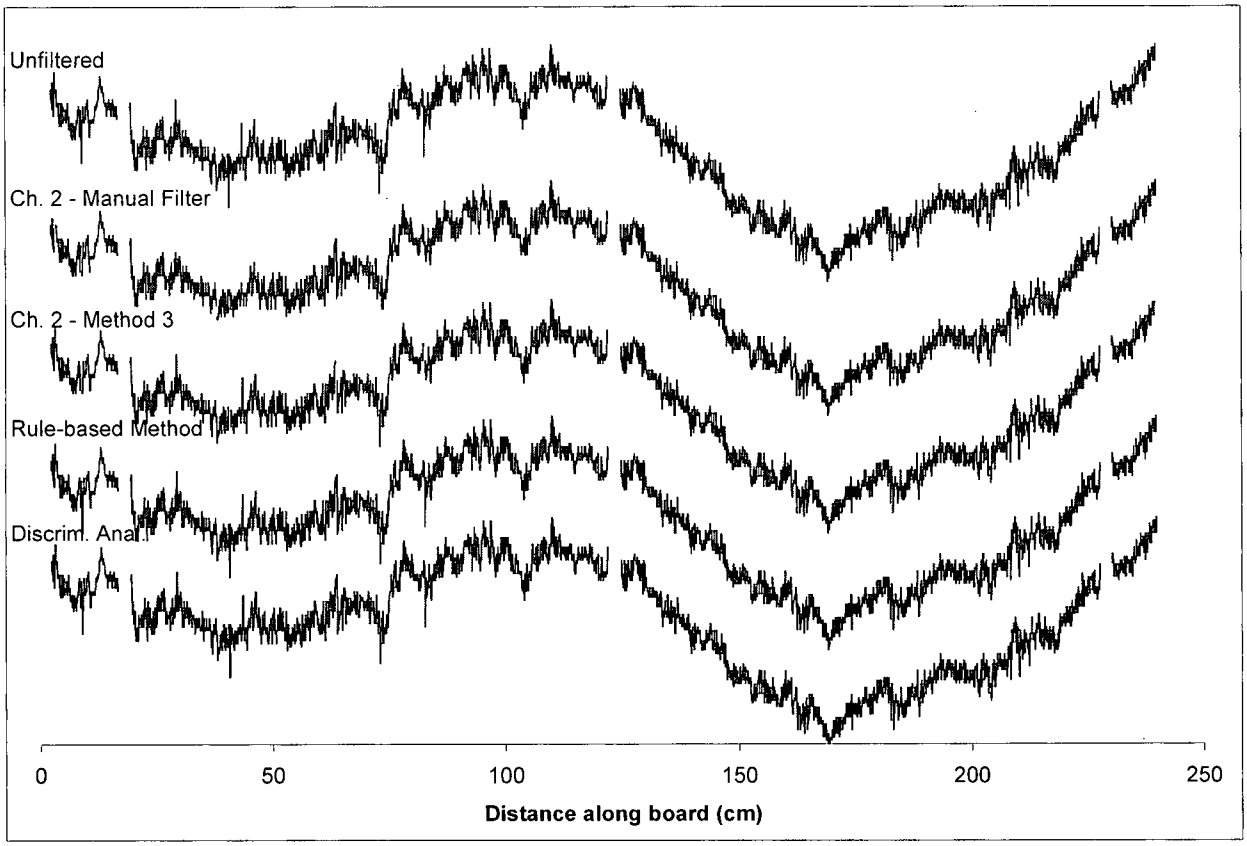



**Figure 3-6. Comparison of Chapter 2 filtering method with camera methods: Board 002.**

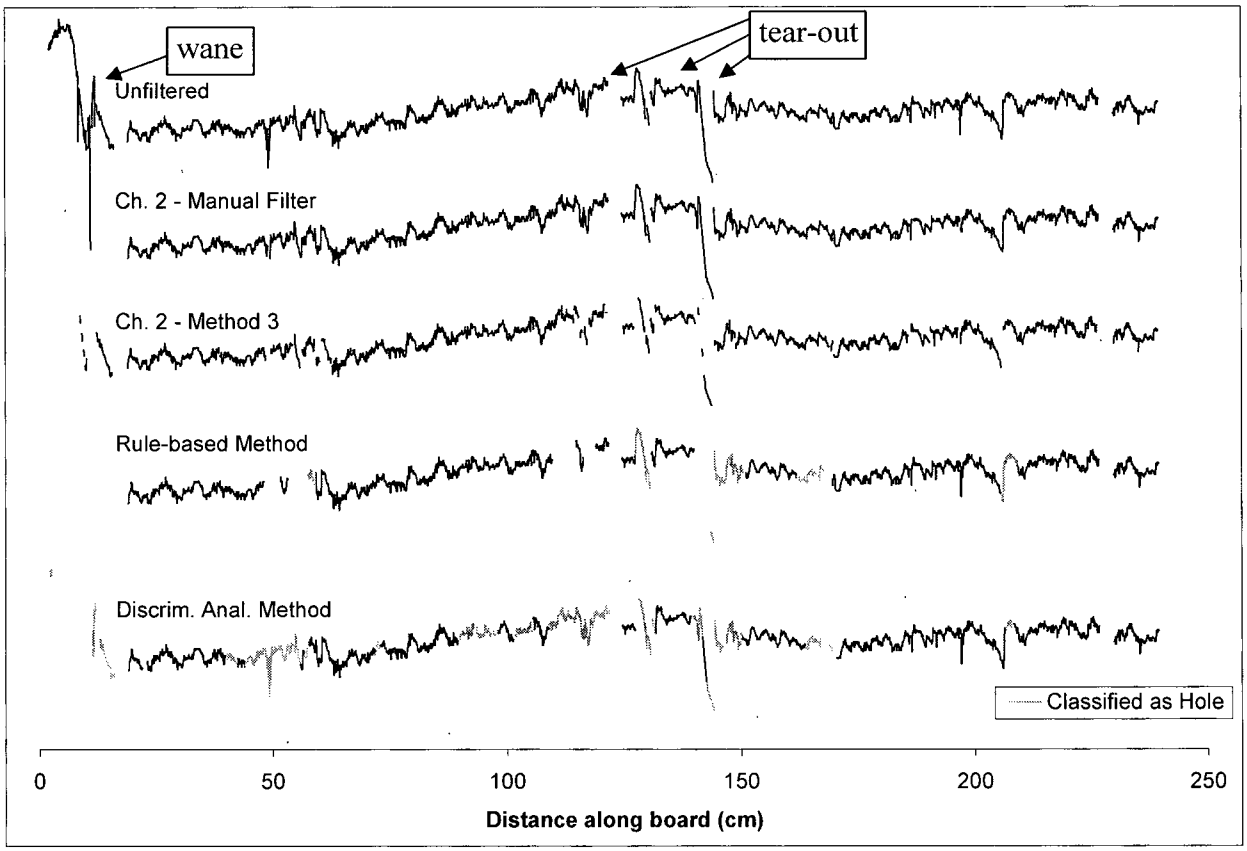



**Figure 3-7. Comparison of Chapter 2 filtering method with camera methods: Board 003.**

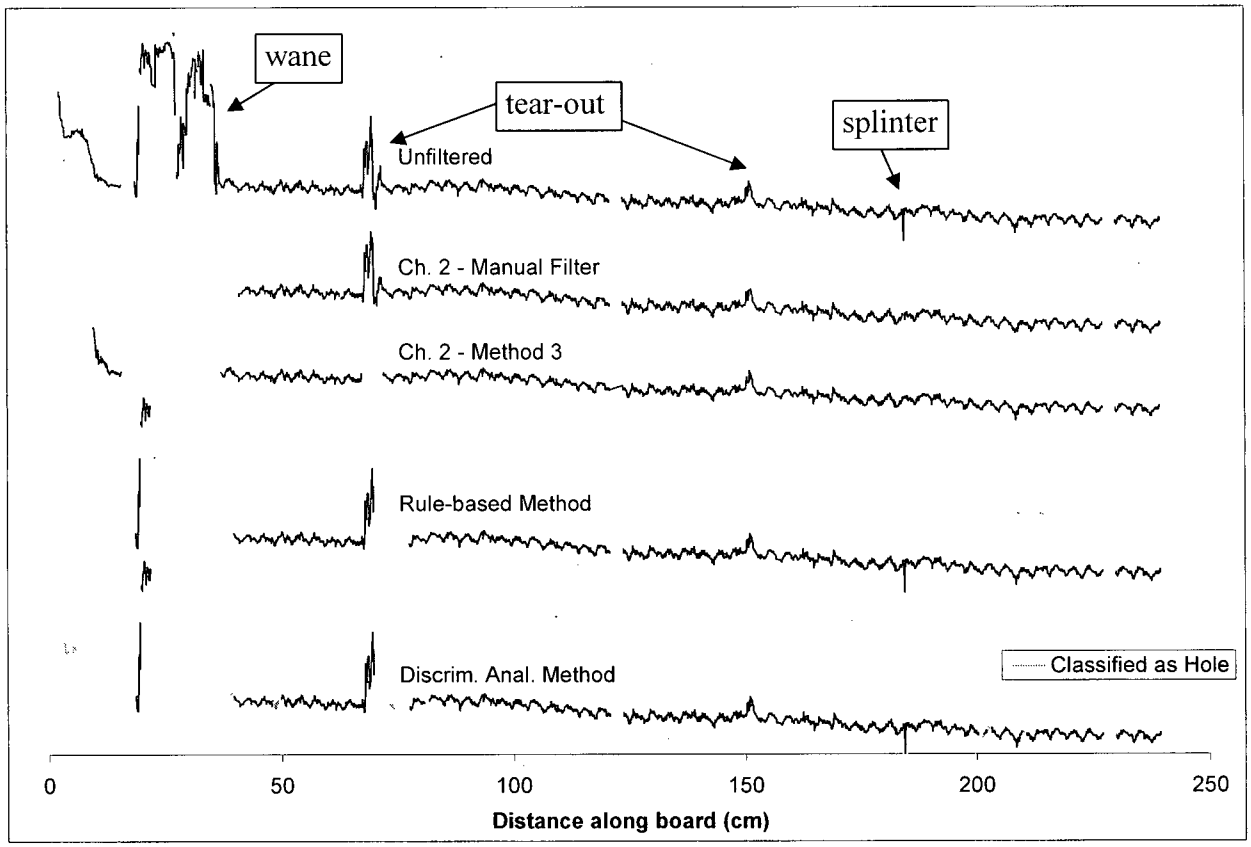



**Figure 3-8. Comparison of Chapter 2 filtering method with camera methods: Board 004.**

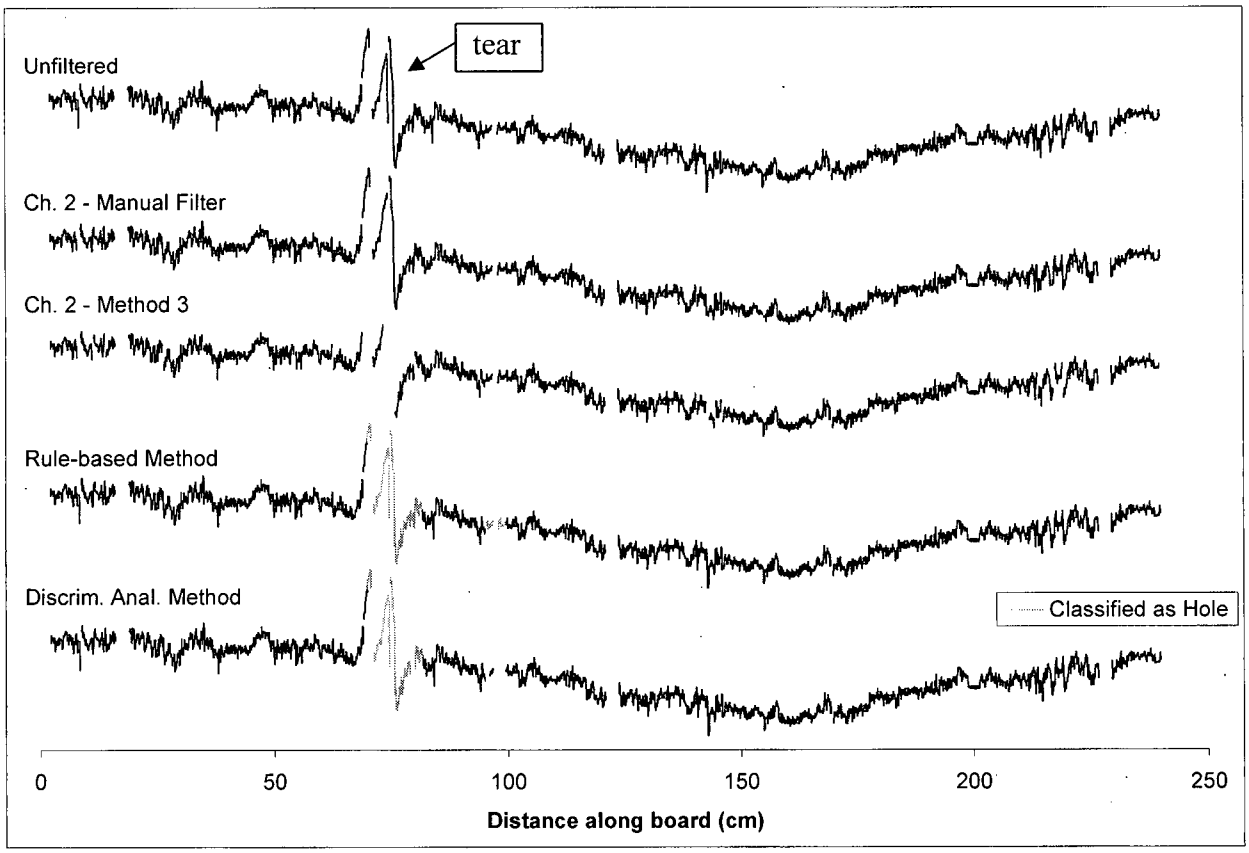



**Figure 3-9. Comparison of Chapter 2 filtering method with camera methods: Board 005.**

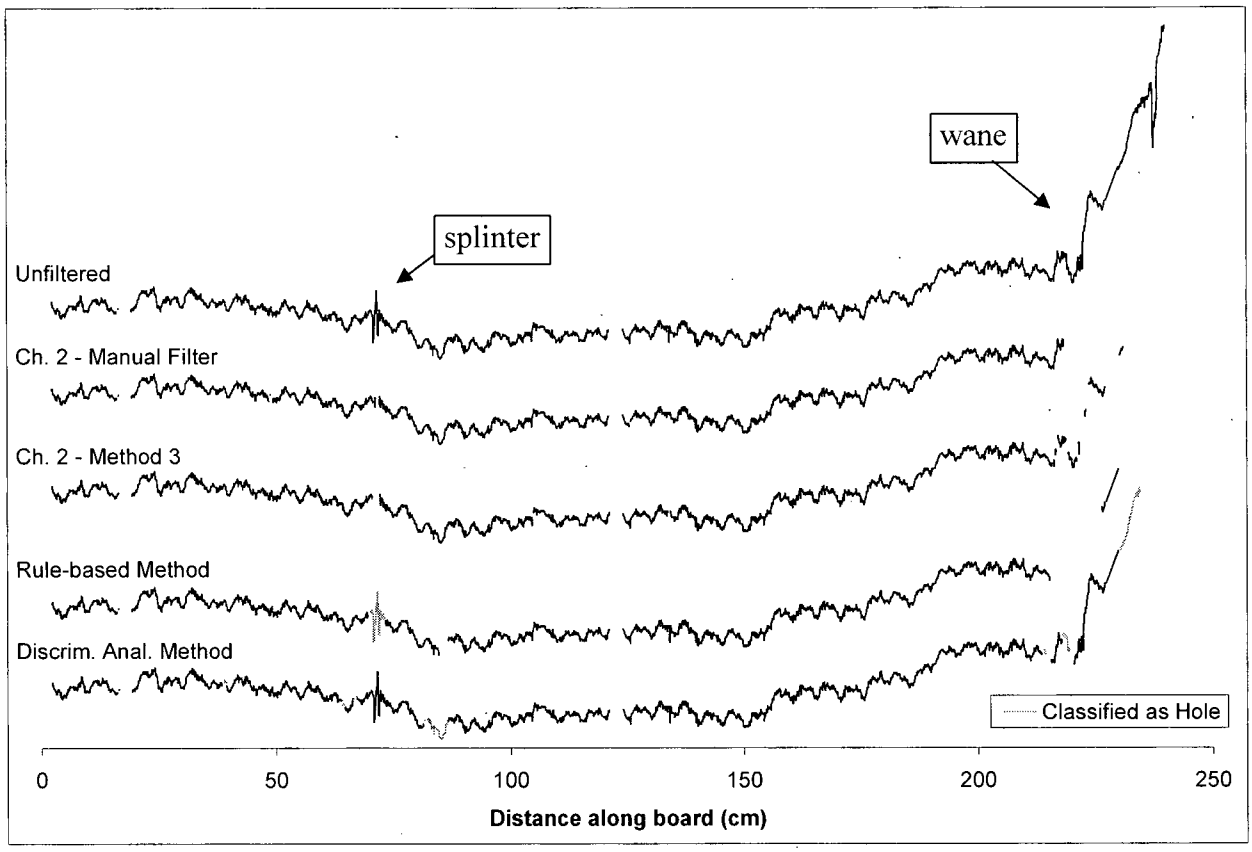



**Figure 3-10. Comparison of Chapter 2 filtering method with camera methods: Board 006.**

Data storage and processing issues limited camera data to about one photograph per centimetre of lumber. The largest amount of classification error occurred with tear-out and holes, which are short-length defects that typically occur over areas of less than 2 cm. With the rule-based method, if only a single frame was marked as a hole or tear, the algorithm re-labeled it as non-defective. This implies that at least some classification failure can be attributed to the algorithm. On the other hand, the discriminant method also performed poorly in identifying holes and tears. This common failure points to a need for better data collection techniques or more advanced hardware components.

Another shortcoming of the methods was that classification rates were better for the bottom laser position versus the top laser position, a result traced to a higher proportion of missing pixel values in the upper portions of the digital pictures. Missing values were likely generated because of incomplete reflectance of the laser line, or reflectance of the laser line at a sub-optimal angle

to the camera lens. This problem would likely only be corrected with upgrades to the camera and/or lens, and would require trial and error type experimentation.

It was interesting to note that classification rates from each of the different saw types were very similar. Chipped surfaces in particular are much rougher, and laser lines projected across the surface of these boards tended to be less straight. Although curve fitting on chipped boards tended to produce splines with more (and changing) curvature, these small perturbations to the curve did not impact results by saw type.

Examination of samples with both high and low classification rates further emphasized the similarities between the two methods. In particular, three of the five boards with the highest classification rates under the rule-based method were also among the top five boards classified under the discriminant method; four of the five boards with the lowest classification rates under the rule-based method were also in the lowest five under the discriminant method.

The samples that had the lowest classification rates included boards with multiple sections of wane and multiple areas of holes and tear-out. Multiple sections of wane often resulted when wane was present along the length of the board, but varied in width, occurring both above and below the LRS position. Close examination of the resulting classifications revealed that the algorithms performed poorly in identifying the multiple starting and ending points of the defective areas.

This result was further confirmed in applying the methods to the six sample boards in the validation dataset from Chapter 2. Of the six boards, those that were defect-free had classification rates of 99.6% and 100% under both methods; the boards with only one defect had rates of 93% and 97% under both methods. In contrast, the rates of correct classification for boards with multiple defects were lower, at only 76-92%. Multiple defects were difficult to

classify regardless of the method chosen. Very small defects, such as splinters were completely missed, and were more appropriately filtered using the Chapter 2 methods.

Overall, the largest source of errors in both methods was misclassification of sawing defects as non-defective. Under the rule-based method and the discriminant method, roughly 2.5% of frames in the validation dataset were classified as non-defective when they contained tear-out. However, in subsequent processing steps, this misclassification would not result in the loss of data, as sawing defects are retained with non-defective areas in the SPC dataset. Areas of wane that are misclassified as non-defective, on the other hand, would be retained when they should have been removed from the dataset. Using the rule-based method with the validation dataset, 2.0% of defect-free frames and frames containing tear-out were classed as wane and therefore would be incorrectly removed from the dataset. Also, 2.0% of frames with wane were incorrectly classed as defect-free or containing tear-out, and thus would be incorrectly retained in the dataset.

Under the discriminant method, misclassification rates for the validation dataset were similar. About 1.8% of frames that were defect-free or contained tear-out were misclassified as wane and would be erroneously removed in subsequent processing steps. About 2.1% of frames with wane were misclassified as non-defective or with holes/tear-out and therefore would erroneously remain in the dataset for subsequent processing steps. The rule-based method, therefore, is slightly more conservative. This method would result in the removal of more defect-free data from the SPC dataset, whereas the discriminant method would incorrectly leave slightly more non-sawing defects in the dataset.

## 3.7 Conclusion

The rule-based and discriminant analysis methods are simple algorithms for classifying defects for the purpose of removing them from real-time SPC data. These methods should perform

accurate defect classification in a real-time environment. Based on a dataset consisting of 30 boards, approximately 87% of the wood along the scan line was defect free, and therefore, assuming all boards were defect free would result in a correct classification rate of 87%. From a practical standpoint, an accuracy rate sufficiently higher than 87% must be achieved in order to justify the cost and computing time of such a system.

The rule-based method is an algorithm consisting of a series of logical steps, each with a distinct purpose in defect recognition. The best-performing discriminant functions, on the other hand, are based on up to seven predictor variables, which are not easily interpretable in terms of defect characterization. Results from the rule-based method were slightly better than that of the discriminant method. Further, the rule-based method is more conservative than the discriminant method; in subsequent processing, using the rule-based method would result in the removal of more non-defective areas than using the discriminant method.

The rule-based method is recommended to remove gross defects from the LRS dataset prior to performing SPC. However, some improvements should be made in order to improve defect classification rates, including:

1. Upgrade camera and/or lens to capture more pixel data in the upper part of the boards;
2. Increase the number of pictures per centimetre to provide better coverage of smaller defects, such as holes, which would require a decrease in the scanning speed and/or an increase in computing speed; and
3. Incorporate additional information to better identify areas of wane. This could be accomplished by using information from the laser scanners themselves, incorporating a multi-sensor approach; however, this would not be possible without increasing the number of pictures taken by the camera.

With these improvements, a camera-based defect detection system could greatly improve the quality of real-time LRS data, which will feed into an expert system for SPC.

## 3.8 Literature Cited


Abbott, A. L., D. L. Schmoldt, P. A. Araman and S. M. Lee. 2001. Automatic scanning of rough hardwood lumber for edging and trimming. *In* Proceedings of ScanTech 2001, the 9th International Conference on Scanning Technology and Process Optimization for the Wood Industry, November 4-6, Seattle, WA. pp. 101-110.

Andersson, C. and F. Walter. 1995. Classification of compression wood using digital image analysis. Forest Products Journal 45(11/12): 87-92.

Brunner, C. C., A. G. Maristany, D. A. Butler, D. VanLeeuwen and J. W. Funck. 1992. An evaluation of color spaces for detecting defects in Douglas-fir veneer. Industrial Metrology 2(3/4): 169-184.

Butler, D. A., C. C. Brunner and J. W. Funck. 1989. A dual-threshold image sweep-and-mark algorithm for defect detection in veneer. Forest Products Journal 39(5): 25-28.

Butler, D. A., C. C. Brunner and J. W. Funck. 2002. Wood surface feature classification via extended-color imagery. Forest Products Journal 52(6): 80-84.

Coates, E. R., S. J. Chang and T. W. Liao. 1998. A quick defect detection algorithm for magnetic resonance images of hardwood logs. Forest Products Journal 48(10): 68-74.

Conners, R. W., T. H. Cho, C. T. Ng, T. H. Dryer, P. A. Araman and R. L. Brisbin. 1992. A machine vision system for automatically grading hardwood lumber. Industrial Metrology 2(3/4): 317-342.

Conners, R. W., D. E. Kline, P. A. Araman and T. H. Drayer. 1997. Machine vision technology for the forest products industry. IEEE Computer 30(7): 43-48.

Forrer, J. B., D. A. Butler, J. W. Funck and C. C. Brunner. 1988a. Image sweep-and-mark algorithms. Part 1. Basic algorithms. Forest Products Journal 38(11/12): 75-79.

Forrer, J. B., D. A. Butler, J. W. Funck and C. C. Brunner. 1988b. Image sweep-and-mark algorithms. Part 2. Performance evaluations. Forest Products Journal 39(1): 39-42.

Gazzarri, J. 2003. Sequential measurements method for moving surfaces profiling. MSc Thesis. Department of Mechanical Engineering, The University of British Columbia, Vancouver, Canada.

Gonzalez-Andrados, J. R., F. Lopes and H. Pereira. 2000. Quality grading of cork planks with classification models based on defect characterisation. Holz als Roh-und Werkstoff 58: 39-45.

Huber, H. A., C. W. McMillin and J. P. McKinney. 1985. Lumber defect detection abilities of furniture rough mill employees. Forest Products Journal 35(11/12): 79-82.

Klecka, W. R. 1980. Discriminant Analysis. Beverly Hills, CA, Sage Publications.

Kline, D. E., C. Surak and P. A. Araman. 2001. Evaluation of a multi-sensor machine vision system for automated hardwood lumber grading. *In* Proceedings of 4th International Conference on Image Processing and Scanning of Wood, August 21-23, 2000, Mountain Lake, VA. pp. 75-87.

Koivo, A. J. and C. W. Kim. 1989. Automatic classification of surface defects on red oak boards. Forest Products Journal 39(9): 22-30.

Lee, S. M., A. L. Abbott and D. L. Schmoldt. 2001. Wane detection of rough lumber using surface approximation. *In* Proceedings of 4th International Conference on Image Processing and Scanning of Wood, August 21-23, 2000, Mountain Lake, VA. pp. 115-126.

Manly, B. F. J. 1994. Multivariate Statistical Methods: a Primer. 2nd ed. London, UK, Chapman & Hall.

Masi, C. G. 2003. Vision inspection grades rough wooden boards. Vision Systems Design 8(7): 33-37.

Nyström, J. and D. E. Kline. 2000. Automatic classification of compression wood in green southern yellow pine. Wood and Fiber Science 32(3): 301-310.

Pham, D. T. and R. J. Alcock. 1998. Automated grading and defect detection: a review. Forest Products Journal 48(4): 34-42.

SAS Institute. 2002. SAS/STAT User's Guide, Version 8. Cary, NC, SAS Publishing.

Schmoldt, D. L., P. Li and A. L. Abbott. 1997. Machine vision using artificial neural networks with local 3D neighborhoods. Computers and Electronics in Agriculture 19: 255-271.

Schmoldt, D. L., H. Song and P. A. Araman. 2001. Real-time value optimization of edging and trimming operations for rough, green hardwood lumber. *In* Proceedings of ScanTech 2001, the 9th International Conference on Scanning Technology and Process Optimization for the Wood Industry, November 4 - 6, Seattle, WA. pp. 87-100.

Silvén, O. and H. Kauppinen. 1996. Recent developments in wood inspection. International Journal of Pattern Recognition and Artificial Intelligence 10(1): 83-95.

Silvén, O., M. Niskanen and H. Kauppinen. 2003. Wood inspection with non-supervised clustering. Machine Vision and Applications 13: 275-285.

Szymani, R. 1985. An overview of scanning technology in sawmilling. *In* Proceedings of 1st International Conference on Scanning Technology in Sawmilling, San Francisco, CA. Miller Freedman Publications. pp. 1-33.

Tabachnick, B. G. and L. S. Fidell. 1983. Using Multivariate Statistics. New York, Harper & Row.

# Chapter 4 Mixed-model Development for Real-time Statistical Process Control Data in Wood Products Manufacturing

## 4.1 Introduction

Statistical process control (SPC) in wood products manufacturing has traditionally been a labour-intensive job, relying on relatively simple equipment and basic statistical methods. In a typical lumber mill, SPC personnel periodically sample a small number of pieces of lumber from each sawing machine, measure the thickness of each piece in 6-10 places with digital calipers, and enter these data into an SPC software package. These packages produce a variety of charts, which display process control limits and give mill personnel feedback on the performance of sawing machines (for basic SPC concepts, see, e.g., Montgomery 2001). SPC is widely accepted in automated lumber manufacturing, as it is used as an analytical tool for problem solving and provides a signal to operators when sawing machines need adjustment. When properly applied, SPC can prevent continued production of sub-standard material.

Recent advances in technology have made extremely accurate non-contact laser range sensors (LRS) affordable to lumber mills. These measurement devices can take up to 1000 measurements per second and can be set up in-line with sawing machines to measure each piece of lumber processed. Under typical mill conditions, about 3000 measurements could be taken on a 2.5 metre (8-foot) board, resulting in a more than 1000-fold increase in the amount of SPC data available. Moreover, these devices can be set up to obtain information specific to each side of each board. Whereas ordinary digital calipers give thickness measurements only, LRS data could possibly describe each surface of a board. This is especially important in modern sawmills, where two different cutting devices usually make the two "face" surfaces of a single piece of lumber.

This advance would seem like a windfall to SPC personnel in that it greatly increases the amount of information available to make decisions about the quality of lumber being produced. However, data from LRSs contain measurement errors (Wehr and Lohr 1999; Burman 2002) and non-sawing defects (e.g., wane[18] and holes), and their distributional properties have not yet been studied. The statistical model that underlies traditional SPC techniques for wood products data was developed for periodic sampling described by a one-way analysis of variance model (Warren 1973). Data from LRS-based real-time systems, on the other hand are virtually continuous, with several sources of variation. Using methods from image processing, filtering techniques have been developed to remove measurement errors and non-sawing defects (Chapters 2 and 3). In order to develop a system for SPC using real-time LRS data, an appropriate statistical model must be derived.

Continuous process data are often highly autocorrelated (Wheeler 1995; Young and Winistorfer 2001; Noffsinger and Anderson 2002); most observations are easily predicted from their neighbours. While autocorrelation in data does not bias estimates of process parameters (e.g., mean lumber thickness), estimates are inefficient as each observation does not represent a new piece of independent information (Schabenberger and Pierce 2001). This causes underestimation of the variation in the process parameters and overestimation of their degrees of freedom, which in turn causes construction of incorrect control limits (Wheeler 1995). Perhaps more important, the components of variance in the statistical model need to be correctly identified, as the sources of variation are keys to understanding the sawing process (Maness et al. 2002). The statistical model must estimate and account for all known sources of variation, in consideration of the inherent autocorrelation in the process.

---

[18] Wane is the natural curvature of the edge of a board sawn from a log. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

This statistical model will be an integral part of an SPC system for real-time LRS data. The ultimate goal of any SPC system is process improvement (Shewhart 1931). The goal of this particular SPC system is to evaluate board surface profiles, providing real-time feedback on sawing performance and monitoring for specific kinds of sawing defects. Better knowledge of the process and faster response to problems when they occur will ultimately enable mills to reduce target sizes and save money.

## 4.2 Objectives

The objective of this chapter is to describe a statistical model for online LRS profile data taken on sawn lumber, which will enable an SPC system to be created. The model and its estimated components of variance (COV) were identified, partitioned, and calculated, providing the basis for a real-time lumber manufacturing SPC protocol.

## 4.3 Materials

### 4.3.1 Lab Scans

A laser measurement apparatus was set up in the Q-Lab of the Department of Wood Science in the Forest Sciences Centre at The University of British Columbia, Vancouver, Canada (Figure 4-1). This apparatus consisted of a moving carriage, encoder, and four laser measurement devices. The motor controlling the carriage could be set to very precise speeds defined by the user, from 0-2.5 m (8 feet) per second. Four Hermary LRS-50 point laser range sensors were mounted securely, two on each side of the carriage. Side 1-Laser 1 and Side 2-Laser 1 were vertically positioned to take measurements 2.54 cm (one inch) above the bottom of the board; Side 1-Laser 2 and Side 2-Laser 2 were vertically positioned to take measurements 2.54 cm (one inch) below the top of the board (Figure 4-2). This measurement apparatus was manufactured to high standards and calibrated periodically to ensure accurate measurements.

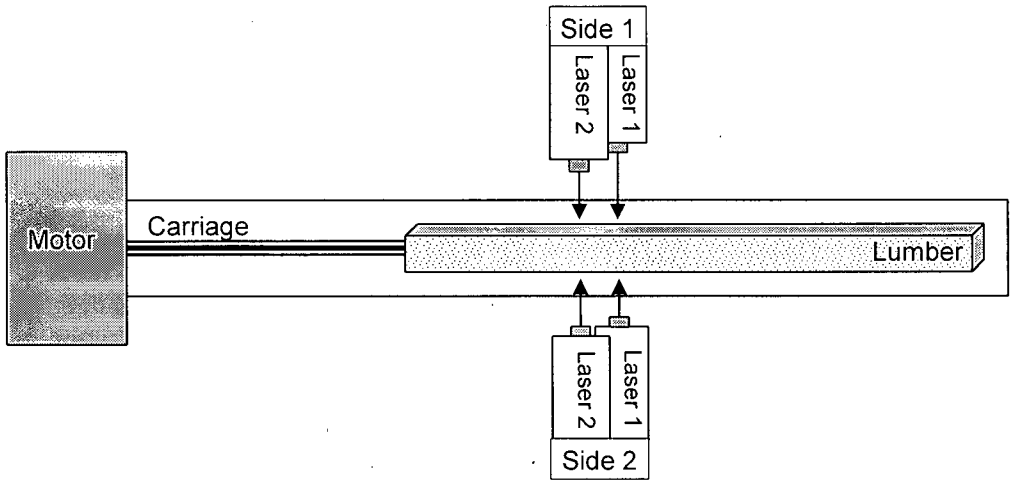



**Figure 4-1. Set up of measurement apparatus in Q-lab, as viewed from above.**

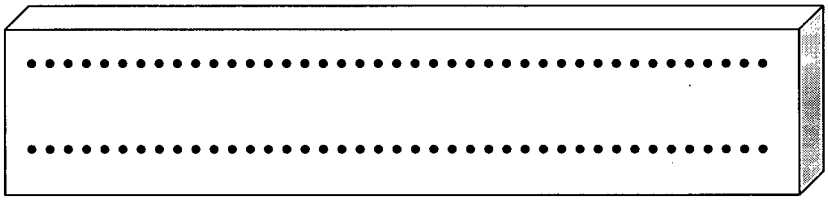


**Figure 4-2. Location of measurements taken on Side 2 of a single board, as viewed from side 2.**

Boards were secured to the carriage with a clamping system, which pushed Side 1 of the board flush against a fixed rail on Side 2. For a perfectly flat surface, the boards would be a constant distance from the Side 1 lasers.

As the carriage moved the boards past the laser scanners at a controlled speed, the four streams of laser measurement data and encoder measurements of the location of the carriage were sent to a data concentrator and passed to a computer via Ethernet cable. The raw LRS data consisted of the distance from each of the four lasers to the wood surface, and the encoder data consisted of the distance along the length of the carriage.

Sample lumber obtained from Weyerhaeuser's New Westminster (British Columbia, Canada) sawmill consisted of 100 arbitrarily selected pieces of *Taruki*, a western hemlock (*Tsuga heterophyla* (Raf) Sarg.) dimension lumber product with target thickness and width dimensions 80 × 135 mm (2 1/32 × 5 5/16 inches), respectively. The green and un-planed samples of *Taruki*

were taken at the sorter. The primary processing of this lumber was done at the quad-bandsaw, where the logs were squared and cut into cants (Figure 4-3).

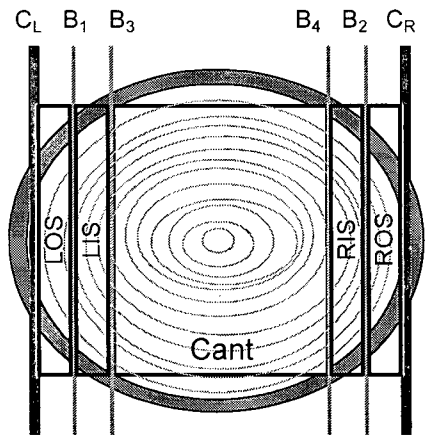



**Figure 4-3. Sawing solution with four sideboards ($C_L$=left chipper-head, $C_R$=right chipper-head, $B_1$=bandsaw #1, $B_2$=bandsaw #2, $B_3$=bandsaw #3, $B_4$=bandsaw #4, LOS=left outside sideboard, LIS=left inside sideboard, RIS=right inside sideboard, ROS=right outside sideboard).**

With this four-board solution, there were four possible saw configurations under which sample boards were sawn. Boards cut from the outermost part of the log (outside sideboards) had one side cut by a chipper-head ($C_L$ or $C_R$) and one side cut by a bandsaw ($B_1$ or $B_2$). Right outside sideboards (ROS) were labeled as saw configuration "BC", and left outside sideboards (LOS) were labeled as saw configuration "CB". Inside sideboards had both sides cut by a bandsaw. Since the lumber was collected at the sorter, left inside sideboards (LIS) and right inside sideboards (RIS) were indistinguishable and labeled as saw configuration "BB". Boards cut from the cant itself had both sides cut in a gang of nineteen circular saws (not shown), and were labeled with saw configuration "RR".

With smaller logs, there were other possible sawing solutions; for example, there could be one sideboard on each side (sawing pattern 1-1 in Figure 4-4) or no sideboards at all (sawing pattern 0-0). In these cases, the cant was always cut with saws $B_3$ and $B_4$. Therefore, samples labeled as

saw configuration BC could have been sawn with $B_2$ and $C_R$, or with $B_4$ and $C_R$, and similarly, CB samples could have been sawn with $B_1$ and $C_L$, or with $B_3$ and $C_L$.

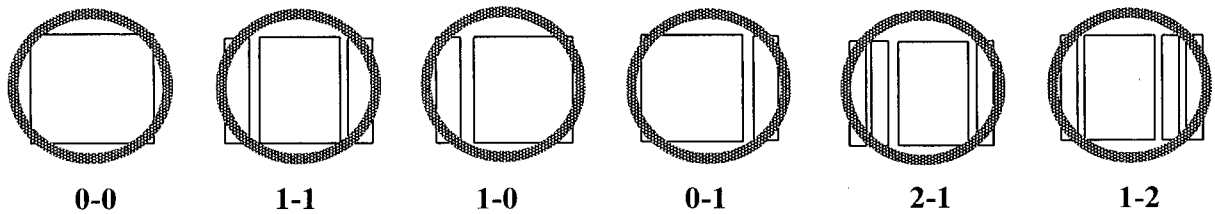



**Figure 4-4. Other possible sawing solutions (0-0 = 0 left and 0 right sideboards, 0-1 = 0 left sideboards and 1 right sideboard, etc.).**

A summary of the possible saw and chipper-head combinations is listed in Table 4-1. For saw configurations CB and BC, the chipper-heads were uniquely identified; bandsaws and circular saws could not be identified with certainty.

**Table 4-1. Summary of saw and/or chipper-head combinations by saw configuration[19].**

| Saw Configuration | Board Type | Side 1 | Side 2 |
|---|---|---|---|
| BB | LIS | $B_1$ | $B_3$ |
| | RIS | $B_4$ | $B_2$ |
| BC | ROS | $B_2$ | $C_R$ |
| | ROS | $B_4$ | $C_R$ |
| CB | LOS | $C_L$ | $B_1$ |
| | LOS | $C_L$ | $B_3$ |
| RR | Cant | $R_1$- $R_{18}$ | $R_2$- $R_{19}$ |

### 4.3.2 Field Scans

Since the lumber for lab scanning was collected at the sorter, the order in which the boards were sawn was unknown. Because sawing order is needed to quantify the autocorrelation between sawn boards, an ordered field sample was obtained. This sample was collected for a separate related real-time laser scanning project, the objective of which was to verify an algorithm that identified and removed vibrations caused by the mill machinery from the online LRS profile data

[19] BB=Bandsaw–Bandsaw configuration, BC=Bandsaw–Chipper-head configuration, CB=Chipper-head–Bandsaw configuration, RR=Circular Saw–Circular Saw configuration, LOS=left outside sideboard, LIS=left inside sideboard, RIS=right inside sideboard, ROS=right outside sideboard, $C_L$=left chipper-head, $C_R$=right chipper-head, $B_1$=bandsaw #1, $B_2$=bandsaw #2, $B_3$=bandsaw #3, $B_4$=bandsaw #4, $R_1$-$R_{19}$=circular saws #1-#19, respectively.

(Gazzarri 2003). Sample data were collected at Weyerhaeuser's New Westminster sawmill using an in-line measurement system. Cants were scanned at the outfeed of the quad bandsaw, using a process similar to that in the Q-Lab. The sample data consisted of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) cants cut by the quad bandsaw with three possible saw configurations: BB, BC, or CB. The profile measurements were taken 5 cm (2 inches) below the top of the cant to avoid potential areas of wane. Because the vibration removal algorithm had complex input requirements, only two streams of LRS data (one per side) were available from the measurement system[20].

## 4.4 Methods

### 4.4.1 Lab Scan Data

Within 48 hours of processing, the sample lumber was delivered to the Q-lab and cut to 2.5-metre (8-foot) lengths. Since some samples were longer than 5 metres (16 feet), this resulted in a total of 110 2.5-metre samples. The lumber was kept at a very wet state (>> 30% moisture content), and data were collected over a period of four days. Qualitative notes were made on the individual boards, including the saw configuration, direction of sawing, obvious saw mark patterns, and defects.

Measurements were taken at time intervals and speeds consistent with conditions in a typical lumber mill. With the carriage motor set at 500 rpm and the lasers set to collect 500 measurements per second, the datasets consisted of ~1200 measurements per metre (375 per foot) from each laser, or ~3000 measurements per board, side, and laser position.

---

[20] See Gazzari (2003) for additional measurement details.

The data captured by the point laser range sensors were distances from the laser to the board surface. At the $m$th measurement point along the board, the four LRS observations were: $l_{11_m}$, $l_{12_m}$, $l_{21_m}$, $l_{22_m}$ (Figure 4-5). The distance between the Side 1 lasers and the Side 2 lasers is denoted $d$. In order to perform SPC on measurements related to the board dimensions, observations from the lasers were translated into four profile quantities, $y_{11_m}$, $y_{12_m}$, $y_{21_m}$, and $y_{22_m}$. Profiles were calculated by first finding the "centre of sawing" for each board via two separate regressions with the two top and two bottom laser quantities. Regressing $l_{11m}$ and $d$ - $l_{21m}$ versus $m$ yielded a line through the centre of the board at the bottom laser height. Similarly, regressing $l_{12m}$ and $d$ - $l_{22m}$ versus $m$ yielded a line through the centre of the board at the top laser height. By drawing a line connecting these two regression lines, the board is essentially split in half vertically. This arbitrary centerline was used as a reference from which to calculate the profiles[21].

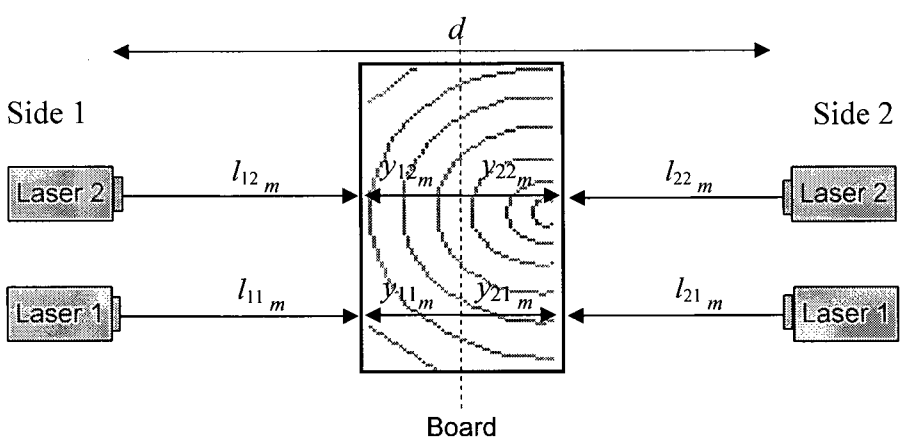



**Figure 4-5. Diagram of laser measurement quantities and derived surface profiles ($l_{11m}$=$m$th measurement from Side 1-Laser 1, $l_{12m}$=$m$th measurement from Side 1-Laser 2, $l_{21m}$=$m$th measurement from Side 2-Laser 1, $l_{22m}$=$m$th measurement from Side 2-Laser 2, $d$=distance from Side 1 lasers to Side 2 lasers).**

[21] Note that this method of profile calculation ensured that for each board, $\bar{y}_{11} = \bar{y}_{21}$ and $\bar{y}_{12} = \bar{y}_{22}$.

Raw data from the laser scanners included non-sawing defects and measurement errors. Filtering for gross defects, such as wane and surface debris was done manually, using the known locations of defects. Filtering for measurement errors was accomplished via the "MSD filter", an iterative outlier removal process (Method 3, Chapter 2). Filtering resulted in data reduction of 1-2% in most cases. Approximately 20% of boards had at least some wane removed; for about 5% of boards, large amounts of wane resulted in the removal of ~20% of profile observations.

### 4.4.2 Field Scan Data

Cants were continuously scanned over a period of several hours. For each cant, the cutting pattern was noted and gross defects such as wane were recorded. After application of Gazzari's (2003) vibration removal algorithm, cant profile data were obtained in the same manner as for the lab scans. Because the field scan data were used only to quantify the autocorrelation between subsequent boards, the data simply consisted of scan numbers, saw configurations, and average profiles.

In total, 208 samples were obtained. Because only relatively large-sized logs were cut during the time of sampling, nearly 95% of cants were bandsawn on both sides (BB). Thus, the analysis was done for this saw configuration only. Wane sections were removed from the scans and the analysis for board to board correlation was performed with 194 cants in the sample pool.

### 4.4.3 Model

Each saw configuration represents a different sawing process, and each side of each board is sawn by two different saws. Thus, models were developed separately for each saw configuration and side. Adding subscripts to allow for multiple boards, and to denote the saw configuration, $y_{ijklm}$ denotes the profile observation from the $i$th saw configuration ($i$ = 1 to 4: BB, BC, CB, and RR), $j$th side ($j$ = 1 to 2), $k$th sample board ($k$ = 1 to $b_i$; $b_i$ =41 for BB, $b_i$ =24 for BC and CB, and

$b_i$ =21 for RR), $l$th laser location ($l$ = 1 to 2), and $m$th distance along the board ($m$ = 1 to $n_{ijkl}$). Using a mixed-effects model, these observations were described as:

$$y_{ijklm} = \mu_{ij} + \beta_{ijk} + \lambda_{ijl} + \beta\lambda_{ijkl} + \varepsilon_{ijklm} \qquad [4\text{-}1]$$

where: $\mu_{ij}$ = the mean profile of the $i$th saw configuration and $j$th side;

$\beta_{ijk}$ = the $k$th board effect from the $i$th saw configuration and $j$th side;

$\lambda_{ijl}$ = the $l$th laser location effect from the $i$th saw configuration and $j$th side;

$\beta\lambda_{ijkl}$ = the interaction of the $k$th sample board and $l$th laser location from the $i$th saw configuration and $j$th side; and

$\varepsilon_{ijklm}$ = the error associated with the $m$th measurement from the $l$th laser location and $k$th sample board, in the $i$th saw configuration and $j$th side.

Each of the eight saw configuration ($i$) × side ($j$) combinations were modelled separately to allow for different error distributions. Under the "usual" mixed-model analysis approach, the effects in each ($ij$) model are assumed independent, identically and normally distributed:

$\beta_{ijk} \sim N(0, \sigma_{\beta_{ij}}^2)$, $\lambda_{ijl} \sim N(0, \sigma_{\lambda_{ij}}^2)$,

$\beta\lambda_{ijkl} \sim N(0, \sigma_{\beta\lambda_{ij}}^2)$, $\varepsilon_{ijklm} \sim N(0, \sigma_{\varepsilon_{ij}}^2)$; and

$\mathrm{Cov}(\varepsilon_{ijklm}, \varepsilon_{ijk'l'm}) = 0$ for $k \neq k'$ and $l \neq l'$.

When these assumptions are met, the variance estimators are unbiased and thus, tests of statistical significance can be performed and confidence limits can be calculated. However, given that the measurements are very closely spaced, with ~3000 measurements taken from each laser per side per board, a high degree of spatial- and auto-correlation among measurements taken from each LRS on a particular board is likely.

The autocorrelation in time-ordered data is measured by the autocorrelation function (ACF). The sample ACF of the profile data from a single board, side, and laser position at lag $t$ measures the similarity between measurements $t$ lags apart:

$$\hat{\rho}_{ijkl_t} = \sqrt{\frac{\sum_{m=1}^{n_{ijkl}} (y_{ijkl,m} - \bar{y}_{ijkl\cdot})(y_{ijkl,m+a} - \bar{y}_{ijkl\cdot})}{\sum_{m=1}^{n_{ijkl}} (y_{ijkl,m} - \bar{y}_{ijkl\cdot})^2}} \qquad [4\text{-}2]$$

where: $t$=1, 2, 3, …; and $\bar{y}_{ijkl\cdot}$ is the mean profile for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position.

In the presence of non-zero autocorrelation, the assumption of independent errors ($\varepsilon_{ijklm}$) is not valid (Schabenberger and Pierce 2001). Measurements from the two lasers on the same board may also be correlated, affecting the distributional assumptions about $\lambda_{ijl}$. Furthermore, if boards were scanned in the order they were sawn, there could be strong correlations from subsequent boards from a particular saw, affecting the distributions of $\beta_{ijk}$ and $\beta\lambda_{ijkl}$. Without accounting for these sources of correlation, the variation in this process could be under-estimated (Schabenberger and Pierce 2001).

Model [4-1] was fit first ignoring sources of autocorrelation in the data. It was then re-fitted using a multi-stage model, where autocorrelation in the errors was explicitly modeled, and the two models were compared.

### 4.4.3.1 Model Ignoring Autocorrelation

Using the lab scan profile data, the parameters of Model [4-1] for each saw configuration and side (eight combinations) were estimated using the SAS procedure PROC MIXED[22]. The PROC

[22] All SAS procedures referred to in this chapter were run with Version 8.2 (SAS Institute 2002).

MIXED procedure allowed for fixed and random effects, and computed estimates of the variance components for all random effects in the model. For each saw configuration and side ($ij$), a variance estimate was found for: board ($\sigma^2_{\beta_{ij}}$), laser position ($\sigma^2_{\lambda_{ij}}$), the interaction ($\sigma^2_{\beta\lambda_{ij}}$), and the residual ($\sigma^2_{\varepsilon_{ij}}$).

### 4.4.3.2 Multi-Stage Model

In multi-stage modeling, the model parameters are estimated in steps. In the first step, lab scan data were modeled within each board and side ($ijk$) with an appropriate autocorrelative structure. In step 2, the autocorrelation between subsequent sample boards ($ijk$ to $ij,k+1$) was investigated using the field scan data. In the final step, Model [4-1] was fit to the lab scan profile data using the estimated autocorrelative structure for the errors ($\varepsilon_{ijklm}$).

#### *4.4.3.2.1 Stage 1: Within Sample Board Model*

Data taken from one side of a single board ($ijk$) are samples from a random field; since different saws are responsible for cutting each side of the board, LRS data from the opposite side of the board are samples from a separate and distinct random field. These fields could be considered in either a spatial or temporal context. Since a single piece of lumber comes from one tree, subject to one set of genetic and environmental conditions, the surface data from a single piece of lumber could be considered spatially correlated. On the other hand, the surface of the lumber was measured in order to monitor and diagnose saw performance. Saws rotate and vibrate at rates and angles that depend on an array of factors, which in turn depend to some degree on the qualitative properties of the wood (e.g., density and moisture content). However, saw performance depends to a great extent on operator controlled factors, such as the feedspeed of the log, saw tension, and sharpness of the sawblade (Schajer 1990). Although feedspeed can

change during the cut, these factors are generally more related to time than to space. Autocorrelation within a single board and side therefore was estimated with time series models.

The autocorrelation structure of the errors was estimated using univariate time series models by saw configuration, side, board, and laser position (*ijkl*). Although a bivariate model for simultaneously fitting the profile data from the top and bottom laser positions by board and side (*ijk*) can be more efficient under certain conditions, it is required when a non-zero correlation exists between the two laser positions. It was assumed that the correlation between laser positions was not significantly different from zero for model fitting, and this assumption was tested. Several time series models were fit (Appendix I). Because each of the three types of saws (Bandsaw, Chipper-head, and Circular Saw) produced different surface variability, different model forms were investigated by saw type. Models were chosen by visually assessing of lack of fit and by computing Akaike's Information Criteria (AIC, Box *et al.* 1994). For data from a single board, side, and laser position, the AIC is calculated as:

$$AIC_{ijkl} = -2\log \ell(\hat{\boldsymbol{\psi}}_{ijkl} \mid \mathbf{y}_{ijkl}) + 2K \qquad [4\text{-}3]$$

where: $\ell(\hat{\boldsymbol{\psi}}_{ijkl} \mid \mathbf{y}_{ijkl})$ is the empirical maximized log likelihood function;

$\hat{\boldsymbol{\psi}}_{ijkl}$ is a row vector of estimated model parameters for the *i*th saw configuration, *j*th side, *k*th board, and *l*th laser position;

$\mathbf{y}_{ijkl}$ is a column vector of profile measurements from the *i*th saw configuration, *j*th side, *k*th board, and *l*th laser position; and

$K$ is the number of parameters in the model.

For each saw type, the best-fitting model form with the lowest AIC for the majority of series in the saw type was chosen.

To simplify notation in fitting separate models by saw configuration, board, side, and laser position, the subscript *ijkl* was replaced by *f*. Ignoring random effects for this step simplified Model [4-1] to:

$$y_{fm} = \mu_f + \varepsilon_{fm} \qquad [4\text{-}4]$$

$\hat{\mu}_f$ was simply estimated as the mean value of the *f*th (saw configuration × board × side × laser position) profile. To account for autocorrelated errors within board and side, $\{\varepsilon_{fm}\}$ were modeled. For bandsawn boards, an autoregressive integrated moving average (ARIMA) model with first order differencing (i.e., the differencing parameter $\delta_f = 1$), one autoregressive parameter ($\phi_f$), and one moving average parameter ($\theta_f$) (ARIMA(1,1,1)) was chosen:

$$(1-\phi_f B)(\varepsilon_{f,m} - \varepsilon_{f,m-1}) = \alpha_f + (1-\theta_f B)\upsilon_{f,m} \qquad [4\text{-}5]$$

where: $B$ is the backshift operator, e.g., $(1 - \phi_f B)\varepsilon_{fm} = \varepsilon_{fm} - \phi_f\, \varepsilon_{f,m-1}$;

$\alpha_f$ is the intercept; and

$\upsilon_{fm}$ is a white noise error process, with $\upsilon_{fm} \sim N(0, \sigma_{\upsilon_f}^2)$.

An ARIMA model was also chosen to model the autocorrelated errors for about half of the circular-sawn boards. In particular, when the circular-sawn boards were free of cyclical patterns resulting from deep saw marks, an ARIMA(0,1,1) model was chosen:

$$\varepsilon_{f,m} - \varepsilon_{f,m-1} = \alpha_f + (1-\theta_f B)\upsilon_{f,m} \qquad [4\text{-}6]$$

For chipped boards and circular-sawn boards exhibiting cyclical patterns, a model with seasonal and long-memory terms was more appropriate. These data were modeled with a seasonal autoregressive fractionally integrated moving average (SARFIMA) model. Whereas $\phi_f$ and $\theta_f$ describe the short-term autoregressive and moving average behaviour of the ARIMA(1,1,1)

series, respectively, $\Phi_f$ and $\Theta_f$ are parameters that describe cyclical (seasonal) behaviour with cycle length $s_f$. Fractional (versus integer) values for the differencing order, $\delta_f$, give the model long memory, and $\delta'_f$ describes the cyclical nature of the long-memory behaviour. The chosen model was of the form SARFIMA(1,$\delta_f$,1)x $(1,\delta'_f,1)_{s_f}$:

$$(1-\Phi_f B^{s_f})(1-\phi_f B)(1-B^{s_f})^{\delta'_f}(1-B)^{\delta_f}\varepsilon_{fm} = \alpha_f + (1-\theta_f B)(1-\Theta_f B^{s_f})\upsilon_{fm} \qquad [4\text{-}7]$$

The parameters of the ARIMA models were estimated using the SAS procedure PROC ARIMA. The parameters for the SARFIMA models were obtained in several steps. First, the seasonal period $s_f$ of each board × side was found by examining the spectrum of the first-differenced data using the SAS procedure PROC SPECTRA. Using the SAS linear regression procedure PROC REG, the parameters $\delta_f$ and $\delta'_f$ were then estimated simultaneously following Andel's (1986) extension of Geweke and Porter-Hudak's (1983) method. Finally, the autoregressive and moving average parameters were estimated with PROC ARIMA.

Although ARIMA and SARFIMA models fit the data well, it should be noted that these models are non-stationary when $\delta_f$ and/or $\delta'_f$ are > 0.5. This implies an unbounded variance of $\varepsilon_{fm}$ as the number of observations along the board, *m,* increases. While these models are valid in a descriptive sense, this property makes inference about individual data points questionable.

After fitting models by board, side, and laser (*f=ijkl*), the bivariate relationship between the top and bottom laser for each board-side combination (*ijk*) was assessed by examining the cross-correlation function (CCF). The CCF measures the similarity between two variables, computed by the sum of the cross products between the two variables at different lags. Plots of the CCF were examined to check the assumption of non-zero correlation between the top and bottom laser positions at a significance level ($\alpha$) of 0.05.

#### *4.4.3.2.2 Stage 2: Between Sample Boards Model*

Since the lab scan samples were not scanned in the same order as sawn, the autocorrelation between sample boards was evaluated using the field scan data. PROC ARIMA was run with the mean values of each board × side (*ijk*) in the order they were sawn. The model form that produced the best (lowest) value of Bayesian Information Criteria (BIC) was found using the sample ACF. For data from a single board, side, and laser position, the BIC is calculated as:

$$BIC_{ijkl} = -2\log \ell(\hat{\boldsymbol{\psi}}_{ijkl} \mid \mathbf{y}_{ijkl}) + K\ln(n_{ijkl}) \qquad [4\text{-}8]$$

The hypothesis of no autocorrelation between subsequent boards was verified when an ARIMA(0,0,0) model produced the lowest BIC.

#### *4.4.3.2.3 Stage 3: Mixed-effects Model*

Ideally, Stage 3 would involve fitting Model [4-1] with the error covariance matrix structure estimated in Stages 1 and 2 using the SAS procedure PROC MIXED. However, due to computing limitations, this was not possible. Instead, Model [4-1] was modified to use the estimates of the residuals from Stages 1 and 2, $\hat{\varepsilon}_{fm}$. Using $\hat{\mu}_f$ estimated from [4-4], $z_{fm}$ was defined to represent the profile measurements without autocorrelation:

$$z_{fm} = \hat{\mu}_f + \hat{\upsilon}_{fm}, \quad \text{or} \qquad [4\text{-}9]$$

$$z_{ijklm} = \hat{\mu}_{ijkl} + \hat{\upsilon}_{ijklm} \qquad [4\text{-}10]$$

where: $\varepsilon_{fm} - \hat{\varepsilon}_{fm} = \hat{\upsilon}_{fm}$.

In Stage 3, PROC MIXED was used to estimate the parameters in the modified model in [4-11] for every saw configuration and side (*ij*):

$$z_{ijklm} = \mu_{ij} + \beta_{ijk} + \lambda_{ijl} + \beta\lambda_{ijkl} + \nu_{ijklm} \qquad [4\text{-}11]$$

where: $\nu_{ijklm}$ are independent, identically distributed errors with $\nu_{ijklm} \sim N(0,\sigma_{\nu_{ij}}^2)$, and other terms are distributed as in Model [4-1].

## 4.5 Results

Table 4-2 shows descriptive statistics for the profile data by each saw configuration and side. The average number of observations ($\bar{n}_{ij..}$) is slightly lower than 3,000 because of filtering for measurement errors and non-sawing defects[23]. Given this large $\bar{n}_{ij..}$, the minimum and maximum show a large range. All statistics shown were somewhat lower for bandsawn boards.

**Table 4-2. Descriptive statistics by saw configuration[24] and side (cm).**

| Saw Configuration | Side | $\bar{n}_{ij..}$ | Minimum | Maximum | Average | Standard Deviation |
|---|---|---|---|---|---|---|
| BB | 1 (bandsawn) | 2773 | 2.380 | 2.896 | 2.575 | 0.0509 |
| | 2 (bandsawn) | 2779 | 2.360 | 2.782 | 2.575 | 0.0502 |
| BC | 1 (bandsawn) | 2750 | 2.389 | 2.906 | 2.632 | 0.0626 |
| | 2 (chipped) | 2706 | 2.316 | 2.938 | 2.633 | 0.0668 |
| CB | 1 (chipped) | 2662 | 2.404 | 2.912 | 2.656 | 0.0630 |
| | 2 (bandsawn) | 2724 | 2.464 | 2.914 | 2.654 | 0.0562 |
| RR | 1 (circular-sawn) | 2727 | 2.434 | 2.883 | 2.650 | 0.0631 |
| | 2 (circular-sawn) | 2778 | 2.442 | 3.023 | 2.650 | 0.0656 |
| Overall | | 2737 | 2.316 | 3.023 | 2.636 | 0.0585 |

Profile observations from a single board × side × laser position combination for three saw configuration × side combinations (BB-Side 1-Board 001-Laser 1, BC-Side 1-Board 002-Laser 1, RR-Side 1-Board 012-Laser 1) are shown in Figure 4-6. These same three samples were arbitrarily selected as examples of each saw type (Bandsaw, Chipper-head, and Circular Saw, respectively) for this section because they were the first wane-free samples measured of each saw type. Since each saw type produced similar results, the remaining five combinations are not shown. The three data series show running patterns typical of autocorrelated time series data.

---

[23] See Chapters 2 and 3.

[24] The following Saw Configuration abbreviations are used in tables throughout the remainder of this chapter: BB=Bandsaw-Bandsaw, BC=Bandsaw-Chipper-head, CB=Chipper-head-Bandsaw, RR=Circular Saw-Circular Saw.

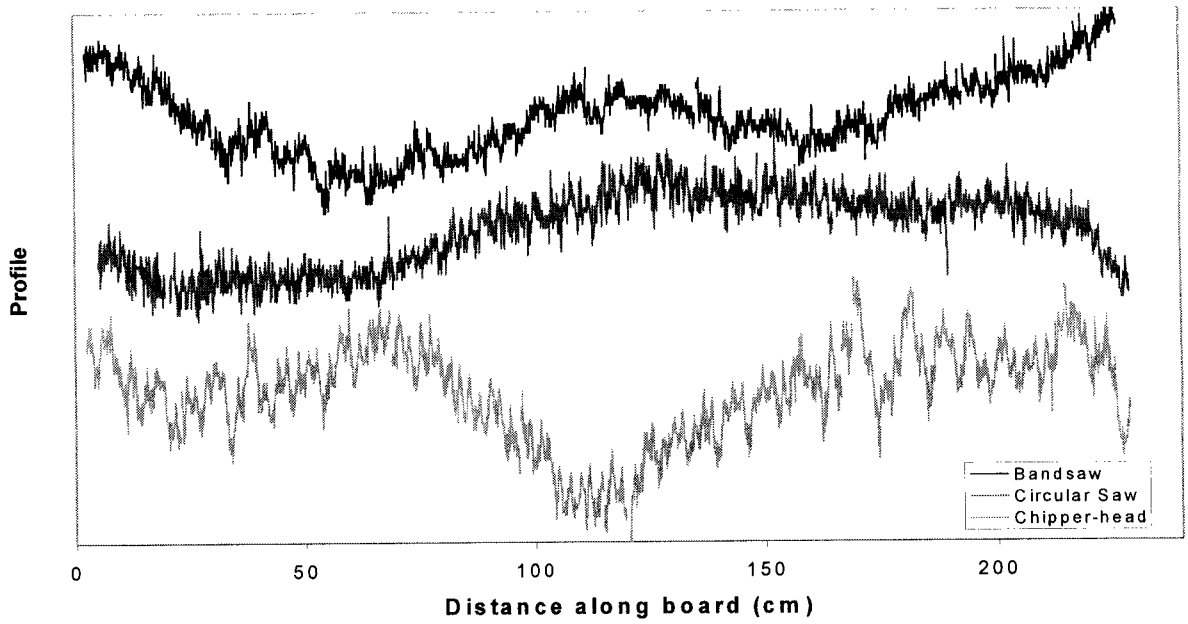



**Figure 4-6. Observations from a single board-side-laser position for each of three saw configuration × side combinations (BB-Side 1-Board 001-Laser 1, BC-Side 1-Board 002-Laser 1, RR-Side 1-Board 012-Laser 1).**

Plots of the sample ACF by saw type for the three samples are shown in Figure 4-7. The ACFs for the bandsawn and circular-sawn data were particularly slow to decay; the ACF was significantly non-zero ($\alpha = 0.05$) for more than 200 lags. The Chipper-head sample shows a somewhat less persistent pattern, with the ACF dying off to a non-significant level at 185 lags.

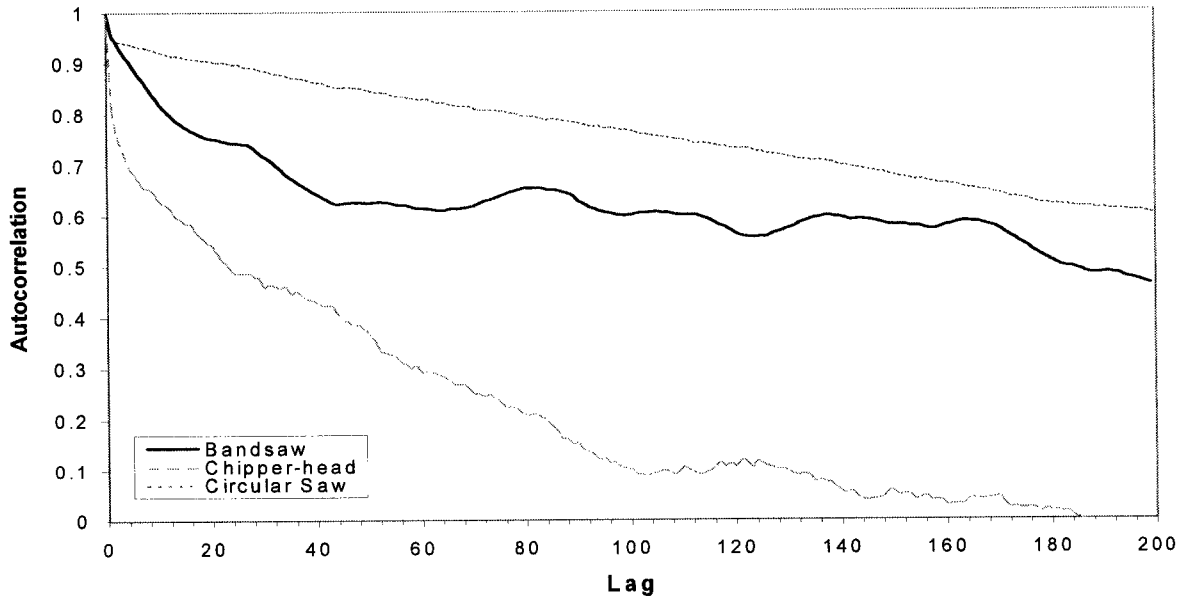



**Figure 4-7. ACF of a single board-side-laser position for each of three saw configuration × side combinations (BB-Side 1-Board 001-Laser 1, BC-Side 1-Board 002-Laser 1, RR-Side 1-Board 012-Laser 1).**

### 4.5.1 Model Ignoring Autocorrelation

Estimates of the mean and components of variance (COV) from [4-1] are listed in Table 4-3. These estimates are in the range of "typical" mill data; for example, Maness *et al.* (2004) used a range of 0.01 to 0.09 cm for the COV of board thicknesses in their SPC simulation study. All effects except that of laser position were significantly different from zero ($\alpha = 0.05$) for all saw configurations and sides. Since the interaction of boards and laser positions was significant, the effect of laser position was not consistent by board for each saw configuration and side combination. For the BB and RR saw configurations, the Side 1 and Side 2 estimates were very close. BC and CB saw configurations produced the most dissimilar estimates, especially for the residual variance. However, this estimated variance is likely biased due to the significant non-zero autocorrelation.

**Table 4-3. Estimated parameters (cm) by saw configuration and side for model ignoring autocorrelation.**

| | Saw configuration | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BB | | BC | | CB | | RR | |
| Parameter | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 |
| $\mu_{ij}$ | 2.575 | 2.575 | 2.632 | 2.634 | 2.656 | 2.655 | 2.651 | 2.650 |
| $\sigma_{\beta_{ij}}$ | 0.0204 | 0.0205 | 0.0289 | 0.0305 | 0.0328 | 0.0338 | 0.0403 | 0.0403 |
| $\sigma_{\lambda_{ij}}$ | 0.0052 | 0.0051 | 0.0083 | 0.0066 | 0.0048 | 0.0061 | 0.0174 | 0.0167 |
| $\sigma_{\beta\lambda_{ij}}$ | 0.0238 | 0.0238 | 0.0292 | 0.0280 | 0.0286 | 0.0274 | 0.0352 | 0.0341 |
| $\sigma_{\varepsilon_{ij}}$ | 0.0403 | 0.0395 | 0.0481 | 0.0529 | 0.0462 | 0.0368 | 0.0357 | 0.0404 |

Estimates of the variability of the profile data were derived using the components of variance, the number of sample boards ($b_i$=41 for BB, $b_i$=24 for BC and CB, and $b_i$=21 for RR), number of laser positions (2), and average number of measurements per board, side, and laser ($\bar{n}_{ij\cdot\cdot} \approx 2800$). For each saw configuration and side, the standard deviation of the mean by board and laser ($\bar{y}_{ijkl\cdot}$), mean by board ($\bar{y}_{ijk\cdot\cdot}$), and overall for the saw configuration × side combination ($\bar{y}_{ij\cdot\cdot\cdot}$) can be constructed as (following Neter *et al.* 1996):

$$\sigma_{\bar{y}_{ijkl\cdot}} = \sqrt{\mathrm{var}(\bar{\beta}_{ijkl\cdot}) + \mathrm{var}(\bar{\lambda}_{ijkl\cdot}) + \mathrm{var}(\overline{\beta\lambda}_{ijkl\cdot}) + \mathrm{var}(\bar{\varepsilon}_{ijkl\cdot})}$$
$$= \sqrt{\sigma^2_{\beta_{ij}} + \sigma^2_{\lambda_{ij}} + \sigma^2_{\beta\lambda_{ij}} + \mathrm{var}(\bar{\varepsilon}_{ijkl\cdot})} \qquad [4\text{-}12]$$

$$\sigma_{\bar{y}_{ijk\cdot\cdot}} = \sqrt{\sigma^2_{\beta_{ij}} + \frac{\sigma^2_{\lambda_{ij}}}{2} + \frac{\sigma^2_{\beta\lambda_{ij}}}{2} + \frac{\mathrm{var}(\bar{\varepsilon}_{ijkl\cdot})}{2}} \qquad [4\text{-}13]$$

$$\sigma_{\bar{y}_{ij\cdot\cdot\cdot}} = \sqrt{\frac{\sigma^2_{\beta_{ij}}}{b_i} + \frac{\sigma^2_{\lambda_{ij}}}{2} + \frac{\sigma^2_{\beta\lambda_{ij}}}{2b_i} + \frac{\mathrm{var}(\bar{\varepsilon}_{ijkl\cdot})}{2b_i}} \qquad [4\text{-}14]$$

When the $y_{ijklm}$ are independent and identically distributed, the variance of the average residuals by board and laser position is given by:

$$\mathrm{var}(\bar{\varepsilon}_{ijkl\cdot}) = \sigma^2_{\varepsilon_{ij}} / \bar{n}_{ij\cdot\cdot} \qquad [4\text{-}15]$$

This variance would be nearly zero for very large $\bar{n}_{ij\cdot\cdot}$. However, given the significant non-zero autocorrelations to 200 lags, [4-15] is a biased estimate of the variance of the average residual ($\bar{\varepsilon}_{ijkl\cdot}$). Assuming that there is no significant long-range dependence, this variance could be estimated using:

$$\mathrm{var}(\bar{\varepsilon}_{ijkl\cdot}) \cong \frac{\sigma^2_{\upsilon_{ij}}}{n_{ijkl}}\left(1 + \frac{2(n_{ijkl}-1)}{n_{ijkl}}\hat{\rho}_{ijkl_1} + \frac{2(n_{ijkl}-2)}{n_{ijkl}}\hat{\rho}_{ijkl_2} + \ldots\right) \qquad [4\text{-}16]$$

However, for very large $n_{ijkl}$, this also approaches zero. Table 4-4 shows estimates of the standard deviations for the profile data using [4-12] – [4-14] and assuming negligible variation from average residuals. Estimates were higher for the circular-sawn data versus chipped or bandsawn boards, and estimates on the chipped sides of BC and CB boards were higher than on the bandsawn sides.

**Table 4-4. Estimated standard deviations (cm) by saw configuration and side for model ignoring autocorrelation.**

| Standard Deviation | Saw Configuration | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BB | | BC | | CB | | RR | |
| | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 |
| $\sigma_{\bar{y}_{ijkl\cdot}}$ | 0.0318 | 0.0319 | 0.0419 | 0.0420 | 0.0438 | 0.0439 | 0.0563 | 0.0554 |
| $\sigma_{\bar{y}_{ijk\cdot\cdot}}$ | 0.0267 | 0.0268 | 0.0360 | 0.0367 | 0.0387 | 0.0392 | 0.0489 | 0.0485 |
| $\sigma_{\bar{y}_{ij\cdot\cdot\cdot}}$ | 0.0267 | 0.0268 | 0.0360 | 0.0367 | 0.0387 | 0.0392 | 0.0489 | 0.0485 |

Table 4-5 shows examples of predicted profile values ($\hat{y}_{ijklm}$) for several saw configurations and sides. The predictions were computed as the sum of the overall mean for the saw configuration, plus the effect by board, by laser, and by laser × board.

**Table 4-5. Examples of predicted profile values (cm) using fixed and random effects estimates for the model ignoring autocorrelation.**

| Saw Configuration | Sample Board | Side | Laser Position | $\hat{\mu}_{ij}$ | $\hat{\beta}_{ijk}$ | $\hat{\lambda}_{ijl}$ | $\hat{\beta\lambda}_{ijkl}$ | $\hat{y}_{ijklm}$ |
|---|---|---|---|---|---|---|---|---|
| Bandsaw-Bandsaw (BB) | 001 | 1 (B) | 1 (bottom) | 2.575 | -0.015 | 0.003 | -0.001 | 2.562 |
| | | | 2 (top) | 2.575 | -0.015 | -0.003 | -0.020 | 2.537 |
| Chipper-head-Bandsaw (CB) | 002 | 1 (C) | 1 (bottom) | 2.656 | -0.054 | -0.002 | -0.035 | 2.564 |
| | | | 2 (top) | 2.656 | -0.054 | 0.002 | -0.007 | 2.597 |
| Circular Saw-Circular Saw (RR) | 012 | 1 (R) | 1 (bottom) | 2.651 | -0.073 | 0.011 | -0.066 | 2.523 |
| | | | 2 (top) | 2.651 | -0.073 | -0.011 | 0.011 | 2.578 |

Residuals were computed as the difference between the actual profile value and the predicted values (from Table 4-5). Plots of the residuals for several saw configuration × side combinations are shown in Figure 4-8 - Figure 4-10. Residual plots for all bandsawn sides (e.g., BB-Side 1 and BB-Side 2, CB-Side 2 and BC-Side 1) were very similar, and thus only one example is shown (Figure 4-8). Plots of both chipped sides and both circular-sawn sides (CB-Side 1 and BC-Side 2, and RR-Side 1 and RR-Side 2) were also very similar, and only one example or each saw type is shown (Figure 4-9 and Figure 4-10). The residuals from all saw configurations and sides show distinctive wave patterns, indicating that Model [4-1] with uncorrelated errors may be inappropriate.

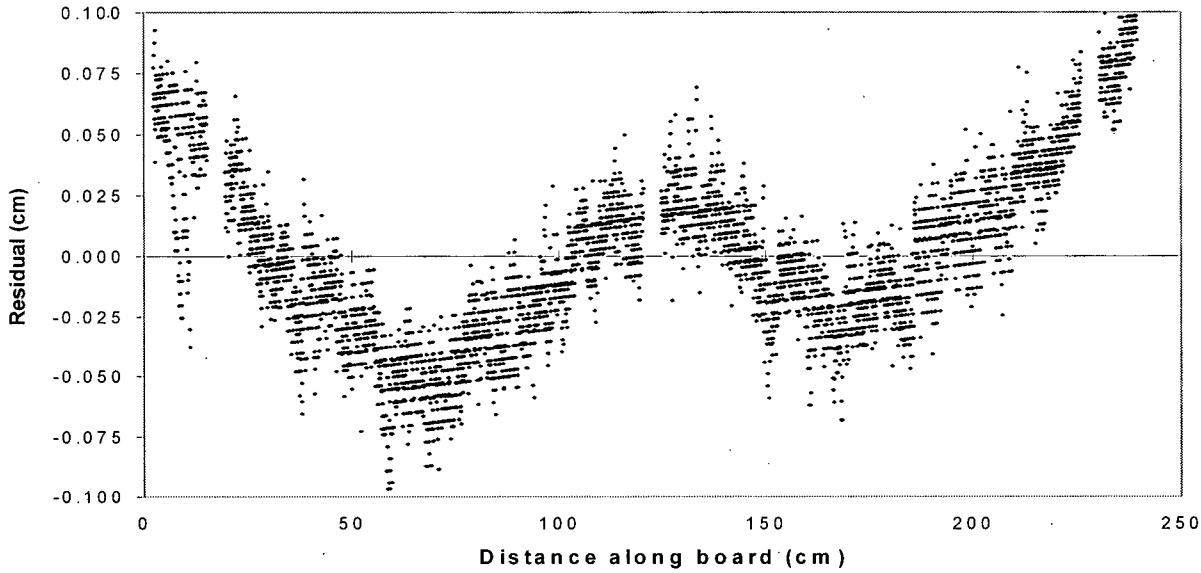



**Figure 4-8. Residuals from model ignoring autocorrelation for Saw Configuration BB-Side 1-Board 001.**

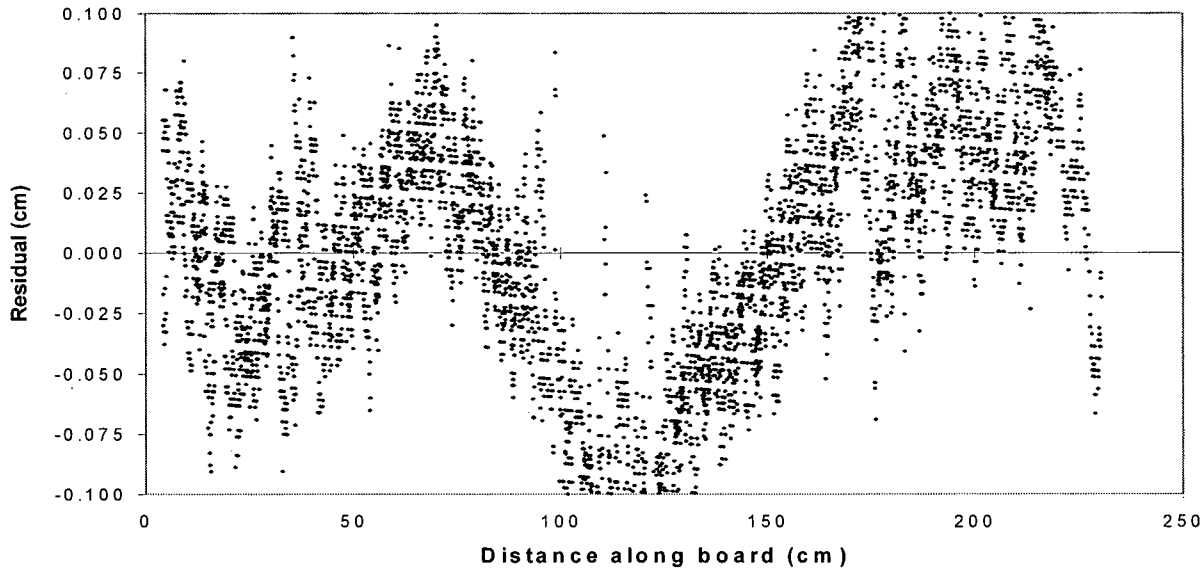



**Figure 4-9. Residuals from model ignoring autocorrelation for Saw Configuration CB-Side 1-Board 002.**

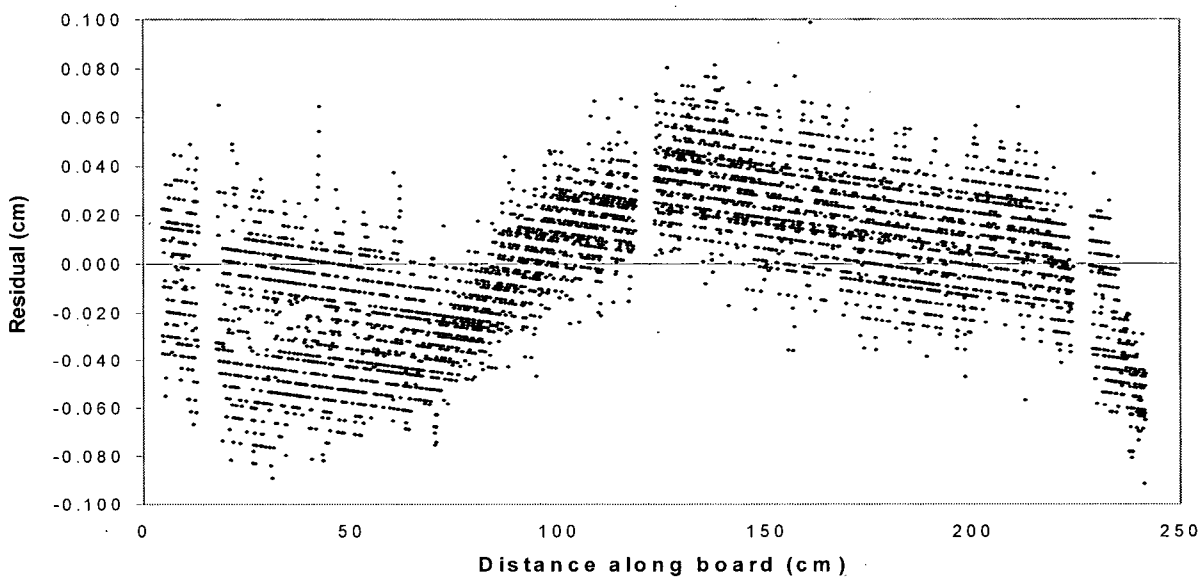



**Figure 4-10. Residuals from model ignoring autocorrelation for Saw Configuration RR-Side 1-Board 012.**

## 4.5.2 Multi-Stage Model

### 4.5.2.1 Stage 1: Within Sample Board Model

The estimated ARIMA parameters for the bandsawn data and circular-sawn data without saw marks are shown in Table 4-6. Statistical tests indicated residuals had significant non-zero autocorrelation ($\alpha = 0.05$) in about 30% of the series; however, the magnitude of the significant correlations was less than 0.05 and the number of observations was large ($n_f \approx 2800$). Therefore, significant correlations were not considered to be of practical importance. The estimate of the parameter $\alpha_f$ was significantly different from zero ($\alpha$=0.05) for nine of 306 (< 3%) series, indicating that linear trend could be assumed zero for all series. The estimated values of $\theta_f$ were significantly different from zero ($\alpha$=0.05) and positive in all cases, indicating a strong mixing process. Estimates of $\phi_f$, on the other hand were slightly more varied.

**Table 4-6. Summary of estimated ARIMA(1,1,1) model [4-5] parameters (mm) for bandsawn data (188 series) and ARIMA(0,1,1) model [4-6] parameters (mm) for circular-sawn data without saw marks (118 series).**

| | Bandsaw | | | | Circular Saw | | |
|---|---|---|---|---|---|---|---|
| | $\hat{\alpha}_f$ | $\hat{\theta}_f$ | $\hat{\phi}_f$ | $\hat{\sigma}_{\upsilon_f}$ | $\hat{\alpha}_f$ | $\hat{\theta}_f$ | $\hat{\sigma}_{\upsilon_f}$ |
| Average | 0.0000 | 0.7027 | 0.2101 | 0.0967 | 0.0001 | 0.5527 | 0.1009 |
| Minimum | -0.0032 | 0.3611 | -0.0893 | 0.0619 | -0.0013 | 0.2520 | 0.0688 |
| Maximum | 0.0031 | 0.8884 | 0.5859 | 0.1385 | 0.0018 | 0.7536 | 0.1366 |
| % Significant | 5% | 100% | 94% | | 0% | 100% | |

Estimates of the SARFIMA model parameters for data from chipped boards and circular-sawn boards with saw marks are shown in Table 4-7. Cycle lengths for most of the circular-sawn and chipped boards were 22-24 measurements, which corresponds to a distance along the board of approximately 2 cm (0.8 inch). Estimates of $\delta_f$ and $\delta'_f$ were wide-ranging; however, they were similar for chipped and circular-sawn boards. On average, most values of $\delta_f$ and $\delta'_f$ were greater than 0.5, indicating that the best-fitting models were non-stationary with unbounded variance. The estimated values of $\theta_f$, $\Theta_f$, $\phi_f$ and $\Phi_f$ were significantly different from zero ($\alpha = 0.05$) for almost all series. The estimated values of $\theta_f$ and $\Theta_f$ were positive in all cases, indicating a strong mixing process; estimates of $\phi_f$ and $\Phi_f$ were slightly more varied. The estimates of $\alpha_f$ were significantly different than zero more often than with the bandsawn data, indicating a very small, but significant thickening or thinning along the length of the board in 14% of the sample series.

**Table 4-7. Summary of estimated SARFIMA(1,$\delta_f$,1)x(1,$\delta'_f$,1)$_s$ model [4-7] parameters (mm) for chipped data (96 series) and circular-sawn data with saw marks (38 series).**

| | $\hat{s}_f$ | $\hat{\delta}_f$ | $\hat{\delta}'_f$ | $\hat{\alpha}_f$ | $\hat{\phi}_f$ | $\hat{\Phi}_f$ | $\hat{\theta}_f$ | $\hat{\Theta}_f$ | $\hat{\sigma}_{\upsilon_f}$ |
|---|---|---|---|---|---|---|---|---|---|
| Chipper-head | | | | | | | | | |
| Average | 24 | 0.677 | 0.207 | 0.0001 | -0.1589 | -0.0972 | 0.7045 | 0.7880 | 0.1425 |
| Minimum | 22 | 0.304 | -0.227 | -0.0006 | -0.3217 | -0.3489 | 0.4284 | 0.5102 | 0.0891 |
| Maximum | 50 | 1.000 | 0.941 | 0.0010 | 0.0791 | 0.1290 | 0.9161 | 0.8911 | 0.2315 |
| % Significant | | 100% | 79% | 13% | 90% | 73% | 100% | 100% | |
| Circular Saw | | | | | | | | | |
| Average | 23 | 0.587 | 0.353 | 0.0000 | -0.1259 | -0.1432 | 0.7647 | 0.7871 | 0.1381 |
| Minimum | 22 | 0.308 | 0.081 | -0.0006 | -0.2905 | -0.2369 | 0.5468 | 0.5715 | 0.0847 |
| Maximum | 48 | 0.975 | 0.555 | 0.0004 | 0.1127 | -0.0502 | 0.9273 | 0.8948 | 0.1788 |
| % Significant | | 100% | 100% | 11% | 95% | 95% | 100% | 100% | |

Cross-correlations were examined to quantify the relationship between the bottom (Laser 1) and top (Laser 2) laser positions. Figure 4-11 compares the cross-correlation of the original first-differenced series with those of the ARIMA model residuals for a bandsawn sample (BB-Side 1). Dotted lines indicate 99% confidence intervals for the cross-correlations, and thus about 30 measurements per board are expected to be outside these lines. Although there were non-zero ($\alpha = 0.05$) cross-correlations at some lags, this is not unexpected given the large number of observations. More important, the ARIMA model did not generate any obvious pattern, which would indicate a cross-correlation between laser positions within board and side.

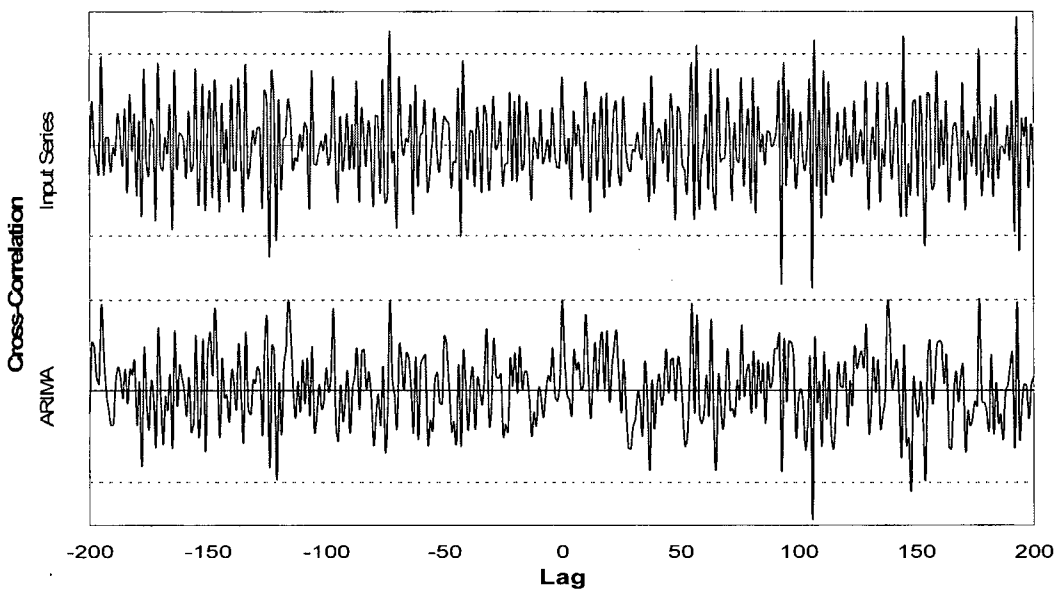



**Figure 4-11. Cross-correlation of Laser 1 versus Laser 2 for Saw Configuration BB-Side 1-Board 001.**

Cross-correlations under the SARFIMA model are shown in Figure 4-12 and Figure 4-13 for chipped and circular-sawn samples, respectively (CB-Side 1 and RR-Side 1). No patterns were obvious for either saw type, and therefore no cross-correlation between laser positions was indicated.

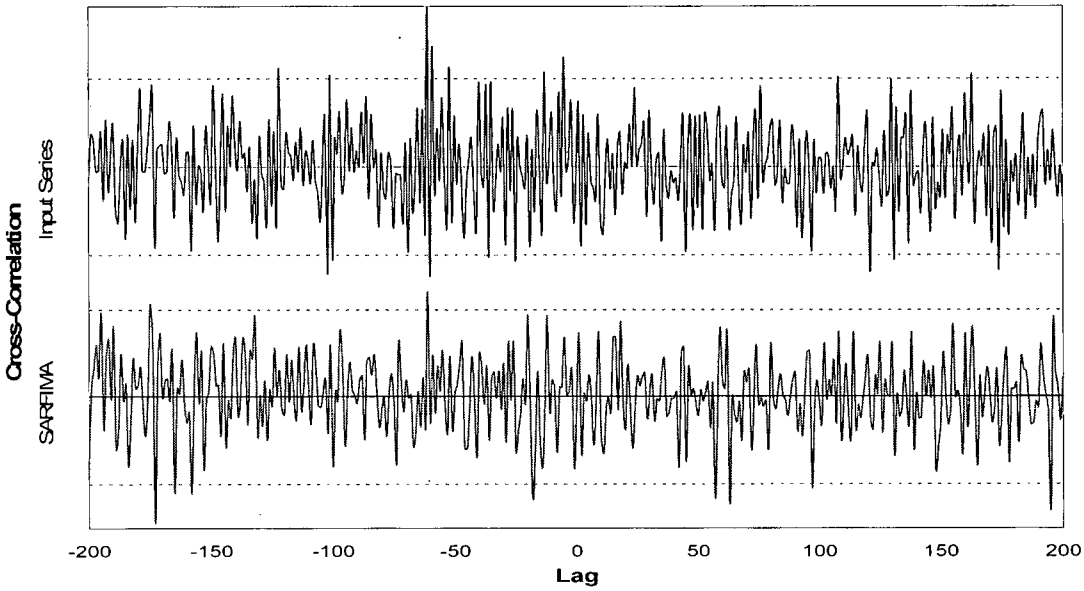



**Figure 4-12. Cross-correlation of Laser 1 versus Laser 2 within board and side for Saw Configuration CB-Side 1-Board 002.**

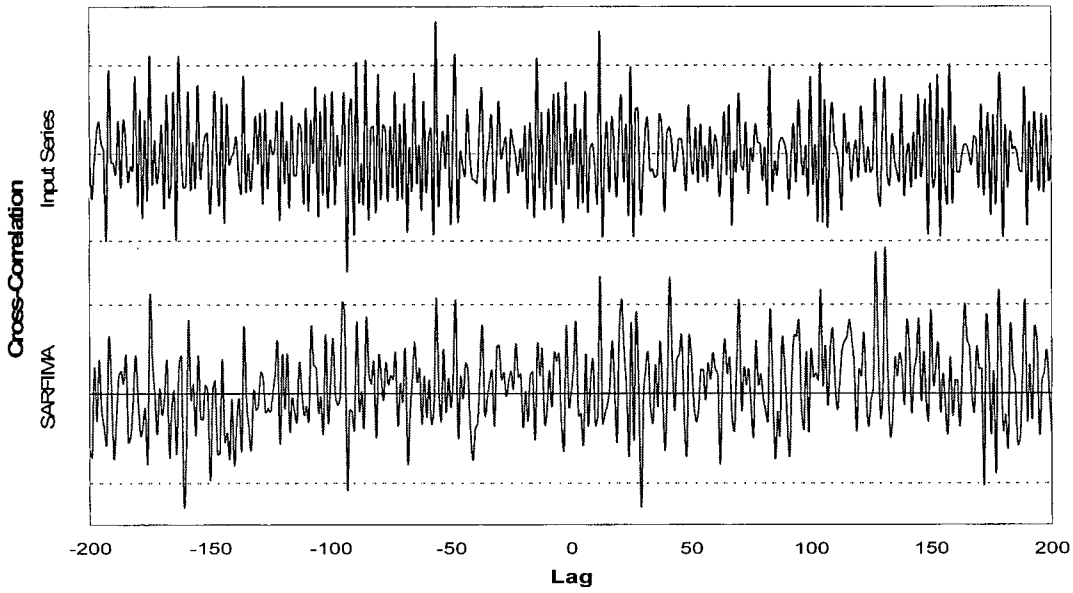



**Figure 4-13. Cross-correlation of Laser 1 versus Laser 2 within board and side for Saw Configuration RR-Side 1-Board 012.**

## 4.5.2.2 Stage 2: Between Sample Boards Model

The series of average cant profiles from the field scan data are plotted in Figure 4-14. Although the 100$^{th}$ and 179$^{th}$ scans were unusual, they are representative of the normal range of variation in cant profile data. There is no apparent trend or pattern indicative of an autocorrelated series.

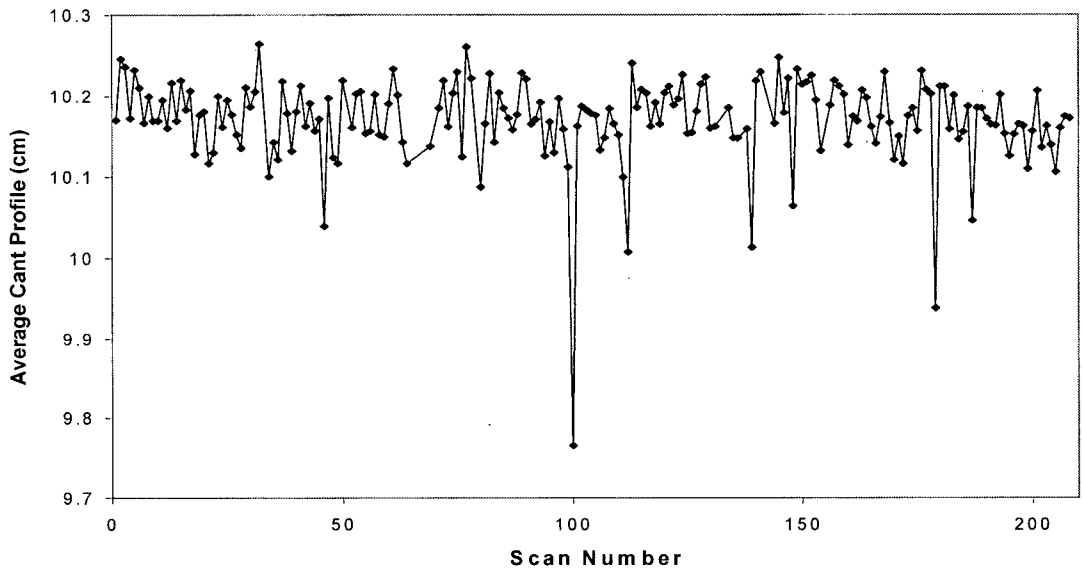



**Figure 4-14. Average cant profiles by scan number for field scan data.**

Lack of autocorrelation was confirmed by results from PROC ARIMA and the ACF of the field scanned series appeared to be white noise (Figure 4-15). The model form with the lowest BIC was an ARIMA(0,0,0), indicating a model without any time series parameters was the best-fitting. Thus, no adjustment to the residuals for between-board correlations was made before estimating the mixed-model effects.

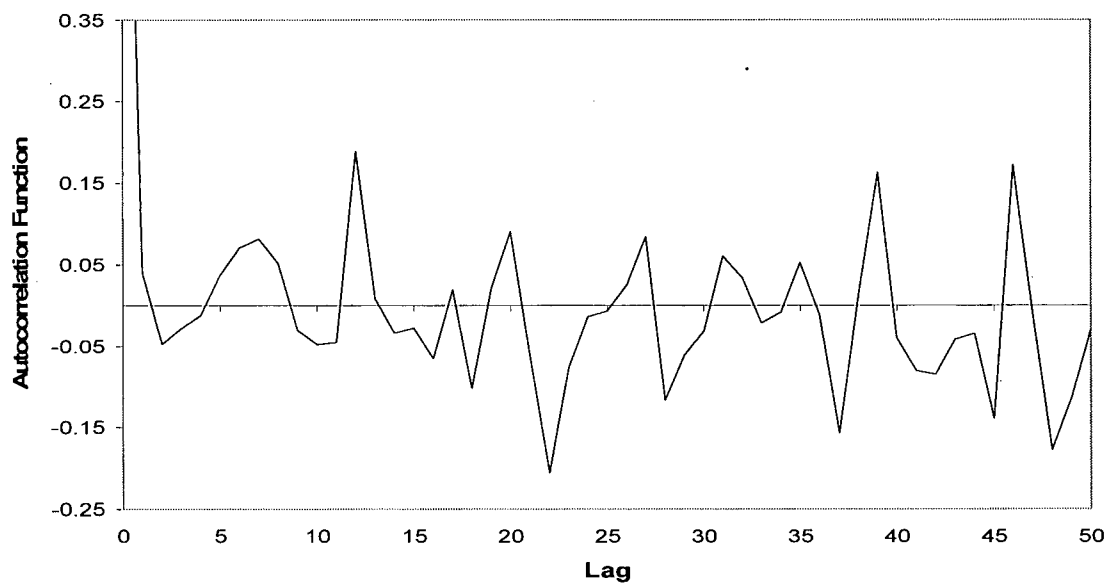



**Figure 4-15. ACF of average cant profiles for field scan data.**

### 4.5.2.3 Stage 3: Mixed-effects Model

The estimated mean and COV for the multi-stage model are listed in Table 4-8. All estimated COV were nearly identical to those of the model ignoring autocorrelation, except for the residual variation, which was considerably smaller. However, the estimates from the two models are not directly comparable because the model ignoring autocorrelation estimated the residual variance as $\sigma^2_{\varepsilon_{ij}}$, whereas the multi-stage model estimated the residual variance as $\sigma^2_{\nu_{ij}}$. As in the model ignoring autocorrelation, all effects except laser were significantly different from zero ($\alpha$ = 0.05).

**Table 4-8. Estimated parameters (cm) by saw configuration and side for multi-stage model.**

| | Saw Configuration | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | BB | | BC | | CB | | RR | |
| Parameter | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 |
| $\mu_{ij}$ | 2.575 | 2.575 | 2.632 | 2.634 | 2.656 | 2.655 | 2.651 | 2.650 |
| $\sigma_{\beta_{ij}}$ | 0.0204 | 0.0205 | 0.0289 | 0.0304 | 0.0326 | 0.0337 | 0.0403 | 0.0403 |
| $\sigma_{\lambda_{ij}}$ | 0.0052 | 0.0052 | 0.0083 | 0.0066 | 0.0050 | 0.0061 | 0.0173 | 0.0167 |
| $\sigma_{\beta\lambda_{ij}}$ | 0.0238 | 0.0238 | 0.0292 | 0.0281 | 0.0287 | 0.0274 | 0.0351 | 0.0340 |
| $\sigma_{\nu_{ij}}$ | 0.0095 | 0.0095 | 0.0103 | 0.0147 | 0.0142 | 0.0098 | 0.0112 | 0.0125 |

Estimates of the variability computed with [4-12] – [4-14] are shown in Table 4-9. As in the model ignoring autocorrelation, the contribution of the variation from average residuals was negligible. While the estimation of $\sigma^2_{\nu_{ij}}$ in the multi-stage model (versus $\sigma^2_{\varepsilon_{ij}}$ in the model ignoring autocorrelation) prevents exact comparisons between the two models, the negligible residual variation makes estimates of the standard deviations of the means effectively comparable. Estimates of the various standard deviations of the means under the multi-stage model were nearly identical to those of the model ignoring autocorrelation; the standard deviations were higher for circular-sawn boards versus chipped and bandsawn boards, and estimates on the chipped sides of BC and CB boards were higher than on the bandsawn sides.

**Table 4-9. Estimated standard deviations by saw configuration and side for multi-stage model (cm).**

| Standard Deviation | BB | | BC | | CB | | RR | |
|---|---|---|---|---|---|---|---|---|
| | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 | Side 1 | Side 2 |
| $\sigma_{\bar{y}_{ijkl\cdot}}$ | 0.0318 | 0.0319 | 0.0419 | 0.0419 | 0.0437 | 0.0439 | 0.0561 | 0.0553 |
| $\sigma_{\bar{y}_{ijk\cdot\cdot}}$ | 0.0267 | 0.0268 | 0.0360 | 0.0366 | 0.0386 | 0.0391 | 0.0489 | 0.0484 |
| $\sigma_{\bar{y}_{ij\cdot\cdot\cdot}}$ | 0.0267 | 0.0268 | 0.0360 | 0.0366 | 0.0386 | 0.0391 | 0.0489 | 0.0484 |

Examples of predicted profile values are not shown, as they are nearly identical to that of the model ignoring autocorrelation (Table 4-5). Plots of the residuals are shown in Figure 4-16 - Figure 4-18 for one sample from each saw type. Since, all "B" sides, "C" sides, and "R" sides from each type of saw configuration were very similar, not all combinations are shown. For the circular-sawn and chipped boards, residuals were not computed for the first ~100 observations, as using the SARFIMA model form with seasonality length > 10 requires a substantial number of initial observations. These residuals appeared to be free of any pattern, and the assumption that the errors ($\nu_{fm}$) were free of autocorrelation appeared to be valid.

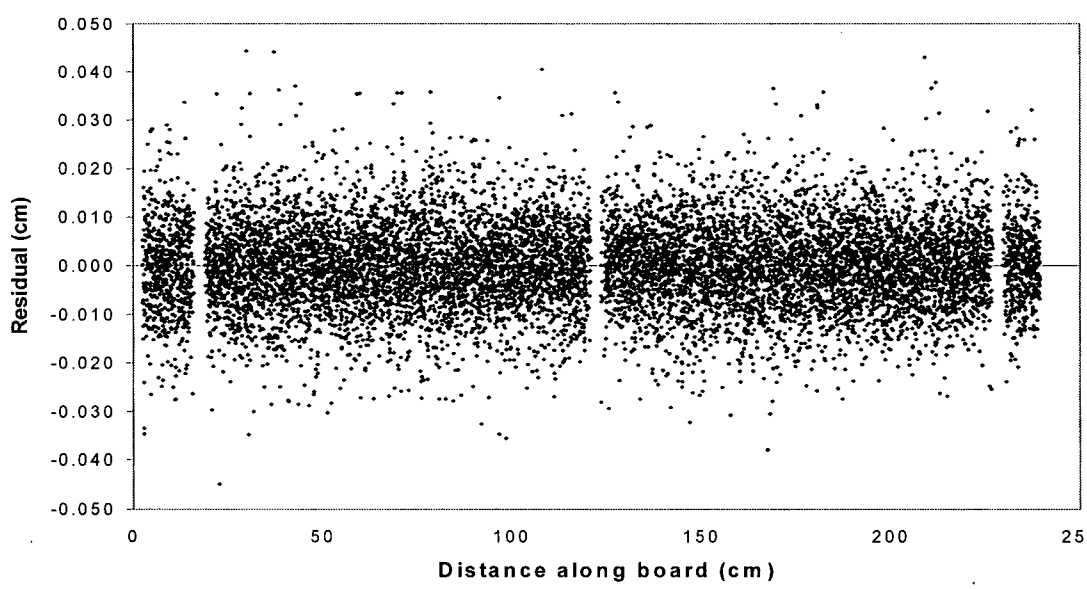



**Figure 4-16. Residuals from multi-stage model for Saw Configuration BB-Side 1-Board 001.**

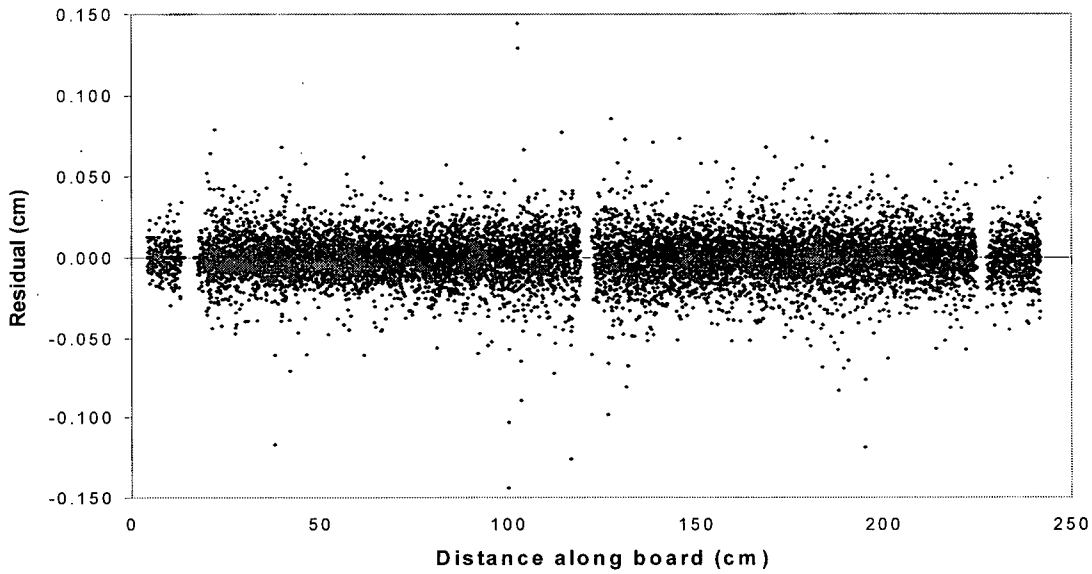



**Figure 4-17. Residuals from multi-stage model for Saw Configuration CB-Side 1-Board 002.**

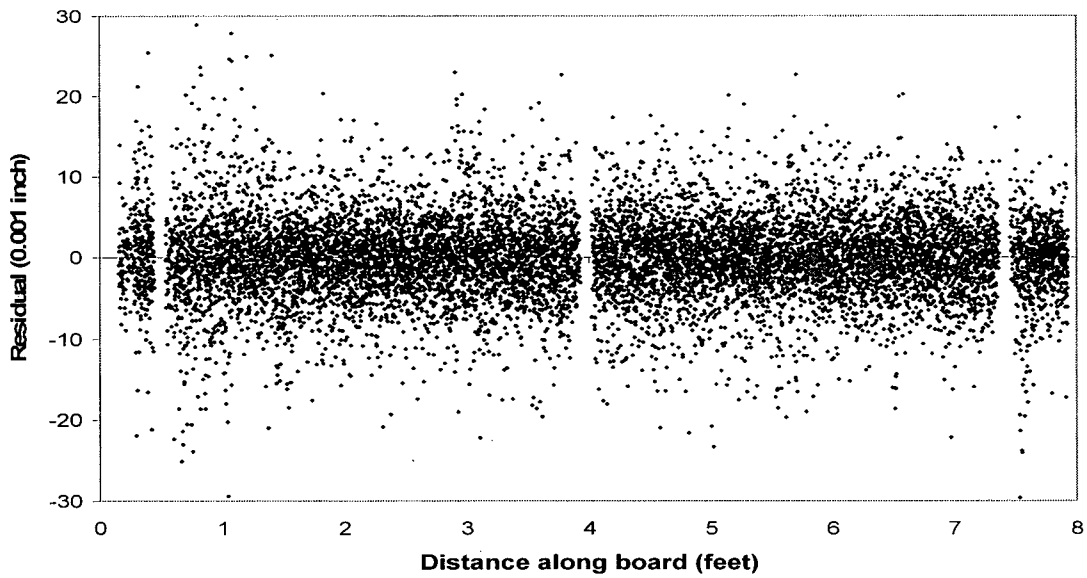



**Figure 4-18. Residuals from multi-stage model for Saw Configuration RR-Side 1-Board 012.**

## 4.5.3 Comparison of Models

The model ignoring autocorrelation and the multi-stage model were compared using Akaike's Information Criteria (AIC) (Table 4-10). In all cases, the model with the lowest AIC was the multi-stage model. Thus, accounting for the autocorrelative structure of the errors reduced the amount of unexplained variation in the model.

**Table 4-10. Comparison of Akaike's Information Criteria (AIC) for model ignoring autocorrelation and multi-stage model.**

| Saw Configuration | Side | Model Ignoring Autocorrelation | Multi-stage Model |
|---|---|---|---|
| Bandsaw–Bandsaw (BB) | 1 (B) | 1,897,850 | 1,242,197 |
| | 2 (B) | 1,903,131 | 1,238,285 |
| Bandsaw–Chipper-head (BC) | 1 (B) | 1,157,610 | 770,185 |
| | 2 (C) | 1,151,176 | 740,684 |
| Chipper-head–Bandsaw (CB) | 1 (C) | 1,070,338 | 721,144 |
| | 2 (B) | 1,104,387 | 748,187 |
| Circular Saw–Circular Saw (RR) | 1 (R) | 977,059 | 671,419 |
| | 2 (R) | 930,981 | 653,388 |

## 4.6 Discussion

The objective of this research was to derive a statistical model to describe real-time laser measurements, which could then be used as the basis for an SPC system for automated lumber manufacturing. This model differs from traditional models for SPC data in that there are several identifiable sources of variation and the data suggest a large and significant autocorrelative structure.

Estimation of the model components of variation was undertaken with and without considering the autocorrelative structure of the errors. Although most of the parameter estimates were nearly identical, important differences were observed. Under the model ignoring autocorrelation, residuals exhibited strong cyclical patterns and there was large and significant non-zero autocorrelation. Using a multi-stage model that quantified and isolated this autocorrelation, residuals were without pattern and not significantly different from zero. Since independent errors are necessary for statistical tests in mixed-effects modeling, the multi-stage model should be used for testing.

Fitting a multi-stage model also provided insight into the sawing process, as the autocorrelative structure was different depending on the saw type. Long and short-term cyclical behaviour was prominent in the Chipper-head data and the Circular-Saw data when saw marks were present. Whereas certain cycles are normal and expected in Chipper-head data, cycles in circular-sawn

data could indicate washboarding or deep saw marks. These sawing defects occur when a saw loses stiffness due to heat and indicate a need for saw maintenance, such as saw tensioning (Schajer 1989). Finding these cycles may help in diagnosing maintenance problems.

On the other hand, the model ignoring autocorrelation is appealing because it does not require multiple steps, and would therefore be much easier to fit if process parameters needed to be updated. Moreover, violation of the statistical assumptions does not invalidate the model as a descriptive tool. While autocorrelation may cause estimates of the error variance to be biased, parameter estimators, such as the mean, are unbiased (Schabenberger and Pierce 2001). This was shown by comparing the COV estimates in the model ignoring autocorrelation to that of the multi-stage model. Estimates of the mean were identical by model form, as were all COV estimates except that of the residual. Although the residual variances were not directly comparable, the standard deviations were comparable in practical terms, and these estimates were found to be nearly identical. This result was not unexpected, since the only significant autocorrelation in the model was within each board × side × laser position.

The real-time LRS data will be used primarily in constructing control charts for monitoring sawing performance. Because of the abundance of real-time data, individual observations cannot be monitored. Instead, averages and trends by board, or other measures of surface profile variation must be used. Alternative methods to monitor the surface profiles could be derived from fitting the within sample board model. For instance, as boards were scanned, the parameters of each board × side × laser model could be estimated and compared to some standard. Spectral analysis could also be used to look for changing cycles in the data. While the cycles may help to uncover maintenance issues for the circular saws, the ARIMA and SARFIMA parameters are less helpful. While ARIMA and SARFIMA models were found to adequately describe the profile data and quantify its autocorrelative structure, neither model form is good for

prediction. Because these models are non-stationary, models with the same parameter estimates can look quite different, and thus ARIMA and SARFIMA parameters are not particularly useful in terms of SPC.

Measures, such as the average profile by board or the average simple linear trend over a board are more useful for SPC applications. Since control limits constructed for these quantities rely on standard deviations of means (e.g., by board and side, or by board, side, and laser position), the model ignoring autocorrelation will be adequate to describe the data for these purposes. As such, it is recommended for SPC applications that rely only on these quantities.

An important step in this analysis was verifying the lack of significant autocorrelation between subsequent boards. This finding was contrary to expected, given a basic understanding of sawing machines and the mechanisms that control them. On the other hand, the sawn logs are an arbitrary sample of fiber, and may represent different growing conditions, moisture contents, and wood densities. Different tapers and log shapes cause operators to make different decisions, and the setworks for the saw are reset for each log. It may be that the random inputs to the process are strong enough to prevent substantial autocorrelation during "normal" operating conditions. Moreover, a significant autocorrelation between boards could be an indicator of quality problems, as machines that become out of adjustment tend to stay out of adjustment.

## 4.7 Conclusions

The development of an SPC protocol for real-time systems presents numerous challenges in data collection, filtering, and analysis. The objective of this chapter was to develop a model that adequately described LRS data in order to use it in a SPC protocol. The recommended model for SPC data applications is an uncorrelated errors model; that is, a model ignoring autocorrelation,

with parameters for the effects of laser position and sample boards and their interactions, and model parameters estimated for each type of saw configuration and side.

In subsequent steps in this research, estimates of the components of variance are important for monitoring the consistency of the sawing process. For example, the setworks of the saw may be monitored by tracking the average profile for each sample board and comparing it to control limits which are constructed with the standard error of the mean sample board. Other sawing defects may be monitored by tracking the individual components of variance. For instance, wedge[25] results in a significant difference between the top and bottom laser positions, and would be indicated by high within-laser variation.

Future research should include a dataset which has information about the specific saws that were engaged during the cut. Although this will require designing an interface with the programmable logic controllers used to control the saws, it will increase the usefulness of the model, as data can be tied to specific saws, rather than specific saw configurations. Data that includes sawing defects should also be collected. For instance, the cycles present in the saw marks from circular-sawn boards could be investigated. Cycle detection via spectral analysis may provide an early warning for saw maintenance concerns and prevent the waste associated with excess washboarding. Moreover, this analysis could be performed without fitting the multi-stage model.

Further research should also include more field sample data. For example, autocorrelation was not significant between sample boards, but it may be an important factor in identifying quality

[25] Wedge is a sawing defect characterized by a gradual thinning (or thickening) across the width of a board or through its thickness. See section 1.2.1 for more detail.

problems in the mill. This possibility should be investigated with a sample taken during a time of saw malfunction.

## 4.8 Literature Cited


Andel, J. 1986. Long memory time series models. Kybernetika 22(2): 105-123.

Box, G. E. P., G. M. Jenkins and G. C. Reinsel. 1994. Time Series Analysis: Forecasting and Control. 3. New York, Prentice Hall.

Burman, H. 2002. Laser strip adjustment for data calibration and verification. Presented at Symposium 2002: ISPRS Commission III. September 9-13, Graz, Austria.

Gazzarri, J. 2003. Sequential measurements method for moving surfaces profiling. MSc Thesis. Department of Mechanical Engineering, The University of British Columbia, Vancouver, Canada.

Geweke, J. and S. Porter-Hudak. 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4(4): 221-238.

Maness, T. C., R. A. Kozak and C. L. Staudhammer. 2004. Reliability testing of statistical process control procedures for manufacturing with multiple sources of variation. Wood and Fiber Science 36(3): 443-458.

Maness, T. C., C. L. Staudhammer and R. A. Kozak. 2002. Statistical considerations for real-time size control systems in wood products manufacturing. Wood and Fiber Science 34(3): 476-484.

Montgomery, D. C. 2001. Introduction to Statistical Quality Control. 4th ed. New York, John Wiley & Sons.

Neter, J., M. H. Kutner, W. Wasserman and C. J. Nachtsheim. 1996. Applied Linear Statistical Models. 4th ed. Chicago, Irwin.

Noffsinger, J. R. and R. B. Anderson. 2002. Effect of autocorrelation on composite panel production monitoring and control: a comparison of SPC techniques. Forest Products Journal 52(3): 60-67.

SAS Institute. 2002. SAS/STAT User's Guide, Version 8. Cary, NC, SAS Publishing.

Schabenberger, O. and F. J. Pierce. 2001. Contemporary Statistical Models for the Plant and Soil Sciences. Boca Raton, CRC Press.

Schajer, G. S. 1989. The importance of circular saw tensioning. Forest Industries/World Wood Technical Report (Sawing Technology): T14-T16.

Schajer, G. S. 1990. Designing a saw that works involves many factors. Forest Industries 117(9): 17-20.

Shewhart, W. A. 1931. Economic Control of Quality of Manufactured Product. New York, NY, Van Nostrand.

Warren, W. G. 1973. How to calculate target thickness for green lumber. Department of the Environment, Canadian Forestry Service, Western Forest Products Laboratory, Vancouver, BC.

Wehr, A. and U. Lohr. 1999. Airborne laser scanning - and introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing 54: 68-82.

Wheeler, D. J. 1995. Advanced Topics in Statistical Process Control. Knoxville, TN, SPC Press.

Young, T. M. and P. M. Winistorfer. 2001. The effects of autocorrelation on real-time statistical process control with solutions for forest products manufacturers. Forest Products Journal 51(11/12): 70-77.

# Chapter 5 SPC Methods for Real-Time Laser Range Sensor Data in Lumber Manufacturing

## 5.1 Introduction

For more than three decades, Shewhart control charts (Shewhart 1931) have helped automated lumber manufacturers to monitor the sawing process and produce lumber to consistent size standards. In order to successfully apply Shewhart's methods, process data must meet three assumptions: normality, independence, and homogeneity of variance (Mastrangelo *et al.* 2001). Under typical mill conditions, statistical process control (SPC) is conducted manually; a small group of boards is taken from the sawing process at infrequent time intervals and measured with digital calipers. Under these conditions, the three assumptions are met. The success of SPC programs has led to their widespread use in modern sawmills, and lumber manufacturers can directly attribute tremendous cost savings to their SPC practices (Maness 1993; Young and Winistorfer 1999).

New technologies for SPC in lumber manufacturing include laser range sensors (LRS), which make real-time measurement of the sawing process possible. When set up in-line with sawing machines, each piece of lumber processed is measured at a very fine scale, making thousands of measurements per sawn piece available. Moreover, these systems can be set up with multiple LRSs, enabling data collection on each side of each board. This is of particular interest in modern mills, where it is standard to cut the opposing sides of each board with different saws.

Many mills are now implementing real-time scanning technologies; however, SPC methods have not been updated to reflect the sampling frequency or the capacity of this new technology. Mills using LRSs have anecdotally reported that control limits must be set manually in order to prevent false out-of-control signals from overwhelming their systems. This is not surprising; the current SPC methods do not transfer directly to this new real-time technology, as the statistical model which describes the real-time data is different from that of manual sampling. Moreover, SPC

methods have not been updated to take advantage of the opportunity to better describe the sawing process with the additional data available.

A statistical model describing real-time LRS measurements taken from multiple boards and multiple surfaces was derived in Chapter 4. While the usual statistical model for SPC contains components for within- and between-board variation, the LRS model contains additional components of variance from laser positions and the interaction between boards and laser positions. These components are the basis of the SPC system for monitoring quantities such as the average board size. Because real-time measurement systems take a very large number of autocorrelated observations on each side of each board, many of the usual inferential statistics associated with SPC charts are not appropriate. Control charts must be specifically developed to take into account this autocorrelation (Montgomery and Mastrangelo 1991), and where appropriate, alternative measures, such as control charts for dependent and/or non-normal data (Padgett and Spurrier 1990; Grimshaw and Alt 1997) or those based on empirical quantiles (Willemain and Runger 1996) need to be developed.

With the additional data from this technology, there is a great opportunity to more thoroughly monitor the sawing process. Systems can be designed to target known causes of sub-standard product by identifying specific sawing defects. For example, Rasmussen *et al.* (2004) documented five sawing defects that may be identifiable with laser scanning technology. Using multiple LRSs, an SPC system for real-time data has the capacity to better describe the sawing process and prevent the production of lumber with specific kinds of defects.

## 5.2 Objective

The objective of this research is to develop SPC systems based on a statistical model of the sequence of real-time LRS measurements. This system would be used to monitor the sawing process, targeting specific kinds of defects common to automated sawmills.

## 5.3 Background

### 5.3.1 Identifying Sawing Defects with SPC

Five common sawing defects were described in detail in Section 1.2.1 (Chapter 1):

1. Machine positioning (or setworks) problems are characterized by a constant deviation in board thickness along the length of a board;
2. Wedge is a gradual thinning (or thickening) across the width of a board;
3. Taper is a gradual increase (or decrease) in thickness along the length of a board;
4. Flare/snipe describe a condition in which a triangular section is added to/removed from the end of a board; and
5. Snake is characterized by an uneven wave pattern on the surface of a board.

Common SPC methods used in automated lumber manufacturing were described in Section 1.2.2 (Chapter 1). Groups of boards are sampled periodically and four charts are routinely used:

1. X-bar chart for monitoring average board thickness;
2. $S_b$ chart for monitoring the between-board variability;
3. $S_w$ chart for monitoring the within-board variability; and
4. R chart for monitoring the range of grouped board thickness averages.

With current methods, only some of the five defects are readily recognizable. If machine positioning problems result in consistent differences from target, they are detected by a shift in the X-bar chart. If these problems are inconsistent, i.e., different from board to board, they will be detected in a chart for between-board variation. Since data from each board are grouped together without regard to the location of the measurements along the board, other sawing defects, such as snake, could easily be undetected by an X-bar chart. However, all of these

defects would likely produce a signal on an $S_w$ or R chart. It would then require further investigation on the part of SPC personnel to determine the exact cause of the chart's signal.

### 5.3.2 Real-time SPC Data from Laser Range Sensors

A real-time scanning apparatus can be configured to scan both "face" sides of each board as it leaves a sawing machine. A measurement apparatus used to scan sample boards in a laboratory setting was described in Section 4.4 (Chapter 4). This system mimicked commercially available real-time systems, using four point laser range sensors (two stacked on each side of the board) to measure the two board surfaces simultaneously. This resulted in four streams of measurement from each board scanned (Figures 4-1 and 4-5).

A statistical model describing the real-time LRS measurements taken from multiple sawing configurations, boards, sides, and laser positions was derived in Chapter 4. This mixed-effects model allowed for different variance components for each saw configuration × side combination. The profile observations, $y_{ijklm}$, from the $i$th saw configuration ($i$ = 1 to 4), $j$th side ($j$ = 1 to 2), $k$th sample ($k$ = 1 to $b_i$), $l$th laser location ($l$ = 1 to 2), and $m$th distance along the board ($m$ = 1 to $n_{ijkl}$) were modeled with random effects for boards ($\beta_{ijk}$), laser positions ($\lambda_{ijl}$), and the interaction of boards and laser positions ($\beta\lambda_{ijkl}$).

The series of measurements taken from a single sample, side, and laser position exhibited a high degree of autocorrelation. While homogeneity of the residual error variances could be assumed for each saw configuration and side (i.e., $\mathrm{Var}(\varepsilon_{ijklm}) = \sigma_{\varepsilon_{ij}}^2$), independence could not. In order to describe individual observations taken by each LRS, the autocorrelation in the data needed to be explicitly accounted for. Using a multi-stage model, the autocorrelation in the errors was estimated with autoregressive integrated moving average (ARIMA) models and ARIMA models modified for seasonal and long-memory effects (seasonal autoregressive fractionally integrated moving average, or SARFIMA models). The parameters from [4-1] were then fit with the

estimated autocorrelated error covariance matrix.

While the autocorrelated errors model provided a good description of the correlative structure in the data, it was only necessary to explicitly account for autocorrelation when performing tests of significance or predicting $y_{ijklm}$. If only summary statistics by sample and side are needed for SPC, the properties of the residual variance are not as important. For example, the X-bar chart described in Chapter 1 uses the standard error of the mean board thickness for a group of boards. For real-time data, this calculation can be made with [4-12]. Assuming that the variance of the average residuals ( $\mathrm{var}(\bar{\varepsilon}_{ijkl\cdot})$ ) approaches zero for large values of $n_{ijkl}$, the contribution of the residual variation ($\sigma_{\varepsilon_{ij}}^2$) to the standard error of the mean is negligible.

Thus, for use in detecting many of the common sawing defects described, the autocorrelation in the model can be ignored with only a negligible change in accuracy. In particular, control limits for monitoring averages and the components of variance (COV) due to boards, laser positions, and the interaction of boards and laser positions ($\sigma_{\beta_{ij}}^2$, $\sigma_{\lambda_{ij}}^2$ , $\sigma_{\beta\lambda_{ij}}^2$, respectively) can be computed assuming independent and identically distributed normal variates. These parameters could then be used directly to monitor machine positioning problems and wedge. Control limits could be constructed for the group of measurements from each board or cant sawn, and for subgroups of boards/cants. Although natural subgroups are not formed in the real-time data collection system, subgroups are necessary to quantify the COV for $\sigma_{\beta_{ij}}^2$. Subgroups could be formed based on time or by board location in a gang of saws, for example.

By their nature, sawing defects, such as taper and snake, are traditionally indicated by high within-board variability that causes a signal in the $S_w$ chart. Because the number of observations × sample × side × laser position is so large, traditional inferential statistics are no longer meaningful in describing the observations from a single board or cant surface. Confidence limits

constructed around $\sigma_{\varepsilon_{ij}}^2$ using a chi-square distribution are not valid where there is significant autocorrelation. If this autocorrelation is ignored, these limits would have degrees of freedom $(b_i - 1)(\bar{n}_{ij\cdot\cdot} - 1)$, where $\bar{n}_{ij\cdot\cdot}$ is the average number of observations per board and laser position for the *i*th saw configuration and *j*th side. For example, consider a subgroup of 4 boards and 2800 LRS measurements per board, side, and laser position. A sawing machine with a target standard deviation of 0.76 mm (0.030 inch) would have control limits of 0.74 and 0.78 mm (0.0293 and 0.0307 inch, or ± 2.3%). Using these limits, an extremely small change in the variability of the surface profile data would cause the $S_w$ chart to signal an assignable cause.

Shewhart-type charts have been derived based on the upper and lower quantiles of standard distributions fit to the distribution of process parameters (Padgett and Spurrier 1990; Grimshaw and Alt 1997). For instance, Levinson (1997) constructed a quantile (Q) chart for monitoring impurities in aluminum products based on a gamma distribution. Developing a Q chart for the within-board variance distribution is tempting because $S_w$ is easy to calculate and simple to use. However, using an $S_w$ chart with real-time data ignores a wealth of information from the LRS data. To illustrate this, Figure 5-1 shows profile measurements from four boards with the same mean and within-board variance ($\sigma_\varepsilon^2$). Board A has a very rough surface, while Board B exhibits trend along the length of the board. Boards C and D have underlying sinusoidal patterns; Board C has a period of 120 cm, while Board D has a period of 60 cm. Although these boards have very different quality problems, they would be indistinguishable on an $S_w$ chart.

Using regression techniques, trends can be extracted from the profile data and used to monitor taper, snipe, and flare. Taper can be detected by examining the trend along the entire length of the board, while snipe and flare are restricted to the last 15 cm (6 inches) of the board. In the presence of these defects, the data series from both the top and bottom laser positions would have an increasing (or decreasing) trend (as in Board B). Therefore, a linear regression fit

simultaneously to both series would have a coefficient associated with the direction of sawing that was significantly different from zero. Instead of using a mean or variance statistic derived from the data, control charts for these defects can be designed for the estimated regression coefficient for slope.

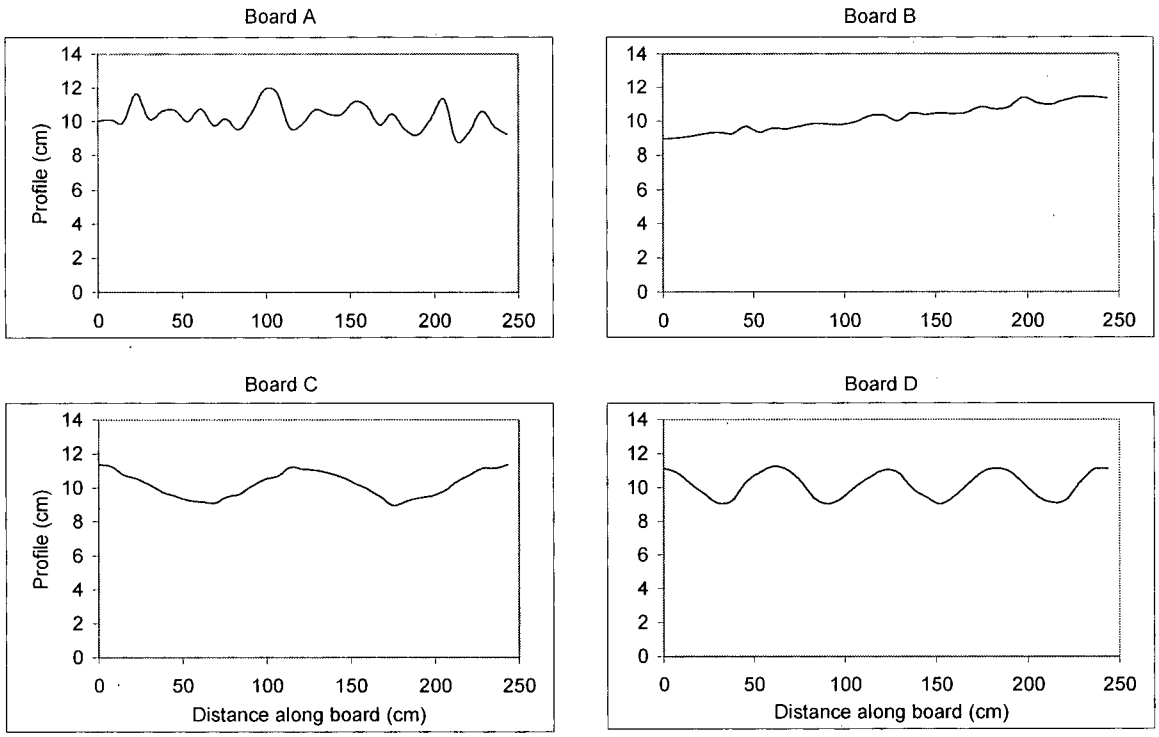



**Figure 5-1. Example of four boards with $\mu$=10 cm and $\sigma_\varepsilon$=0.60 cm.**

The idea of monitoring model parameters is not new. For example, parameter estimates from principal components analyses and other multivariate statistical techniques have been used to monitor multivariate processes in the chemical industry (Baseville 1988; Negiz *et al.* 1994). These researchers used parameter monitoring in the interest of parsimony; a large number of independent quality control variables were reduced to a few principal components to make the number of parameters monitored more reasonable. In real-time SPC for automated lumber manufacturing, parameter monitoring will instead be used for inference; a specific quality problem will be targeted with a single parameter inferred from the process data.

In addition to extracting trend, surface profile measurements can be decomposed into roughness

and waviness components to target quality problems described in Boards A, C, and D. As defined in industrial metrology, roughness is the high frequency (short wavelength, or closely spaced) repetitive or random deviations from the "normal" surface, whereas waviness is the medium-to-long frequency deviations (long wavelength). Some authors have tied the different wavelengths to different parts of the manufacturing process (Raja *et al.* 2002). For instance, in wood processing, roughness may be due to the sawblades, whereas waviness may be due to saw guides or hold down mechanisms. Roughness parameters have been used in wood products technology to measure product smoothness from a visual grading standpoint (Funck *et al.* 1993), for assessment of tool wear (Lemaster and Taylor 1999), and to relate a numerical measure to tactile roughness (Fujiwara *et al.* 2001; Sandak *et al.* 2003).

As an alternative to an $S_w$ chart, the profile measurements could be decomposed into trend, roughness, and waviness. Based on the distribution of these quantities and practical considerations, Shewhart-type Q charts for these measures could be derived. Non-parametric charts have also been developed that use the empirical quantiles of a bootstrap distribution (Willemain and Runger 1996). However, these methods have been found to be adequate only with sample sizes of 300 or more (Vermaat *et al.* 2003).

Multivariate charts are useful when individual charts do not provide enough information to decide if a process is in control (Wheeler 1995). However, single multivariate charts have been found to be poor operationally, as out-of-control signals still must be investigated via univariate control charts in order to determine the cause of the signals (Does *et al.* 1999). For each sawing defect type discussed, univariate control charts are suggested as a part of a multi-chart SPC system for real-time size control; a single multivariate chart is not suggested. Univariate charts targeting specific defects give more information than simple in-control/out-of-control signals in that each chart is related to a specific sawing problem that can be addressed by mill staff.

## 5.4 Materials and Methods

Real-time SPC data were obtained with the laser measurement apparatus described in Section 4.4 (Chapter 4). The sample boards consisted of 110 green and un-planed western hemlock (*Tsuga heterophyla* (Raf) Sarg.) boards with dimensions 51.5 mm x 135 mm ($2\frac{1}{32}$ x $5\frac{5}{16}$ inches). The sample data were profile measurements taken simultaneously with four LRSs. Areas containing non-sawing defects, such as wane[26], were removed from these data manually, using the known positions of these defects mapped at the time of data collection. These data were then filtered for measurement errors using the MSD Method (Method 3, Chapter 2). Results were obtained separately for each saw configuration × side combination; the four saw configurations were BB (both sides of the board bandsawn), BC (Side 1 bandsawn, Side 2 chipped), CB (Side 1 chipped, Side 2 bandsawn), and RR (both sides circular-sawn).

Control limits and other descriptors were developed to best detect the five common sawing defects listed in Section 5.3. For the first two defects (machine positioning problems and wedge), summary statistics within board were used. For the remaining three defects, profile observations within each board were examined. For each defect type, several candidate control charts are presented.

For machine positioning problems and wedge, control limits were based on the traditional 3-sigma control charts originally developed by Shewhart (1931), and extended to processes exhibiting between- and within-part-size variability (Maness *et al.* 2003). The basis for these control charts are the average profile values by board and by laser, and the components of variance from the statistical model [4-1].

Control charts of this type for lumber manufacturing have traditionally been based on

---

[26] Wane is the natural curvature of the edge of a board sawn from a log. Additional definitions are given in the List of Nomenclature and Abbreviations for this thesis.

subgrouping. Natural subgrouping occurred because only small groups of boards were pulled periodically for SPC measurement. However, in real-time data collection, there is no obvious natural subgrouping, as production of lumber is continuous with the exception of shift changes and breaks. To reflect this continuity, control charts for individuals were investigated where possible. Subgrouping is necessary if control charts for the between-board variance are to be constructed. Control charts for moving statistics, such as moving average and moving standard deviation, have been used in continuous processes; however, these charts tend to over-signal due to correlation introduced by using overlapping observations (Wheeler 1995). Therefore, artificial subgroups were created by taking groups of subsequent boards.

The primary reason for choosing these types of charts is their ease of use and familiarity to mill personnel (Young and Winistorfer 1999). Although more modern control charting techniques, such as the CUSUM and EWMA charts could have been investigated, many authors (e.g., Wheeler 1995) have shown that these charts provide only marginal improvements, if any, over traditional Shewhart charts, and are notoriously difficult to develop, maintain, and interpret. Some SPC software packages are capable of maintaining moving centerline EWMA charts, such as those advocated by Montgomery (2001); however, these packages are rarely used in lumber mill applications.

Charts specific to taper, snipe/flare, and snake were developed using non-traditional methods. For trend-related defects, such as taper and snipe/flare, charts were constructed for slopes of regression lines fit to the profile data along the length of the board. For snake, charts were developed to monitor several measures of roughness and waviness. Control charts for these defects were Shewhart-type Q charts.

The parameters that are monitored in the proposed taper, snake, and snipe/flare charts were suggested because they describe specific defects that cause quality problems in automated mills.

Quantile control charts were proposed because the distributional properties of the proposed parameters to be monitored are unknown. Control limits for these charts are relatively simple to derive and easily explained to mill staff, appearing no different from Shewhart charts in an operational sense (Levinson 1997).

A summary of the proposed control charts is given in Table 5-1. Detailed derivations are given in the following sections by defect type.

**Table 5-1. Summary of proposed control charts.**

| Defect Targeted | Name of Chart | Statistic monitored | Equation |
|---|---|---|---|
| Machine positioning | X-bar$_{ind}$ | Individual board averages: $\bar{y}_{i\cdot k\cdot\cdot}$ | [5-1] |
| | X-bar$_{grp}$ | Subgrouped board averages: $\bar{y}_{i\cdot g\cdot\cdot\cdot}$ | [5-4] |
| | MR$_{\beta}$ | Moving range of successive board averages: $MR(\bar{y}_{i\cdot k\cdot\cdot})_k$ | [5-7] |
| | R$_{\beta_{grp}}$ | Range of subgrouped board averages: $R(\bar{y}_{i\cdot gk\cdot\cdot})\Big\|_{k=1}^{G}$ | [5-9] |
| | S$_{\beta}$ | Between board variation of subgrouped boards: $S^2_{\beta_{ijg}}$ | [5-11] |
| Wedge | R$_{\lambda_{ind}}$ | Range of laser position averages within board by side: $R(\bar{y}_{ijkl\cdot})\Big\|_{l=1}^{2}$ | [5-14] |
| | R$_{\lambda_{grp}}$ | Range of laser position averages within subgroup by side: $R(\bar{y}_{ijg\cdot l\cdot})\Big\|_{l=1}^{2}$ | [5-16] |
| | MR$_{\beta\lambda}$ | Moving range of successive board averages by side and laser position: $MR(\bar{y}_{ijkl\cdot})_k$ | [5-18] |
| | R$_{\beta\lambda_{grp}}$ | Range of subgrouped board averages by side and laser position: $R(\bar{y}_{ijgkl\cdot})\Big\|_{k=1}^{G}$ | [5-20] |
| | S$_{\lambda}$ | Between laser position variation for subgrouped boards by side: $S^2_{\lambda_{ijg}}$ | [5-22] |
| | S$_{\beta\lambda}$ | Interaction of board × laser position variation for subgrouped boards by side: $S^2_{\beta\lambda_{ijg}}$ | [5-25] |
| Taper | Q$_{\tau}$ | Slope by board and side: $\hat{\tau}_{1ijk}$ | [5-28] |
| Snipe/flare | Q$_{\tau'}$ | Slope of last 15 cm of board by board and side: $\hat{\tau}'_{1ijk}$ | [5-29] |
| Snake | Q$_{r_a}$ | Average roughness by board and side: $\overline{r^{(a)}_{ijk\cdot}}$ | [5-38] |
| | Q$_{r_q}$ | RMS roughness by board and side: $\overline{r^{(q)}_{ijk\cdot}}$ | [5-39] |
| | Q$_{r_p}$ | Peak-to-peak roughness by board and side: $\overline{r^{(p)}_{ijk\cdot}}$ | [5-40] |
| | Q$_{w_a}$ | Average waviness by board and side: $\overline{w^{(a)}_{ijk\cdot}}$ | [5-41] |
| | Q$_{w_q}$ | RMS waviness by board and side: $\overline{w^{(q)}_{ijk\cdot}}$ | [5-42] |
| | Q$_{w_p}$ | Peak-to-peak waviness by board and side: $\overline{w^{(p)}_{ijk\cdot}}$ | [5-43] |

### 5.4.1 Proposed Charts for Machine Positioning Problems

Machine positioning problems are indicated when boards are consistently thicker or thinner along their entire lengths. Thus, monitoring for machine positioning problems was performed with control charts based on average board values. These values included: (1) the average profile by board, (2) the range and moving range of subsequent board averages, and (3) the between board variance.

#### 5.4.1.1 X-bar$_{ind}$ Chart

Machine positioning problems are indicated when individual boards have higher (or lower) average thickness values. Therefore, an X-bar chart for individual boards (X-bar$_{ind}$ chart) was developed. Without using subgroups, an average profile was computed for each board: $\bar{y}_{i \cdot k \cdot \cdot}$. These values were plotted on the X-bar$_{ind}$ chart, with control limits for a particular saw configuration ($i$) given by:

$$\begin{aligned} CL &= T_i \\ LCL &= CL - 3\hat{\sigma}_{(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})/2} \\ UCL &= CL + 3\hat{\sigma}_{(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})/2} \end{aligned} \qquad [5\text{-}1]$$

where: $CL$ is the centreline of the chart;

$LCL$ is the lower control limit of the chart;

$UCL$ is the upper control limit of the chart;

$T_i$ is the target surface profile value;

$\hat{\sigma}_{(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})/2}$ is the estimated standard error of the average profile value by board for the $i$th sawing configuration; and

$\bar{y}_{i1k\cdot\cdot}$ and $\bar{y}_{i2k\cdot\cdot}$ are the average Side 1 and Side 2 profiles values for the $i$th saw configuration and $k$th board.

The target surface profile value is half the target thickness value for the sawing configuration. In this case, it was calculated as the long-term average of the two average profiles by side for the *i*th saw configuration ( $\hat{\mu}_{i1\cdots}$ and $\hat{\mu}_{i2\cdots}$ from [4-1], respectively). These values were obtained from estimating the parameters of the mixed model in [4-1] with the SAS procedure PROC MIXED[27].

Although $\bar{y}_{i1k\cdot\cdot} = \bar{y}_{i2k\cdot\cdot}$, $y_{i1klm}$ and $y_{i2klm}$ are independent and therefore, the parameters in [4-1] are estimated by saw configuration and side. Thus, the standard errors in [5-1] were calculated using two sets of estimated model parameters:

$$\hat{\sigma}_{(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})/2} = \sqrt{\mathrm{var}\left(\frac{\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot}}{2}\right)} = \sqrt{\frac{1}{4}\mathrm{var}(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})} = \frac{1}{2}\sqrt{\hat{\sigma}^2_{\bar{y}_{i1k\cdot\cdot}}+\hat{\sigma}^2_{\bar{y}_{i2k\cdot\cdot}}} \quad [5\text{-}2]$$

Using a components of variance, with the number of laser positions per side = 2, [5-2] becomes[28]:

$$\begin{aligned}\hat{\sigma}_{(\bar{y}_{i1k\cdot\cdot}+\bar{y}_{i2k\cdot\cdot})/2} &= \frac{1}{2}\sqrt{\hat{\sigma}^2_{\beta_{i1}}+\frac{\hat{\sigma}^2_{\lambda i1}+\hat{\sigma}^2_{\beta\lambda_{i1}}}{2}+\frac{\mathrm{var}(\bar{\varepsilon}_{i1kl\cdot})}{2}+\hat{\sigma}^2_{\beta_{i2}}+\frac{\hat{\sigma}^2_{\lambda i2}+\hat{\sigma}^2_{\beta\lambda_{i2}}}{2}+\frac{\mathrm{var}(\bar{\varepsilon}_{i2kl\cdot})}{2}} \\ &\cong \frac{1}{2}\sqrt{\hat{\sigma}^2_{\beta_{i1}}+\frac{\hat{\sigma}^2_{\lambda i1}+\hat{\sigma}^2_{\beta\lambda_{i1}}}{2}+\hat{\sigma}^2_{\beta_{i2}}+\frac{\hat{\sigma}^2_{\lambda i2}+\hat{\sigma}^2_{\beta\lambda_{i2}}}{2}}\end{aligned} \quad [5\text{-}3]$$

where: $\hat{\sigma}^2_{\beta_{ij}}, \hat{\sigma}^2_{\lambda_{ij}}$, and $\hat{\sigma}^2_{\beta\lambda_{ij}}$ are estimates of the COV $\sigma^2_{\beta_{ij}}, \sigma^2_{\lambda_{ij}}$, and $\sigma^2_{\beta\lambda_{ij}}$, respectively, obtained from estimating the parameters of the mixed model in [4-1] with PROC MIXED.

### 5.4.1.2 X-bar$_{grp}$ Chart

Sawing defects caused by machine positioning problems tend to occur in subsequently sawn boards. Thus, X-bar charts were developed that used groups of boards (X-bar$_{grp}$ chart). In manual SPC, periodic subgroups are taken, and thus, X-bar charts based on subgroups are

---

[27] All SAS procedures referred to in this chapter were run with Version 8.2 (SAS Institute 2002).

[28] See Section 4.5.1 for an explanation of this approximation.

commonly constructed. For this real-time data, artificial subgroups were created by taking successive groups of $G$ boards. A range of subgrouping values was investigated, with $G$ =4, 6, 8, 10, 12, 16, and 20. Adding an additional subscript to denote the $g$th group of $G$ boards, the group average profile for the $i$th saw configuration was $\bar{y}_{i \cdot g \cdots}$. These values were plotted on the X-bar$_{grp}$ chart with control limits for a particular saw configuration ($i$) given by:

$$\begin{aligned} CL &= T_i \\ LCL &= CL - 3\hat{\sigma}_{(\bar{y}_{i1g\cdots}+\bar{y}_{i2g\cdots})/2} \\ UCL &= CL + 3\hat{\sigma}_{(\bar{y}_{i1g\cdots}+\bar{y}_{i2g\cdots})/2} \end{aligned} \qquad [5\text{-}4]$$

where: $\hat{\sigma}_{(\bar{y}_{i1g\cdots}+\bar{y}_{i2g\cdots})/2}$ is the estimated standard error of the average profile value by subgroup for the $i$th sawing configuration; and

$\bar{y}_{i1g\cdots}$ and $\bar{y}_{i2g\cdots}$ are the average Side 1 and Side 2 profiles values for the $i$th saw configuration, $k$th board, and $g$th group.

As in [5-1], the centreline is the target value $T_i$, and the standard error term in [5-4] uses components from the models from each side. Using a components of variance approach and the estimated parameters from [4-1], the standard error for [5-4] was calculated as:

$$\begin{aligned} \hat{\sigma}_{(\bar{y}_{i1g\cdots}+\bar{y}_{i2g\cdots})/2} &= 1/2\sqrt{\hat{\sigma}^2_{\bar{y}_{i1g\cdots}} + \hat{\sigma}^2_{\bar{y}_{i2g\cdots}}} \\ &\cong \frac{1}{2}\sqrt{\frac{\hat{\sigma}^2_{\beta_{i1}} + \hat{\sigma}^2_{\beta_{i2}}}{G} + \frac{\hat{\sigma}^2_{\lambda_{i1}} + \hat{\sigma}^2_{\lambda_{i2}}}{2} + \frac{\hat{\sigma}^2_{\beta\lambda_{i1}} + \hat{\sigma}^2_{\beta\lambda_{i2}}}{2G}} \end{aligned} \qquad [5\text{-}5]$$

### 5.4.1.3 MR$_\beta$ Chart

The moving range (MR) is defined as the absolute difference between successive observations. Machine positioning problems are indicated by large differences between the average size of subsequently sawn boards, and thus, a moving range chart based on board averages (MR$_\beta$ chart) was constructed for detecting this sawing defect. For monitoring average profiles from

individual boards, the moving range between the $k$th and ($k$+1)th successive board averages in the $i$th saw configuration was computed as:

$$MR(\bar{y}_{i\cdot k\cdot\cdot})_k = |\,\bar{y}_{i\cdot,k,\cdot\cdot} - \bar{y}_{i\cdot,k+1,\cdot\cdot}\,| \qquad [5\text{-}6]$$

These values were plotted on the $MR_\beta$ chart with control limits for a particular saw configuration ($i$) given by:

$$\begin{aligned} CL &= \overline{MR(\bar{y}_{i\cdot k\cdot\cdot})_k}_{\,i} \\ LCL &= \left({}^{D_{0.001}}\!/_{d_2}\right)\overline{MR(\bar{y}_{i\cdot k\cdot\cdot})_k}_{\,i} \\ UCL &= \left({}^{D_{0.999}}\!/_{d_2}\right)\overline{MR(\bar{y}_{i\cdot k\cdot\cdot})_k}_{\,i} \end{aligned} \qquad [5\text{-}7]$$

where: $\overline{MR(\bar{y}_{i\cdot k\cdot\cdot})_k}_{\,i}$ is the average of the moving ranges between successive boards for the $i$th saw configuration;

$d_2$ is the bias correction factor for range[29]; and

$D_{0.001}$ and $D_{0.999}$ are the 0.1 th and 99.9th percentiles for the distribution of ranges (Harter 1960)[30].

### 5.4.1.4 $R_{\beta_{grp}}$ Chart

Machine positioning problems can also be found by examining the range of averages in a group of boards. Therefore, a range chart for subgrouped board averages ($R_{\beta_{grp}}$ chart) was constructed. Adding a subscript for groups, the board average in the $g$th subgroup and $i$th saw configuration was $\bar{y}_{i\cdot gk\cdot\cdot}$. The range of these board averages was calculated as:

---

[29] See Montgomery (2001) for an explanation of control chart constants, such as $d_2$ and $c_4$.

[30] In order to make direct comparisons to other proposed charts with 3-Sigma limits, control limits that give a false alarm rate of ~0.27% are desired. This implies upper and lower quantiles of 0.0135% and 0.9865%; however, tabulated values for the distribution of ranges are only available at 0.001 increments, and thus, the closest values were used for this chart, as well as all other Range and Moving Range charts.

$$R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G} = \text{Range}(\bar{y}_{i\cdot g1\cdot\cdot}, \bar{y}_{i\cdot g2\cdot\cdot}, \ldots, \bar{y}_{i\cdot gG\cdot\cdot}) \qquad [5\text{-}8]$$

These values were plotted on the $R_{\beta_{grp}}$ chart with control limits for a particular saw configuration ($i$) given by:

$$\begin{aligned} CL &= \overline{R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G}}_{\,i} \\ LCL &= \left({}^{D_{0.001}}\!/_{d_2}\right)\overline{R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G}}_{\,i} \\ UCL &= \left({}^{D_{0.999}}\!/_{d_2}\right)\overline{R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G}}_{\,i} \end{aligned} \qquad [5\text{-}9]$$

where: $\overline{R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G}}_{\,i}$ is the average of the $R(\bar{y}_{i\cdot gk\cdot\cdot})\Big|_{k=1}^{G}$ values in the $i$th saw configuration.

### 5.4.1.5 $S_\beta$ Chart

Machine positioning problems are indicated by an increase in the between-board variation. Thus, a control chart to monitor the variation due to boards was developed ($S_\beta$ chart). Using artificial subgroups, the estimated between board variation for the $i$th saw configuration, $j$th side, and $g$th group ($S^2_{\beta_{ijg}}$) is a linear combination of the mean squares for board and board × laser in that group ($MS_{\beta_{ijg}}$ and $MS_{\beta\lambda_{ijg}}$, respectively[31]):

$$S^2_{\beta_{ijg}} = \frac{MS_{\beta_{ijg}} - MS_{\beta\lambda_{ijg}}}{\sum_{l=1}^{2} \bar{n}_{ijg\cdot l}} \qquad [5\text{-}10]$$

where: $\bar{n}_{ijg\cdot l}$ is the average number of observations per board in the $g$th group and $l$th laser position, for the $i$th saw configuration and $j$th side; and

$S^2_{\beta_{ijg}}$ was restricted to be non-negative.

[31] Computations for these mean squares, as well as all other mean squares required for the construction of the proposed control charts, are given in Appendix II.

Values of $S^2_{\beta_{ijg}}$ were computed and plotted on the $S_\beta$ chart. The control limits for a subgroup of size $G$ in a particular saw configuration ($i$) and side ($j$) were calculated as:

$$
\begin{aligned}
CL &= \hat{\sigma}^2_{\beta_{ij}} \\
LCL &= \hat{\sigma}^2_{\beta_{ij}}\, \chi^2_{[0.00135;df(\beta_{ij})_G]} \Big/ df(\beta_{ij})_G \\
UCL &= \hat{\sigma}^2_{\beta_{ij}}\, \chi^2_{[0.99865;df(\beta_{ij})_G]} \Big/ df(\beta_{ij})_G
\end{aligned}
\qquad [5\text{-}11]
$$

where: $\hat{\sigma}^2_{\beta_{ij}}$ was obtained from fitting the mixed-effects model in [4-1];

$\chi^2_{(0.00135;df(\beta_{ij})_G)}$ and $\chi^2_{(0.99865;df(\beta_{ij})_G)}$ are cumulative probability values for a Chi-square distribution with $df(\beta_{ij})_G$ degrees of freedom; and

$df(\beta_{ij})_G$ are the estimated degrees of freedom of the Chi-square distribution for $\sigma^2_{\beta_{ij}}$ in the $i$th saw configuration and $j$th side, with subgroup size $G$.

The degrees of freedom were approximated using the Satterthwaite procedure (Gaylor and Hopper 1969)[32]:

$$
df(\beta_{ij})_G = \frac{\left(2\bar{n}_{ij..}\hat{\sigma}^2_{\beta_{ij}}\right)^2}{\dfrac{MS^2_{\beta_{ij}}}{G-1} + \dfrac{MS^2_{\beta\lambda_{ij}}}{G-1}} \qquad [5\text{-}12]
$$

where: $\bar{n}_{ij..}$ is the average number of observations per board and laser position, for the $i$th saw configuration and $j$th side; and

$MS_{\beta_{ij}}$ and $MS_{\beta\lambda_{ij}}$ are the non-grouped mean squares underlying the estimated parameters from the mixed-effects model in [4-1].

---

[32] This approximation is appropriate when $MS_{\beta_{ij}}/MS_{\beta\lambda_{ij}} > F_{(G\text{-}1,\, G\text{-}1;\, 0.99865)} * F_{(G\text{-}1,\, G\text{-}1;\, 0.5)}$ (Gaylor and Hopper 1969).

### 5.4.2 Proposed Charts for Wedge

Wedge is indicated by a difference between the top and the bottom laser position measurements. Thus, monitoring for wedge involved comparing the profile values from the top versus bottom laser positions. This was accomplished with control charts for (1) ranges, and (2) the between-laser and board × laser interaction variances, $\sigma_{\lambda_{ij}}^2$ and $\sigma_{\beta\lambda_{ij}}^2$, respectively Ranges were computed for average profile measurements between laser positions within board, and by board and laser position between subsequent boards.

#### 5.4.2.1 $R_{\lambda_{ind}}$ Chart

Wedge results in a difference between the average profile measurements from the top laser versus those from the bottom laser. Therefore, an R chart for laser position averages within individual boards ($R_{\lambda_{ind}}$ chart) was developed. Since there are only two laser positions, the range between laser positions within each board for the *i*th saw configuration, *j*th side, and *k*th board was computed as:

$$R(\bar{y}_{ijkl\cdot})\Big|_{l=1}^{2} = \left|\bar{y}_{ijk1\cdot} - \bar{y}_{ijk2\cdot}\right| \qquad [5\text{-}13]$$

These values were plotted on the $R_{\lambda_{ind}}$ chart with control limits for a particular saw configuration (*i*) and side (*j*) given by:

$$\begin{aligned} CL &= \overline{R(\bar{y}_{ijkl\cdot})\Big|_{l=1}^{2}}_{ij} \\ LCL &= \left({}^{D_{0.001}}/_{d_2}\right)\overline{R(\bar{y}_{ijkl\cdot})\Big|_{l=1}^{2}}_{ij} \\ UCL &= \left({}^{D_{0.999}}/_{d_2}\right)\overline{R(\bar{y}_{ijkl\cdot})\Big|_{l=1}^{2}}_{ij} \end{aligned} \qquad [5\text{-}14]$$

where: $\overline{R(\bar{y}_{ijkl\cdot})\Big|_{l=1}^{2}}_{ij}$ is the average of all $R(\bar{y}_{ijkl\cdot})\Big|_{l=1}^{2}$ values for the *i*th saw configuration, *j*th side.

#### 5.4.2.2 $R_{\lambda_{grp}}$ Chart

In the presence of wedge, subsequent groups of boards will exhibit differences between the average profile computed for the top and bottom laser positions, and thus, a range chart was developed to monitor laser position averages using artificial subgroups ($R_{\lambda_{grp}}$ chart). Adding a subscript to denote the *g*th subgroup formed from *G* successive sample boards, the average profile by group and laser position for the *i*th saw configuration and *j*th side is $\bar{y}_{ijg\cdot l\cdot}$. Since there are only two laser positions, the range of these averages was calculated for each group as:

$$R(\bar{y}_{ijg\cdot l\cdot})\Big|_{l=1}^{2} = \left|\bar{y}_{ijg\cdot 1\cdot} - \bar{y}_{ijg\cdot 2\cdot}\right| \qquad [5\text{-}15]$$

These values were plotted on the $R_{\lambda_{grp}}$ chart with control limits for a particular saw configuration (*i*) and side (*j*) given by:

$$\begin{aligned} CL &= \overline{R(\bar{y}_{ijg\cdot l\cdot})\Big|_{l=1}^{2}}_{ij} \\ LCL &= \left({}^{D_{0.001}}\!/_{d_2}\right)\overline{R(\bar{y}_{ijg\cdot l\cdot})\Big|_{l=1}^{2}}_{ij} \\ UCL &= \left({}^{D_{0.999}}\!/_{d_2}\right)\overline{R(\bar{y}_{ijg\cdot l\cdot})\Big|_{l=1}^{2}}_{ij} \end{aligned} \qquad [5\text{-}16]$$

where: $\overline{R(\bar{y}_{ijg\cdot l\cdot})\Big|_{l=1}^{2}}_{ij}$ is the average of all $R(\bar{y}_{ijg\cdot l\cdot})\Big|_{l=1}^{2}$ values for the *i*th saw configuration, *j*th side.

#### 5.4.2.3 $MR_{\beta\lambda}$ Chart

Wedge is also indicated by a change in the average values of subsequent profile measurements by board and laser position. Thus, a moving range chart was developed for the average profile values by board and laser position ($MR_{\beta\lambda}$ chart). Using individual board values, a moving range between subsequent boards for the *i*th saw configuration and *j*th side was computed for each *l*th laser position as:

$$MR(\bar{y}_{ijkl\cdot})_k = \left|\bar{y}_{ij,k,l\cdot} - \bar{y}_{ij,k+1,l\cdot}\right| \qquad [5\text{-}17]$$

These values were plotted on the $MR_{\beta\lambda}$ chart with control limits for a particular saw configuration ($i$), side ($j$), and laser position ($l$) given by:

$$\begin{aligned} CL &= \overline{MR(\bar{y}_{ijkl\cdot})_k}_{ijl} \\ LCL &= \left({}^{D_{0.001}}\!/_{d_2}\right)\overline{MR(\bar{y}_{ijkl\cdot})_k}_{ijl} \\ UCL &= \left({}^{D_{0.999}}\!/_{d_2}\right)\overline{MR(\bar{y}_{ijkl\cdot})_k}_{ijl} \end{aligned} \qquad [5\text{-}18]$$

where: $\overline{MR(\bar{y}_{ijkl\cdot})_k}_{ijl}$ is the average of all $MR(\bar{y}_{ijkl\cdot})_k$ values for the $i$th saw configuration, $j$th side, and $l$th laser position.

### 5.4.2.4 $R_{\beta\lambda_{grp}}$ Chart

The change in average profile measurements by board and laser position can also be monitored by group. Thus, range charts were constructed for subgroups of board by laser averages ($R_{\beta\lambda_{grp}}$ chart). Adding a subscript for the $g$th subgroup, the average profile was computed for each board × laser within each subgroup: $\bar{y}_{ijgkl\cdot}$. The range of these averages within each subgroup for the $i$th saw configuration, $j$th side, and $l$th laser position was calculated as:

$$R(\bar{y}_{ijgkl\cdot})\Big|_{k=1}^{G} = \text{Range}(\bar{y}_{ijg1l\cdot}, \bar{y}_{ijg2l\cdot}, \ldots, \bar{y}_{ijgGl\cdot}) \qquad [5\text{-}19]$$

These values were plotted on the $R_{\beta\lambda_{grp}}$ chart with control limits for a particular saw configuration ($i$), side ($j$), and laser position ($l$) given by:

$$\begin{aligned} CL &= \overline{R(\bar{y}_{ijgkl\cdot})\Big|_{k=1}^{G}}_{ijl} \\ LCL &= \left({}^{D_{0.001}}\!/_{d_2}\right)\overline{R(\bar{y}_{ijgkl\cdot})\Big|_{k=1}^{G}}_{ijl} \\ UCL &= \left({}^{D_{0.999}}\!/_{d_2}\right)\overline{R(\bar{y}_{ijgkl\cdot})\Big|_{k=1}^{G}}_{ijl} \end{aligned} \qquad [5\text{-}20]$$

where: $\overline{R(\bar{y}_{ijgkl\cdot})\big|_{k=1}^{G}}_{ijl}$ is the average of the $R(\bar{y}_{ijgkl\cdot})\big|_{k=1}^{G}$ values in the $i$th saw configuration, $j$th side, and $l$th laser position.

### 5.4.2.5 $S_\lambda$ Chart

Wedge results in high laser-to-laser variation. Thus, a chart was developed to monitor the between laser variation ($S_\lambda$ chart). For the $i$th saw configuration, $j$th side, and $g$th group, the between laser variation ($S^2_{\lambda ijg}$) is a linear combination of the mean squares for laser and laser × board in that group ($MS_{\lambda_{ijg}}$ and $MS_{\beta\lambda_{ijg}}$, respectively):

$$S^2_{\lambda ijg} = \frac{MS_{\lambda ijg} - MS_{\beta\lambda ijg}}{\sum_{k=1}^{G} \bar{n}_{ijgk\cdot}} \qquad [5\text{-}21]$$

where: $\bar{n}_{ijgk\cdot}$ is the average number of observations per laser in the $g$th group and $k$th board, for the $i$th saw configuration and $j$th side; and

$S^2_{\lambda ijg}$ was restricted to be non-negative.

Values of $S^2_{\lambda ijg}$ were computed and plotted on the $S_\lambda$ chart, with control limits for subgroup size $G$ in a particular saw configuration ($i$) and side ($j$) calculated as:

$$\begin{aligned} CL &= \hat{\sigma}^2_{\lambda ij} \\ LCL &= \hat{\sigma}^2_{\lambda ij}\, \chi^2_{[0.00135;df(\lambda_{ij})_G]} \big/ df(\lambda_{ij})_G \\ UCL &= \hat{\sigma}^2_{\lambda ij}\, \chi^2_{[0.99865;df(\lambda_{ij})_G]} \big/ df(\lambda_{ij})_G \end{aligned} \qquad [5\text{-}22]$$

where: $\hat{\sigma}^2_{\lambda ij}$ was obtained from fitting the mixed-effects model in [4-1]; and

$df(\lambda_{ij})_G$ are the estimated degrees of freedom of the Chi-square distribution for $\sigma^2_{\lambda ij}$ in the $i$th saw configuration, and $j$th side, with subgroup size $G$.

The degrees of freedom were approximated using the Satterthwaite procedure (Gaylor and Hopper 1969)[33]:

$$df(\lambda_{ij})_G = \left(G\bar{n}_{ij\cdot\cdot}\hat{\sigma}^2_{\lambda ij}\right)^2 \Big/ \left(MS^2_{\lambda ij}/1 + MS^2_{\beta\lambda_{ij}}/(G-1)\right): \qquad [5\text{-}23]$$

where: $MS_{\lambda_{ij}}$ and $MS_{\beta\lambda_{ij}}$ are the non-grouped mean squares underlying the estimated parameters from the mixed-effects model in [4-1].

### 5.4.2.6 $S_{\beta\lambda}$ Chart

Wedge also results in high variation of the interaction of boards and laser positions. Thus, a chart was developed to monitor this variation ($S_{\beta\lambda}$ chart). The board × laser position variance for the *i*th saw configuration, *j*th side, and *g*th group ($S^2_{\beta\lambda_{ijg}}$) is a linear combination of mean squares for board × laser and mean squares residual for that group ($MS_{\beta\lambda_{ijg}}$ and $MS_{\varepsilon_{ijg}}$, respectively):

$$S^2_{\beta\lambda_{ijg}} = \frac{MS_{\beta\lambda_{ijg}} - MS_{\varepsilon ijg}}{\bar{n}_{ijg\cdot\cdot}} \qquad [5\text{-}24]$$

where: $\bar{n}_{ijg\cdot\cdot}$ is the average number of observations per board and laser position in the *g*th group, for the *i*th saw configuration and *j*th side; and

$S^2_{\beta\lambda_{ijg}}$ is restricted to be non-negative.

Values of $S^2_{\beta\lambda_{ijg}}$ were computed and plotted on the $S_{\beta\lambda}$ chart. The control limits for a subgroup of size *G* in a particular saw configuration (*i*) and side (*j*) were calculated as:

---

[33] This approximation is appropriate when $MS_{\lambda_{ij}}/MS_{\beta\lambda_{ij}} > F_{(1,\,G\text{-}1;\,0.99865)} * F_{(G\text{-}1,\,1;\,0.5)}$ (Gaylor and Hopper 1969).

$$CL = \hat{\sigma}^2_{\beta\lambda_{ij}}$$
$$LCL = \hat{\sigma}^2_{\beta\lambda_{ij}}\ \chi^2_{[0.00135;df(\beta\lambda_{ij})_G]} \Big/ df(\beta\lambda_{ij})_G \qquad [5\text{-}25]$$
$$UCL = \hat{\sigma}^2_{\beta\lambda_{ij}}\ \chi^2_{[0.99865;df(\beta\lambda_{ij})_G]} \Big/ df(\beta\lambda_{ij})_G$$

where: $\hat{\sigma}^2_{\beta\lambda_{ij}}$ was obtained from fitting the mixed-effects model in [4-1]; and

$df(\beta\lambda_{ij})_G$ are the estimated degrees of freedom of the Chi-square distribution for $\sigma^2_{\beta\lambda_{ij}}$ in the $i$th saw configuration and $j$th side, with subgroup size $G$.

The degrees of freedom were approximated using the Satterthwaite procedure (Gaylor and Hopper 1969)[34]:

$$df(\beta\lambda_{ij})_G = (\bar{n}_{ij\cdot\cdot}\hat{\sigma}^2_{\beta\lambda_{ij}})^2 \Big/ \left(MS^2_{\beta\lambda_{ij}}/(G-1) + MS^2_{\varepsilon_{ij}}/\left(2G(\bar{n}_{ij\cdot\cdot}-1)\right)\right) \qquad [5\text{-}26]$$

where: $MS_{\beta\lambda_{ij}}$ and $MS_{\varepsilon_{ij}}$ are the non-grouped mean squares underlying the estimated parameters from the mixed-effects model in [4-1].

### 5.4.3 Proposed Charts for Detecting Trend-related Defects

Since taper, flare, and snipe are all indicated by increasing (decreasing) trend along the length of the board, regression techniques were used to construct charts. A surface was simultaneously fit to the profile data for both top and bottom laser positions on each side of each board. The regression model to detect linear trend was a function of the horizontal distance along the board ($x_1$) and the vertical position of the laser measurement ($x_2$, e.g., =22 mm (⅞ inch) for the bottom laser position ($l$=1), and 106 mm (4 3/16 inches) for the top laser position ($l$=2)):

$$y_{ijklm}(x_1, x_2) = \tau_{0_{ijk}} + \tau_{1_{ijk}}x_1 + \tau_{2_{ijk}}x_2 + \zeta_{ijklm} \qquad [5\text{-}27]$$

---

[34] This approximation is appropriate when $MS_{\beta\lambda_{ij}}/MS_{\varepsilon_{ij}} > F_{(G-1,\ 2G(\bar{n}_{ij\cdot\cdot}-1);\ 0.99865)} * F_{(2G(\bar{n}_{ij\cdot\cdot}-1),\ G-1;\ 0.5)}$ (Gaylor and Hopper 1969).

where: $\tau_{0_{ijk}}$, $\tau_{1_{ijk}}$, and $\tau_{2_{ijk}}$ are the regression coefficients for the *i*th saw configuration, *j*th side, and *k*th board; and

$\zeta_{ijklm}$ is the prediction error, assumed ~ N(0, $\sigma_{\zeta_{ij}}^2$).

### 5.4.3.1 $Q_\tau$ Chart

Using the SAS procedure PROC REG, estimates of $\tau_{0_{ijk}}$, $\tau_{1_{ijk}}$, and $\tau_{2_{ijk}}$ ($\hat{\tau}_{0ijk}$, $\hat{\tau}_{1ijk}$, and $\hat{\tau}_{2ijk}$) were found using all observations from the each board. The $Q_\tau$ chart was developed to monitor the estimated regression coefficient associated with the horizontal distance along the board ($\hat{\tau}_{1ijk}$) by saw configuration and side. Control limits were constructed from the distribution of $\hat{\tau}_{1ijk}$ values by saw configuration and side (*ij*). Under in-control conditions, it is expected that boards have zero trend (E($\bar{\hat{\tau}}_{1ij\cdot}$)=0). Since these distributions were approximately normal, control limits for the *i*th saw configuration and *j*th side were set using 3-sigma limits:

$$\begin{aligned} CL &= 0 \\ LCL &= -3\hat{\sigma}_{\hat{\tau}1ijk} \\ UCL &= +3\hat{\sigma}_{\hat{\tau}1ijk} \end{aligned} \qquad [5\text{-}28]$$

where: $\hat{\sigma}_{\hat{\tau}1ijk}$ is the standard error of the $\hat{\tau}_{1ijk}$ values for the *i*th saw configuration, *j*th side.

### 5.4.3.2 $Q_{\tau'}$ Chart

Whereas control charts for detecting taper were constructed based on measurements from the entire length of the board, control charts for detecting snipe and flare were based on measurements from the last 15 cm of each board only. Using this subset of the data yielded an estimate of the slope for each board in the *i*th saw configuration and *j*th side, $\hat{\tau}'_{1ijk}$, and control limits for this chart (the $Q_{\tau'}$ chart) were constructed similarly to [5-28]:

$$CL = 0$$
$$LCL = -3\hat{\sigma}_{\hat{\tau}'1_{ijk}} \qquad [5\text{-}29]$$
$$UCL = +3\hat{\sigma}_{\hat{\tau}'1_{ijk}}$$

where: $\hat{\sigma}_{\hat{\tau}'1_{ijk}}$ is the standard error of the $\hat{\tau}'_{1_{ijk}}$ values for the *i*th saw configuration and *j*th side.

### 5.4.4 Proposed Charts for Detecting Snake

To detect snake, the profile data were first de-trended using the estimated slope along the board from [5-27][35]:

$$y'_{ijklm} = y_{ijklm} - \hat{\tau}_{1ijk} x_1 \qquad [5\text{-}30]$$

Then, $y'_{ijklm}$ was decomposed into estimates of roughness and waviness via filtering. In the electronics industry, the waviness component of the surface profile is found using ISO standard 4287 (Raja *et al.* 2002). However, in wood products, these standards have not been widely applied, primarily because of the variability in wood surfaces (Funck *et al.* 1993). Instead, Gaussian, moving average, digital band-pass, and median filters have been preferred (Fujiwara *et al.* 2001). A filter was chosen subjectively with trial and error; a simple two-pass moving average filter with a window of 50 observations gave reasonable results and was computationally efficient for estimating waviness ($w_{ijklm}$). Roughness ($r_{ijklm}$) was then calculated as the difference between the de-trended surface profile and estimated waviness:

$$r_{ijklm} = y'_{ijklm} - w_{ijklm} \qquad [5\text{-}31]$$

Roughness can be measured in a variety of ways. The simplest measure is the arithmetic average roughness, which was estimated for each board and laser position (*kl*) for a particular saw configuration and side (*ij*):

---

[35] Under in-control conditions, this step is not necessary, i.e., the trend is zero.

$$r_{ijkl}^{(a)} = \frac{1}{n_{ijkl}} \sum_{m=1}^{n_{ijkl}} \left| r_{ijklm} - \bar{r}_{ijkl\cdot} \right| \qquad [5\text{-}32]$$

where: $\bar{r}_{ijkl\cdot}$ is the average of all roughness values ($r_{ijklm}$) for the $k$th board and $l$th laser position in the $i$th saw configuration and $j$th side.

The root mean square (RMS) roughness measures the average squared departure from the mean roughness:

$$r_{ijkl}^{(q)} = \sqrt{\frac{1}{n_{ijkl}} \sum_{m=1}^{n_{ijkl}} \left( r_{ijklm} - \bar{r}_{ijkl\cdot} \right)^2} \qquad [5\text{-}33]$$

The peak-to-peak roughness is the sum of the height of the highest and depth of the lowest points of roughness:

$$r_{ijkl}^{(p)} = \max_m (r_{ijklm} - \bar{r}_{ijkl\cdot}) + \min_m (\bar{r}_{ijkl\cdot} - r_{ijklm}) \qquad [5\text{-}34]$$

The arithmetic average waviness, RMS waviness, and peak-to-peak waviness for each board and laser position ($kl$) in a particular saw configuration and side ($ij$) are defined similarly to that of roughness:

$$w_{ijkl}^{(a)} = \frac{1}{n_{ijkl}} \sum_{m=1}^{n_{ijkl}} \left| w_{ijklm} - \bar{w}_{ijkl\cdot} \right| \qquad [5\text{-}35]$$

$$w_{ijkl}^{(q)} = \sqrt{\frac{1}{n_{ijkl}} \sum_{m=1}^{n_{ijkl}} \left( w_{ijklm} - \bar{w}_{ijkl\cdot} \right)^2} \qquad [5\text{-}36]$$

$$w_{ijkl}^{(p)} = \max_m (w_{ijklm} - \bar{w}_{ijkl\cdot}) + \min_m (\bar{w}_{ijkl\cdot} - w_{ijklm}) \qquad [5\text{-}37]$$

where: $\bar{w}_{ijkl\cdot}$ is the average of all waviness values ($w_{ijklm}$) for the $k$th board and $l$th laser position in the $i$th saw configuration and $j$th side.

For each roughness and waviness measure, the two values by laser position on each side of each board were averaged to obtain one value by board for each saw configuration and side, e.g., $\overline{w_{ijk\cdot}^{(p)}} = (w_{ijk1}^{(p)} + w_{ijk2}^{(p)})/2$. Using distributions of each roughness and waviness measure, quantile charts were developed. The control limits for these charts were found using a method similar to Levinson (1997). The probability distribution function of each quantity was fit to an appropriate theoretical probability distribution, and then the 0.1% and 99.9% quantiles were estimated.

Taking the peak-to-peak waviness as an example, a histogram of the estimated values of $\overline{w_{ijk\cdot}^{(p)}}$ was constructed using all boards in the lab scanned dataset by saw configuration and side. Because the number of boards in each type of saw configuration was <<300, an estimate of the 0.1th percentile could only be approximated. Thus, using the SAS procedure PROC UNIVARIATE, gamma, Weibull, and lognormal distributions were fit to approximate the distribution. These distributions were chosen because they are flexible and positively skewed, ensuring reasonable fits for the six roughness and waviness measures, which are strictly positive. These distributions were recommended by Levinson (1997), who favoured the gamma, as it is the continuous counterpart to the situation where defects arise following a Poisson distribution. Although formal goodness-of-fit tests are available, the best fitting distribution was chosen visually because of the importance of fitting the extreme tails of the distribution. The best fit distribution of the $\overline{w_{ijk\cdot}^{(p)}}$ values was chosen by sawing configuration and side (*ij*), and the lower and upper 0.1% quantiles were calculated ($Q_{ij}\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{0.1\%}$ and $Q_{ij}\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{99.9\%}$). Setting lower and upper limits at these values roughly approximated the false alarm rate of a 3-sigma control

chart[36]. The centreline was set at the median ($Q_{ij}\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{50\%}$). For the three measures of roughness and three measures of waviness, six control charts were constructed ($Q_{r_a}$, $Q_{r_q}$, $Q_{r_p}$, $Q_{w_a}$, $Q_{w_q}$, and $Q_{w_p}$) with control limits calculated for the *i*th saw configuration and *j*th side using Equations [5-38] – [5-43], respectively:

$$\begin{aligned} CL &= Q\left(\overline{r_{ijk\cdot}^{(a)}}\right)_{50\%} \\ LCL &= Q\left(\overline{r_{ijk\cdot}^{(a)}}\right)_{0.1\%} \\ UCL &= Q\left(\overline{r_{ijk\cdot}^{(a)}}\right)_{99.9\%} \end{aligned} \qquad [5\text{-}38]$$

$$\begin{aligned} CL &= Q\left(\overline{r_{ijk\cdot}^{(q)}}\right)_{50\%} \\ LCL &= Q\left(\overline{r_{ijk\cdot}^{(q)}}\right)_{0.1\%} \\ UCL &= Q\left(\overline{r_{ijk\cdot}^{(q)}}\right)_{99.9\%} \end{aligned} \qquad [5\text{-}39]$$

$$\begin{aligned} CL &= Q\left(\overline{r_{ijk\cdot}^{(p)}}\right)_{50\%} \\ LCL &= Q\left(\overline{r_{ijk\cdot}^{(p)}}\right)_{0.1\%} \\ UCL &= Q\left(\overline{r_{ijk\cdot}^{(p)}}\right)_{99.9\%} \end{aligned} \qquad [5\text{-}40]$$

$$\begin{aligned} CL &= Q\left(\overline{w_{ijk\cdot}^{(a)}}\right)_{50\%} \\ LCL &= Q\left(\overline{w_{ijk\cdot}^{(a)}}\right)_{0.1\%} \\ UCL &= Q\left(\overline{w_{ijk\cdot}^{(a)}}\right)_{99.9\%} \end{aligned} \qquad [5\text{-}41]$$

$$\begin{aligned} CL &= Q\left(\overline{w_{ijk\cdot}^{(q)}}\right)_{50\%} \\ LCL &= Q\left(\overline{w_{ijk\cdot}^{(q)}}\right)_{0.1\%} \\ UCL &= Q\left(\overline{w_{ijk\cdot}^{(q)}}\right)_{99.9\%} \end{aligned} \qquad [5\text{-}42]$$

---

[36] Assuming an underlying normal distribution, 3-Sigma control limits give a false alarm rate of ~0.27%, implying upper and lower quantiles for these limits at 0.0135% and 0.9865%. Since PROC UNIVARIATE automatically computes quantiles at 0.1% intervals, the closest quantiles (0.1% and 99.9%) were used.

$$CL = Q\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{50\%}$$
$$LCL = Q\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{0.1\%} \qquad [5\text{-}43]$$
$$UCL = Q\left(\overline{w_{ijk\cdot}^{(p)}}\right)_{99.9\%}$$

### 5.4.5 Evaluation of Proposed Charts

Using the lab scan data, the SAS procedure PROC MIXED gave estimates of all mixed-effects and components of variance from the model in Equation [4-1]. These parameter estimates were used to simulate LRS data arising from [4-1], as well as to construct control limits. Monte Carlo simulation was used to evaluate the performance of the charts under both in-control and out-of-control conditions. Average profiles by board, side, and laser position were simulated for each saw configuration and side using the following steps:

1. For each simulated board, a random board effect was generated: $B_{ijk} \sim N(0, \hat{\sigma}^2_{\beta_{ij}})$;

2. For each simulated laser position, a random laser position effect was generated: $L_{ijl} \sim N(0, \hat{\sigma}^2_{\lambda_{ij}})$;

3. For each simulated board and laser position, a random board × laser effect was generated: $BL_{ijkl} \sim N(0, \hat{\sigma}^2_{\beta\lambda_{ij}})$; and

4. Using the estimate of the overall average profile value by saw configuration and side ($\hat{\mu}_{ij}$), the simulated average profile by board × side × laser position was calculated as:

$$\tilde{\bar{y}}_{ijkl} = \hat{\mu}_{ij} + B_{ijk} + L_{ijl} + BL_{ijkl}$$

Simulated data were created for various subgrouping scenarios. One thousand sample groups were created, with 1, 4, 6, 8, 10, 12, 16, and 20 boards per group. To generate simulated profile observations within each group, board, side, and laser position, simulated autocorrelated errors ($e_{ijklm}$) were added to the simulated profile averages (from Step 4):

$$\tilde{y}_{ijklm} = \tilde{\bar{y}}_{ijkl} + e_{ijklm}$$

In Appendix I, ARIMA and SARFIMA models were fit to the autocorrelated errors in model [4-1] ($\varepsilon_{ijklm}$). The best fitting model forms were the ARIMA(1,1,1) for bandsawn boards and the ARIMA(0,1,1) for circular sawn boards. Thus, these model forms were used to generate simulated errors within board, side, and laser position for these types of saws. Although a seasonal fractional (SARFIMA) model produced the fit best for chipped boards, simulating data using this model requires storage of 100+ lags per observation and is therefore very computationally expensive to generate. A seasonal ARIMA (SARIMA) model provided nearly the same fit at a fraction of the computation time and was therefore used to generate the errors for chipped boards. The model forms used to simulate the $e_{ijklm}$ are listed in Table 5-2.

**Table 5-2. Model forms for simulated profile observations within-board, side, and laser position.**

| Saw Type | Model | Model Form |
|---|---|---|
| Bandsaw | ARIMA(1,1,1) | $(1-\phi_f B)(\varepsilon_{f,m} - \varepsilon_{f,m-1}) = \alpha_f + (1-\theta_f B)\upsilon_{f,m}$ |
| Circular saw | ARIMA(0,1,1) | $\varepsilon_{f,m} - \varepsilon_{f,m-1} = \alpha_f + (1-\theta_f B)\upsilon_{f,m}$ |
| Chipper-head | SARIMA(1,1,1)x(1,1,1)$_s$ | $(1-\phi_f B)(1-\Phi_f B^{s_f})(\varepsilon_{f,m} - \varepsilon_{f,m-1}) =$ $\alpha_f + (1-\theta_f B)(1-\Theta_f B^{s_f})\upsilon_{f,m}$ |

where: $f=ijkl$;
$B$ is the backshift operator, e.g., $(1 - \phi_f B)\varepsilon_{fm} = \varepsilon_{fm} - \phi_f \varepsilon_{f,m-1}$;
$\alpha_f$ is the intercept;
$\phi_f$ and $\Phi_f$ are the autoregressive parameters;
$\theta_f$ and $\Theta_f$ are the moving average parameters;
$s_f$ is the length of the period of seasonality; and
$\upsilon_{f,m}$ is a white noise error process, with $\upsilon_{f,m} \sim N(0,\sigma_{\upsilon_f}^2)$.

For each saw configuration and side ($ij$), an average estimated white noise error process variance was computed: $\bar{\hat{\sigma}}^2_{\upsilon_{ij..}}$. Random numbers, $u_{ijkl_m}$, were then drawn from a $N(0,\bar{\hat{\sigma}}^2_{\upsilon_{ij..}})$ population. Using the average of the estimated autoregressive, moving average, and intercept parameters by saw configuration and side ($\bar{\hat{\phi}}_{ij..}, \bar{\hat{\theta}}_{ij..}, \bar{\hat{\alpha}}_{ij..}$), simulated errors were generated for a bandsawn board × side × laser position as follows:

1. The first two errors were generated as the intercept plus random noise:

$$e_{ijkl_1} = \bar{\hat{\alpha}}_{ij\cdot\cdot} + u_{ijkl_1} \text{ and } e_{ijkl_2} = \bar{\hat{\alpha}}_{ij\cdot\cdot} + u_{ijkl_2}$$

2. The remaining observations were generated using the previous observations, e.g.,:

$$e_{ijkl_3} = e_{ijkl_2} + \bar{\hat{\phi}}_{ij\cdot\cdot}(e_{ijkl_2} - e_{ijkl_1}) + \bar{\hat{\alpha}}_{ij\cdot\cdot} + u_{ijkl_3} - \bar{\hat{\theta}}_{ij\cdot\cdot} u_{ijkl_2}$$

Simulated errors for chipped boards and circular sawn boards were generated similarly, using the models and parameters in Table 5-2.

Because ARIMA and SARIMA models are non-stationary with unbounded variance, the simulated series generated in this way can be extremely unstable. To ensure that simulations included only series that matched observed data, the variance and range of the simulated series were calculated and compared to the variance and range of the actual lab scanned data. If a simulated series was unrealistic, a replacement series was generated.

In order to assure that the simulated data series were representative of the lab scan data, descriptive statistics were computed. For each saw configuration × side combination, the average simulated profiles were computed by board and laser position. The variance of these averages was found for each laser position in each saw configuration × side combination. Values were then compared to that of the actual lab data. A procedure like this is necessary when data are generated with unbounded variance; however, this procedure is not wholly satisfactory (H. Joe[37], personal communication, 2004). Future research is needed for a better approach that matches inference and model fit, which is a difficult theoretical problem.

For evaluating the ability of the charts to detect specific sawing defects, out-of-control data were generated by modifying the simulated data to include each type of defect using graduated levels of severity. To generate boards with machine positioning problems and wedge, the simulated

---

[37] Professor, Department of Statistics, The University of British Columbia, Vancouver, B.C. Canada.

profile observations were modified uniformly along the length of the board. For machine positioning problems, this was accomplished by adding an amount, $\Delta_m$, to each board, side, and laser position observation. For wedge, an amount, $\Delta_w/2$, was added to the top laser position observations, while the same amount, $\Delta_w/2$, was subtracted from the bottom laser position observations.

To generate boards with taper, flare, and snipe, an amount was added incrementally to both the top and bottom laser position measurements as a function of their positions along the length of the board. Using $x_1$ to denote the distance along the board in cm, taper was generated by adding an amount $\Delta_t x_1/(244\text{ cm})$ to each profile measurement, so that the full amount of the increase $\Delta_t$ was reached at the end of the board (244 cm). Snipe and flare were generated by adding an amount $\Delta_f(x_1\text{-}229\text{ cm})/(15\text{ cm})$ to the last 15 cm of the board.

Snake was generated by adding a sinusoidal wave along the length of the board. A wave of period $P$ and amplitude $A$ was incorporated into the profile observations by adding an amount $A\sin\left(\frac{2\pi}{P}x_1\right)$ to each observation along the length of the board.

The values of $\Delta_m$, $\Delta_w$, $\Delta_t$, $\Delta_f$, $P$, and $A$ were chosen to represent a range of defect severities, from small to severe. These values were chosen in consultation with industry sawing experts (G.S. Shajer[38] and D.C. Wong[39], personal communication, 2004), and are shown in Table 5-3.

The nineteen proposed charts were evaluated using the simulated in-control and out-of-control data. For in-control data, the false alarm rate was evaluated. For out-of-control data, the rate of chart signaling was evaluated. To evaluate the out-of-control performance of the five charts

---

[38] Professor, Department of Mechanical Engineering, The University of British Columbia, Vancouver, B.C. Canada.

[39] Wood Machining Scientist, Forintek Canada Corporation, 2665 East Mall, Vancouver, B.C. Canada V6T 1W5

proposed for machine positioning problems and the six charts proposed for wedge, all eleven charts were subjected to both the machine positioning defect deviations and the wedge defect deviations. To evaluate the out-of-control performance of the two charts proposed for taper and flare/snipe, and the six charts proposed for snake, all eight charts were subjected to each of the taper, snipe/flare, and snake defect deviations. Within-board defects (taper, snipe/flare, snake) were not considered in the evaluation of charts proposed for machine positioning problems or wedge, as these defects were generated such that they had no affect on the summary statistics on which these particular charts are based. Machine positioning and wedge defects were not considered in the evaluation of charts proposed for taper, snipe/flare, or snake, as they were generated such that they had no affect on the trend, waviness, or roughness of observations within board and side.

**Table 5-3. Investigated ranges of defect severities.**

| Defect | Parameter | Range (mm) | Range (inches) |
|---|---|---|---|
| Machine positioning | $\Delta_m$ | ± 0.25, 0.50, 0.75, 1.00* | ± 0.010, 0.020, 0.030, 0.040* |
| Wedge | $\Delta_w$ | ± 0.25, 0.50, 0.75, 1.00* | ± 0.010, 0.020, 0.030, 0.040* |
| Taper | $\Delta_t$ | ± 0.5, 1.0, 1.5, 2.0 | ± 0.020, 0.040, 0.060, 0.080 |
| Snipe/flare | $\Delta_f$ | ± 0.5, 1.0, 1.5, 2.0 | ± 0.020, 0.040, 0.060, 0.080 |
| Snake | $P$ | 900, 1800, 2700 | 36, 72, 108 |
| | $A$ | ± 0.5, 1.0, 1.5, 2.0 | ± 0.020, 0.040, 0.060, 0.080 |

* Since this amount is added to both sides of the board, the change in thickness detected is 2 × Range

## 5.5 Results

Descriptive statistics for the simulation are compared to that of the original lab scanned data in Table 5-4. For most saw configuration × side × laser position combinations, the difference between standard deviations for simulation averages and those of the actual data was minimal. Ten combinations were within 0.01 cm of the actual, and only one combination (Saw Configuration BB, Side 2, Laser 1) deviated from the actual by more than 0.14 cm).

**Table 5-4. Descriptive statistics for simulated versus actual data.**

| Saw Configuration | Side | Laser Position | Average Profile Values (cm) | Standard Deviation of Average Simulation. | Actual | Difference |
|---|---|---|---|---|---|---|
| Bandsaw – Bandsaw (BB) | 1 | 1 | 2.571, 2.525, 2.646, . . . | 12.51 | 12.52 | 0.005 |
| | | 2 | 2.563, 2.515, 2.577, . . . | 12.59 | 12.52 | 0.075 |
| | 2 | 1 | 2.568, 2.582, 2.567, . . . | 12.81 | 12.54 | 0.269 |
| | | 2 | 2.543, 2.643, 2.536, . . . | 12.43 | 12.54 | 0.110 |
| Bandsaw – Chipper-head (CB) | 1 | 1 | 2.563, 2.668, 2.611, . . . | 16.52 | 16.51 | 0.009 |
| | | 2 | 2.538, 2.672, 2.601, . . . | 16.64 | 16.51 | 0.130 |
| | 2 | 1 | 2.673, 2.627, 2.638, . . . | 16.55 | 16.50 | 0.049 |
| | | 2 | 2.588, 2.652, 2.684, . . . | 16.45 | 16.50 | 0.054 |
| Chipper-head – Bandsaw | 1 | 1 | 2.665, 2.622, 2.678, . . . | 17.32 | 17.22 | 0.109 |
| | | 2 | 2.642, 2.61, 2.668, . . . | 17.24 | 17.22 | 0.027 |
| | 2 | 1 | 2.642, 2.706, 2.645, . . . | 17.21 | 17.28 | 0.073 |
| | | 2 | 2.643, 2.707, 2.638, . . . | 17.22 | 17.28 | 0.055 |
| Circular saw – Circular saw (RR) | 1 | 1 | 2.624, 2.59, 2.767, . . . | 22.11 | 22.10 | 0.006 |
| | | 2 | 2.62, 2.633, 2.612, . . . | 21.99 | 22.10 | 0.108 |
| | 2 | 1 | 2.625, 2.574, 2.657, . . . | 21.63 | 21.77 | 0.133 |
| | | 2 | 2.613, 2.601, 2.725, . . . | 21.80 | 21.77 | 0.035 |

### 5.5.1 Proposed Charts for Machine Positioning Problems

#### 5.5.1.1 In-control Evaluation

The in-control performance of the X-bar charts for individuals ($\text{X-bar}_{ind}$, Equation [5-1]) and groups ($\text{X-bar}_{grp}$, Equation [5-4]) is shown in Figure 5-2 by saw configuration, with $G$=1 corresponding to $\text{X-bar}_{ind}$. The expected proportion of out-of-control signals is shown as a reference line drawn at 0.27%. While the number of out-of-control signals by subgroup size varied, the overall average was on target at 0.3%, and there was no consistent trend or pattern by saw configuration or subgroup size.

The in-control results for the $MR_\beta$ and $R_{\beta_{grp}}$ charts (Equations [5-7] and [5-9], respectively) are shown in Figure 5-3 (with $G$=1 corresponding to the moving range chart). The expected number of out-of-control signals is shown as a reference line at 0.2%; on average, the number of simulated out-of-controls 0.25%. There was no obvious consistent trend by number of boards per subgroup or saw configuration.

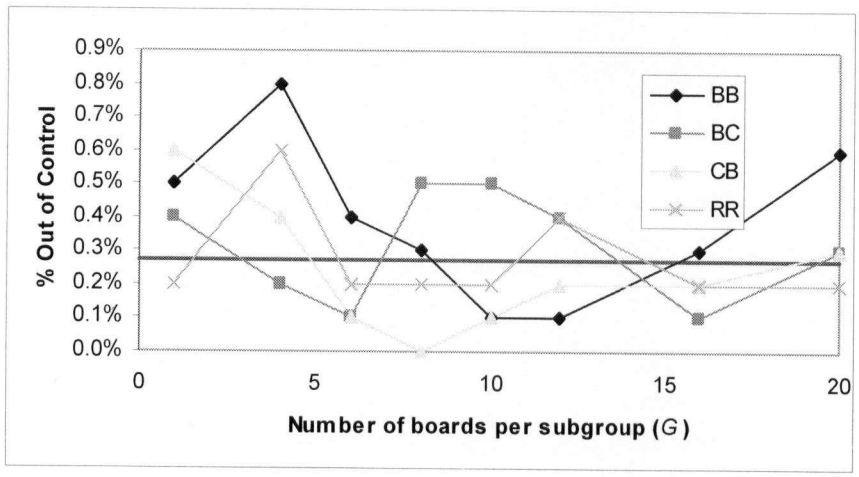



**Figure 5-2. Percent out of control for X-bar$_{ind}$ ($G$=1) and X-bar$_{grp}$ charts by subgroup size ($G$) and saw configuration[40].**

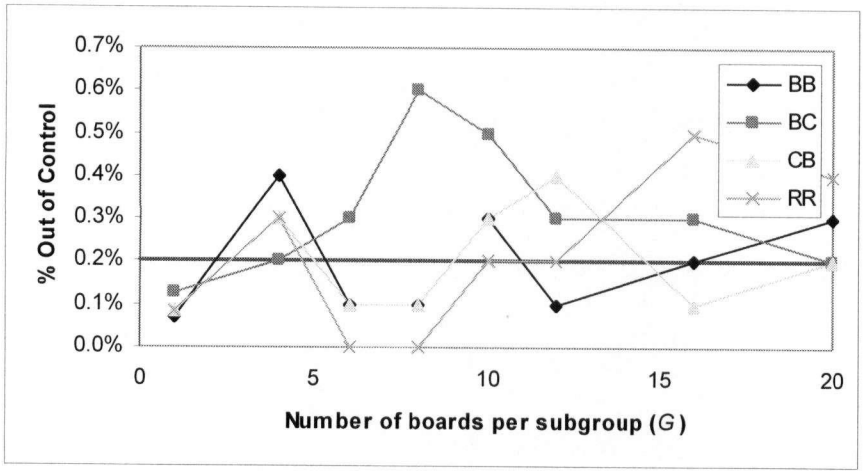



**Figure 5-3. Percent out of control for MR$_\beta$ ($G$=1) and R$_{\beta_{grp}}$ charts by subgroup size ($G$) and saw configuration.**

Results for the S$_\beta$ chart (Equation [5-11]) using in-control simulations are shown in Figure 5-4. This chart was affected by group size, performing best with moderate group sizes. At most group sizes, the number of out-of-control signals was well above the reference line for the expected number of out-of-controls (0.27%).

---

[40] The following abbreviations are used in figures throughout the remainder of this chapter: BB=Bandsaw-Bandsaw Saw Configuration, BC=Bandsaw-Chipper-head Saw Configuration, CB=Chipper-head-Bandsaw Saw Configuration, RR=Circular saw-Circular saw Saw Configuration, BB-1=Side 1 of BB Saw Configuration, BB-2=Side 2 of BB Saw Configuration, etc.

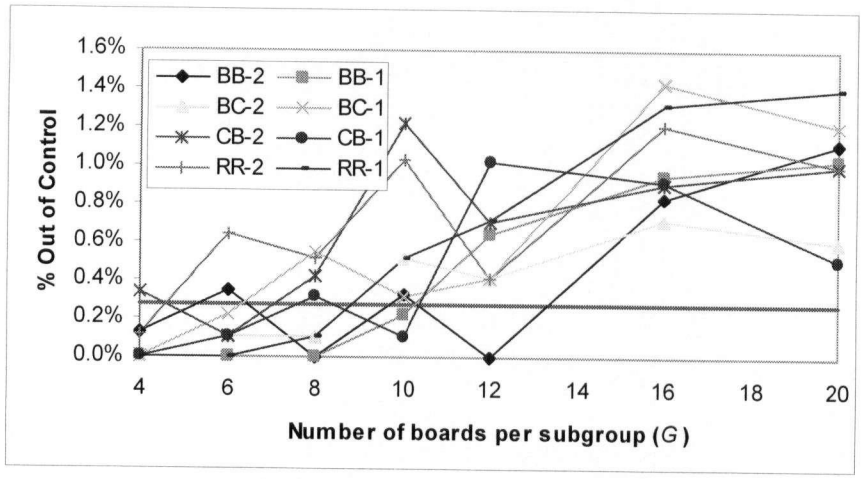



**Figure 5-4. Percent out of control for $S_\beta$ chart by subgroup size ($G$), saw configuration, and side.**

### 5.5.1.2 Out-of-control Evaluation

"Power curves" for the X-bar$_{ind}$ and X-bar$_{grp}$ charts for simulated BB boards are shown in Figure 5-5 by subgroup size (with X-bar$_{ind}$ shown as $G$=1). These curves show the power of the chart to detect machine positioning deviations over various values of $\Delta_m$. Since the response of the chart to negative values of $\Delta_m$ was almost identical to that of the positive, only positive values are shown. It is not surprising that the chart with the largest subgroup size had more out-of-control signals at a smaller level of deviation, since this chart was updated with information from twenty boards at a time. Since the behaviour of this chart was nearly identical for each type of saw configuration, results for the remaining configurations (BC, CB, and RR) are not shown.

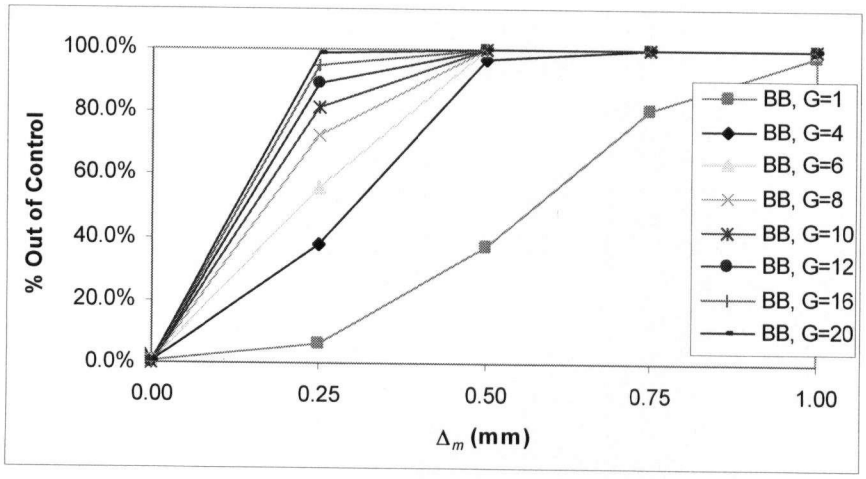



**Figure 5-5. Percent out of control for X-bar$_{ind}$ ($G$=1) and X-bar$_{grp}$ charts by subgroup size ($G$) and size of simulated machine positioning deviation ($\Delta_m$) for Saw Configuration BB.**

With the introduction of machine positioning deviations, the only indication in the $MR_\beta$ and $R_{\beta_{grp}}$ charts was a single out-of-control signal the first time the deviation was introduced. Moreover, the machine positioning change was missed by the $R_{\beta_{grp}}$ chart entirely unless the deviation was introduced mid-subgroup. The $S_\beta$ chart responded similarly; a single out-of-control was signaled at the first group with machine positioning problems only if the deviation was introduced in the middle of the subgroup. As in the X-bar charts, charts with larger sub-groups were more likely to signal, as their control limits were narrower.

None of the charts proposed to target other defects were impacted by the addition of machine positioning deviations in the simulations. This was a direct result of the method used to simulate machine positioning deviations, with a uniform impact over both sides of the board and both laser positions. The introduction of other types of defects (e.g., snake, taper), increased residual variation within each board, but this had no impact on charts for machine positioning problems.

### 5.5.2 Proposed Charts for Wedge

#### 5.5.2.1 In-control Evaluation

The in-control performance of the $R_{\lambda_{ind}}$ and $R_{\lambda_{grp}}$ charts (Equations [5-14] and [5-16]) are shown in Figure 5-6 (with $R_{\lambda_{ind}}$ shown as $G$=1). The number of out-of-control signals for the $R_{\lambda_{ind}}$ and $R_{\lambda_{grp}}$ charts averaged 0.1%, slightly lower than the reference line at the expected value of 0.2%. The rate of out-of-controls was stable, with no obvious trend by number of boards per subgroup or saw configuration.

The results of in-control simulations for the $MR_{\beta\lambda}$ and $R_{\beta\lambda_{grp}}$ charts (Equations [5-18] and [5-20]) are shown in Figure 5-7 by saw configuration and side, with $MR_{\beta\lambda}$ shown as $G$=1. $MR_{\beta\lambda}$ and $R_{\beta\lambda_{grp}}$ charts were produced for each laser position; however, since their performance was nearly identical by laser, the average performance is shown. On average, the number of out-of-controls

was on target with the expected value, shown as a reference line at 0.2%. Like the previous chart, there was no obvious trend by number of boards per subgroup or saw configuration.

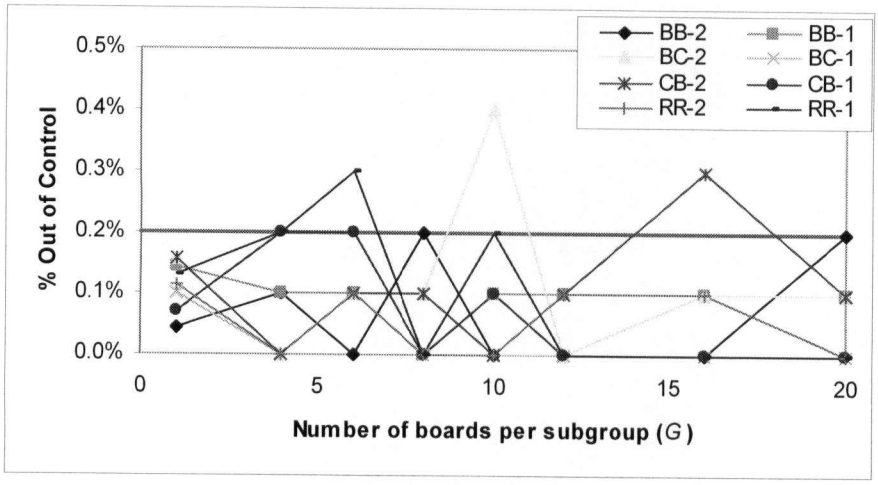



**Figure 5-6. Percent out of control for $R_{\lambda_{ind}}$ ($G$=1) and $R_{\lambda_{grp}}$ charts by subgroup size ($G$), saw configuration, and side.**

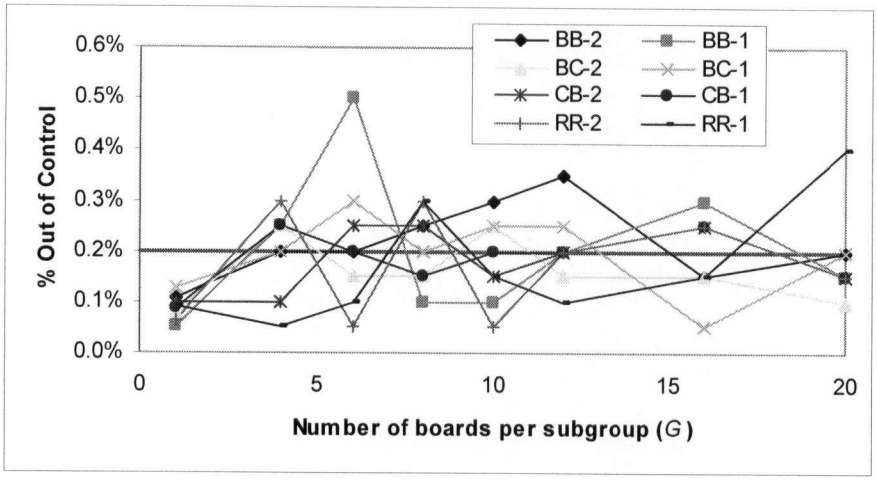



**Figure 5-7. Percent out of control for the $MR_{\beta\lambda}$ ($G$=1) and $R_{\beta\lambda_{grp}}$ charts by subgroup size ($G$), saw configuration, and side.**

The in-control $S_\lambda$ chart for the between-laser variation and $S_{\beta\lambda}$ chart for the board × laser interaction variation (Equations [5-22] and [5-25], respectively) are shown in Figure 5-8 and Figure 5-9, respectively. The $S_\lambda$ chart was greatly affected by subgroup size; it performed very poorly with small subgroups, but improved to an average out-of-control rate of ~4% with subgroup sizes of 14 or more. This improvement was due to increased degrees of freedom at larger subgroup sizes. The $S_{\beta\lambda}$ chart was more stable; the numbers of out-of-control signals were more in line with expected (0.2% on average).

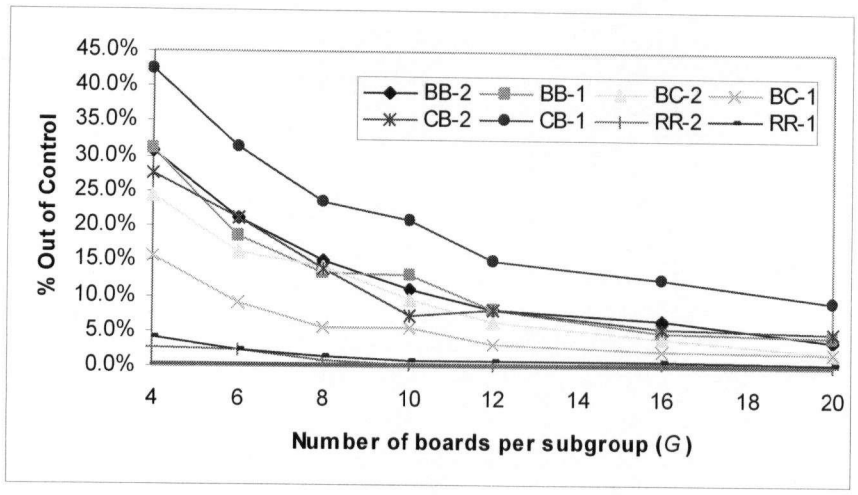



**Figure 5-8. Percent out of control for $S_\lambda$ chart by subgroup size ($G$), saw configuration, and side.**

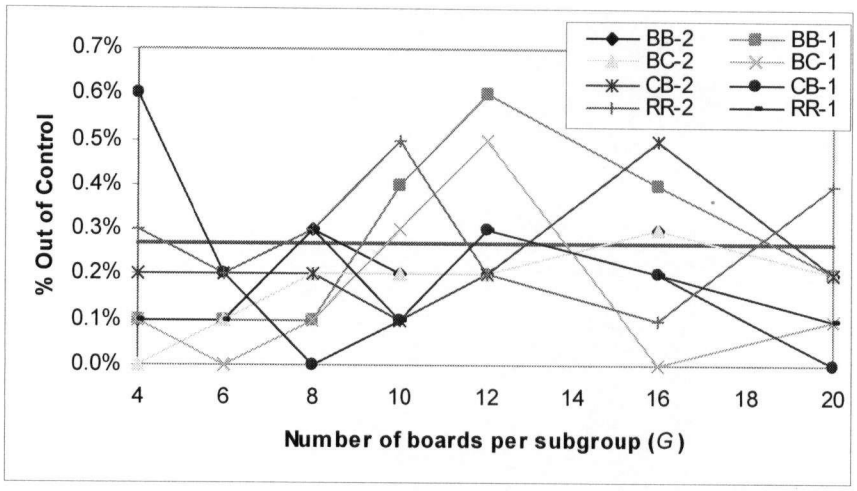



**Figure 5-9. Percent out of control for $S_{\beta\lambda}$ chart by subgroup size ($G$), saw configuration, and side.**

### 5.5.2.2 Out-of-control Evaluation

Because the $R_{\lambda_{ind}}$ and $R_{\lambda_{grp}}$ charts responded very similarly for different subgroup sizes, the result of introducing the wedge deviations is shown averaged over all subgroup sizes (Figure 5-10). As in the X-bar chart with $\Delta_m$, results for negative values of $\Delta_w$ were very similar to positive values, and thus only positive values are shown. As expected, larger deviations in $\Delta_w$ produced more out-of-control signals. The BB saw configurations were the most responsive to the size of $\Delta_w$, while the RR saw configurations were less responsive.

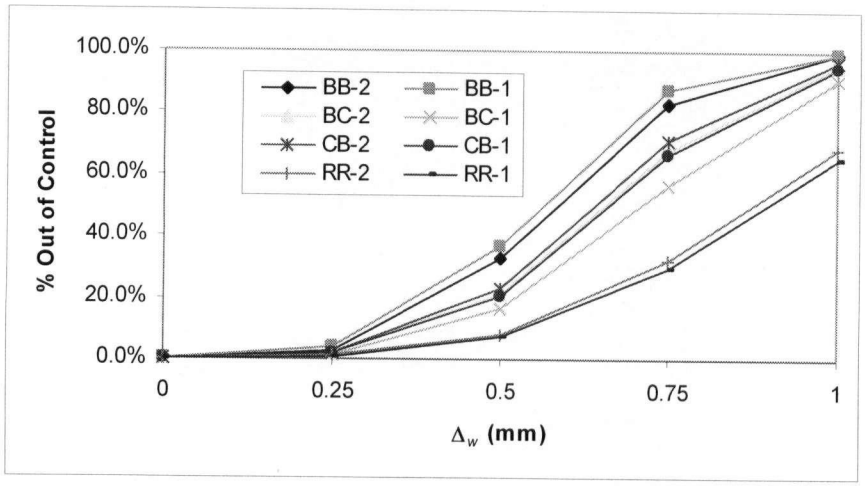



**Figure 5-10. Percent out of control for $R_{\lambda_{grp}}$ chart (average of all subgroups is shown) by size of simulated wedge deviation ($\Delta_w$), saw configuration, and side.**

The response of the $MR_{\beta\lambda}$ and $R_{\beta\lambda_{grp}}$ charts varied by subgroup size. Because results had similar patterns by saw configuration and side, only one combination (BB-Side 2) is shown (Figure 5-11). As in the previous charts, results for negative values of $\Delta_w$ were very similar to positive values, and thus only positive values are shown. When subgroup sizes were larger, the charts tended to signal more often, with more reasonable results obtained for smaller subgroups.

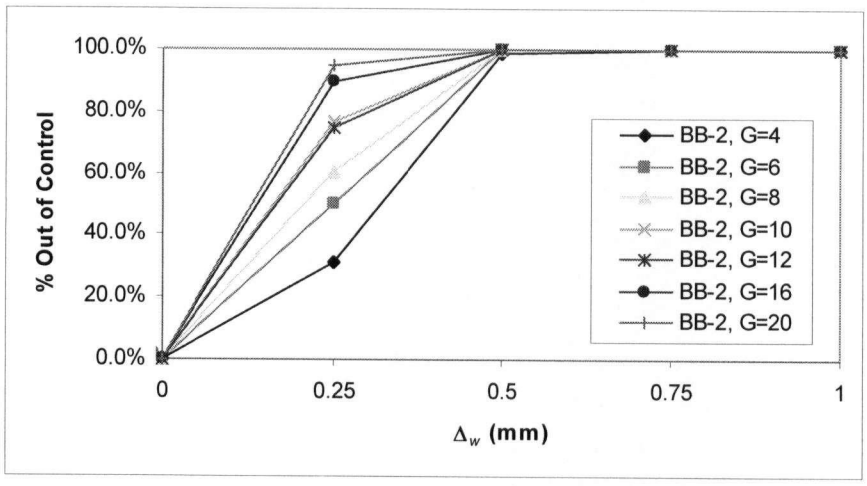



**Figure 5-11. Percent out of control for $R_{\beta\lambda_{grp}}$ chart by size of simulated wedge deviation ($\Delta_w$), for Saw Configuration BB-Side 2.**

The average result over all subgroups for the $MR_{\beta\lambda}$ and $R_{\beta\lambda_{grp}}$ charts is shown by saw configuration and side in Figure 5-12. As in the previous chart, larger deviations in $\Delta_w$ produced

more out-of-control signals, and results for the RR saw configurations were less responsive to the size of $\Delta_w$ than the other saw configurations.

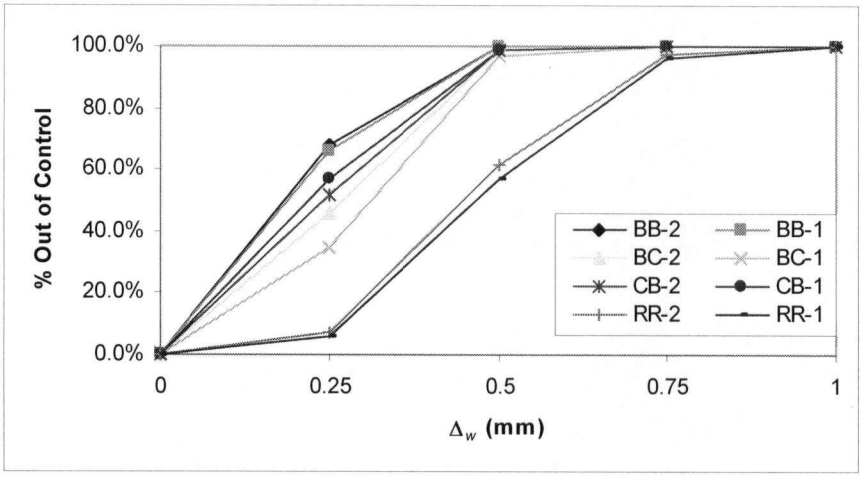



**Figure 5-12. Percent out of control for $R_{\beta\lambda_{grp}}$ chart (average of all subgroups is shown) by size of simulated wedge deviation ($\Delta_w$), saw configuration, and side.**

The results of introducing a wedge deviation to the $S_\lambda$ chart are shown in Figure 5-13 by size of deviation ($\Delta_w$) and subgroup size, for one saw configuration – side combination (RR-Side 1). As in the previous charts, results for negative values of $\Delta_w$ were similar to that of the positive values, and thus only positive values are shown.

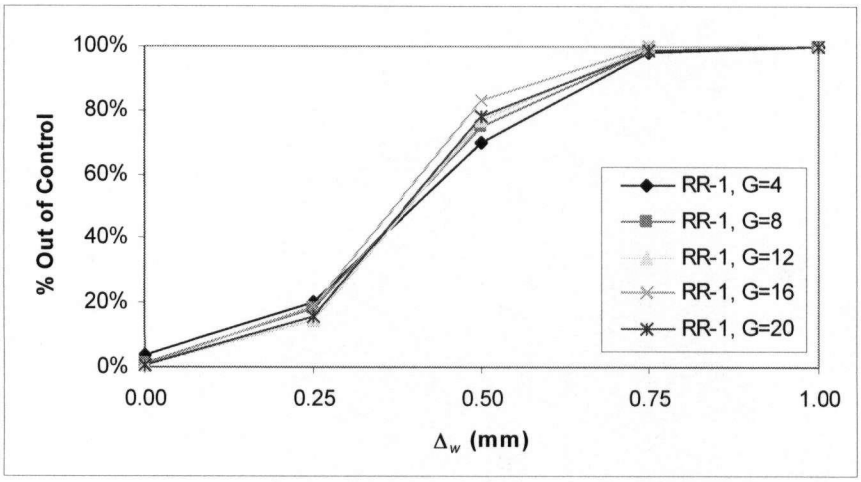



**Figure 5-13. Percent out of control for $S_\lambda$ chart by subgroup size ($G$) and size of simulated wedge deviation ($\Delta_w$), for Saw Configuration RR-Side 1.**

Although results for the $S_\lambda$ chart were affected by subgroup, the effect was similar within saw configuration and side and minimal in comparison to the effect of $\Delta_w$. Thus, results averaged

over all subgroups are shown in Figure 5-14 by saw configuration and side. As expected, larger deviations in $\Delta_w$ produced more out-of-control signals. As in the range charts, results for the RR saw configurations were less responsive to the size of $\Delta_w$ than those of other saw configurations.

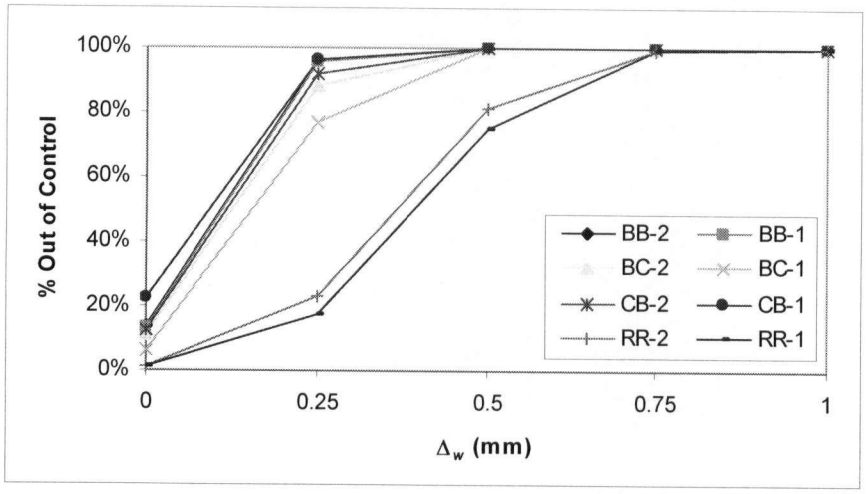



**Figure 5-14. Percent out of control for $S_\lambda$ chart (average of all subgroups is shown) by size of simulated wedge deviation ($\Delta_w$), saw configuration, and side.**

The introduction of wedge deviations did not affect the performance of the $S_{\beta\lambda}$ chart, and thus, out-of-control results are not shown for this chart. Because machine positioning deviations were introduced uniformly to each board and laser position, the differences between laser positions by board remained the same. Therefore, the introduction of machine positioning deviations did not affect the performance of any of the charts proposed for wedge. As in the case of the machine positioning charts, the addition of defects such as snake, which increased the residual variation within each board, side, and laser position, did not affect charts proposed for wedge.

### 5.5.3 Proposed Charts for Detecting Trend-related Defects

#### 5.5.3.1 In-control Evaluation

Estimated slopes along the lab scan boards were found for the whole board and only the last 15 cm (6 inches) of the board, for taper and snipe/flare, respectively. Using three-sigma limits, the in-control performance of the $Q_\tau$ and $Q_{\tau'}$ charts (Equations [5-28] and [5-29]) by saw configuration and side is shown in Table 5-5. For the $Q_\tau$ chart, 0.5% out-of-control signals

occurred, on average. For $Q_{\tau'}$ chart, the overall average was higher, which was driven by a high rate (4%) of out-of-control signals for circular-sawn boards.

**Table 5-5. Percent out of control for $Q_\tau$ and $Q_{\tau'}$ charts by saw configuration and side.**

| | | % Out of Control | |
|---|---|---|---|
| Saw Configuration | Side | $Q_\tau$ Chart | $Q_\tau$ Chart |
| Bandsaw–Bandsaw | 1 | 0.9% | 0.7% |
| | 2 | 1.1% | 0.9% |
| Bandsaw–Chipper-head | 1 | 0.2% | 1.4% |
| | 2 | 1.0% | 2.4% |
| Chipper-head–Bandsaw | 1 | 0.2% | 1.4% |
| | 2 | 0.5% | 0.8% |
| Circular saw–Circular saw | 1 | 0.1% | 4.0% |
| | 2 | 0.0% | 4.0% |
| Average | | 0.5% | 2.0% |

### 5.5.3.2 Out-of-control Evaluation

As in the previous charts, the impact of negative defect deviations was very similar to the positive, and they are not shown. As expected, increasing the amount of taper deviation ($\Delta_t$) increased the number of out-of-controls on the $Q_\tau$ chart (Figure 5-15). The RR saw configurations were the fastest to respond; this was not surprising given the tendency of the RR $Q_\tau$ charts to over-signal under in-control conditions.

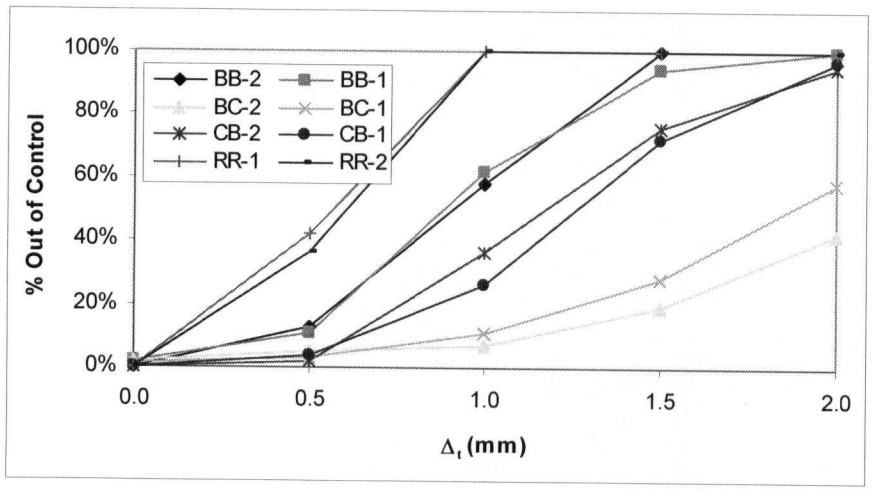



**Figure 5-15. Percent out of control for $Q_t$ chart by size of simulated taper deviations ($\Delta_t$), saw configuration, and side.**

The $Q_\tau$ chart was also evaluated under the various scenarios for snipe and flare, and for snake. When snipe was introduced, the impact to the $Q_\tau$ chart was minor, except at the highest

deviations (Figure 5-16). The impact of snake was substantial for some conditions, and there was an interaction between the amplitude and period of the snake (Figure 5-17). When the period of snake was long, the charts almost always signaled; when the period was short-medium, the chart only signaled with a large amplitude snake. For both snipe/flare and snake deviations, the impact on the $Q_\tau$ charts for the RR saw configurations was more pronounced than for other saw configurations. This was not surprising given the tendency of the RR $Q_\tau$ charts to over-signal under in-control conditions.

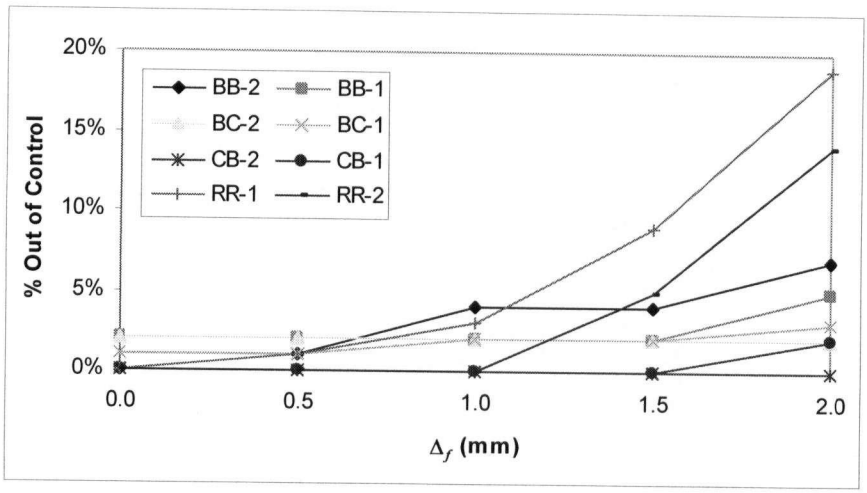



**Figure 5-16. Percent out of control for $Q_t$ chart by size of snipe/flare deviations ($\Delta_f$), saw configuration, and side.**

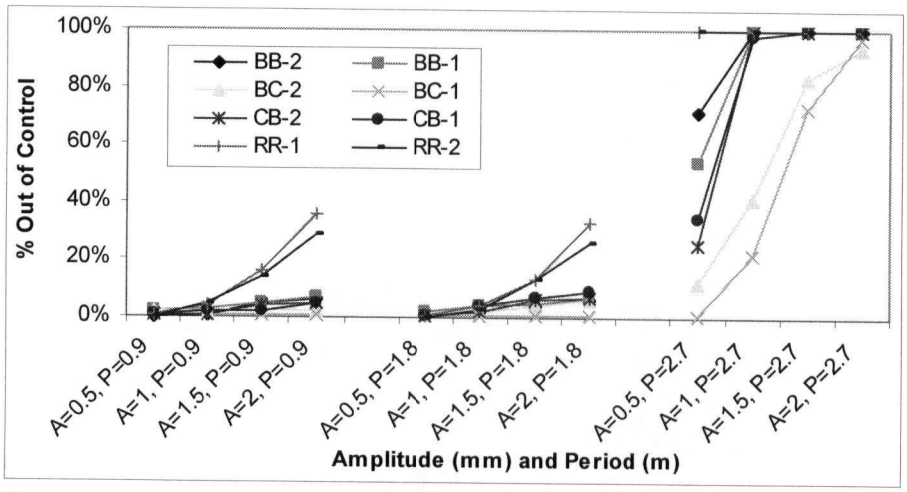



**Figure 5-17. Percent out of control for $Q_t$ chart by size of snake deviations (*A*=amplitude and *P*=period), saw configuration, and side.**

The $Q_{\tau'}$ chart was also evaluated under the various scenarios for taper, snipe, and flare. As expected, the addition of taper had almost no affect on the $Q_{\tau'}$ chart (not shown), and the addition

of snipe caused the chart to signal more frequently as a function of $\Delta_f$ (Figure 5-18). The addition of snake only affected the $Q_{\tau'}$ chart when the period was low (Figure 5-19). This is not unexpected, since flare/snipe are measured only over a short distance.

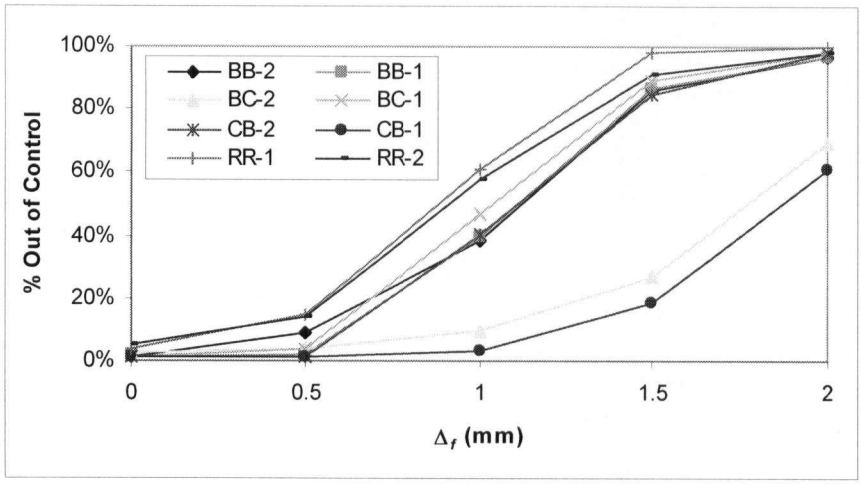



**Figure 5-18. Percent out of control for $Q_{\tau'}$ chart by size of snipe/flare deviations ($\Delta_f$), saw configuration, and side.**

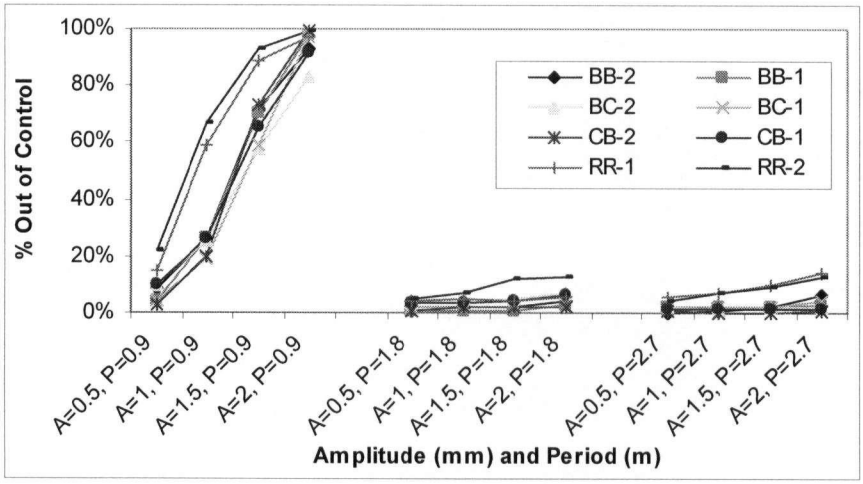



**Figure 5-19. Percent out of control for $Q_{\tau'}$ chart by size of simulated snake deviations (*P*=period and *A*=amplitude), saw configuration, and side.**

## 5.5.4 Proposed Charts for Snake

### 5.5.4.1 In-control Evaluation

The distributions of all six measures of waviness and roughness were fit and compared to distributions for the gamma, lognormal, and Weibull distributions. As shown in Figure 5-20, the distribution for average waviness ( $\overline{w_{ijk\cdot}^{(a)}}$ ) for BB-Side 1 boards was close to that of a normal,

whereas the best fitting distribution for BC-Side 2 boards was one that accommodated the long-tailed characteristics of the distribution. This result was not very surprising, since average waviness is a truncated distribution (positive-valued only); with smaller numbers of observations (circular-sawn and chipper-head data), the distribution is skewed, but approaches normality as the number of observations increases (bandsawn data). The gamma distribution described all the distributions fairly well, and this distribution can roughly approximate a normal distribution. Therefore, the gamma distribution was used to find percentiles and compute control limits for $\overline{w_{ijk\cdot}^{(a)}}$.

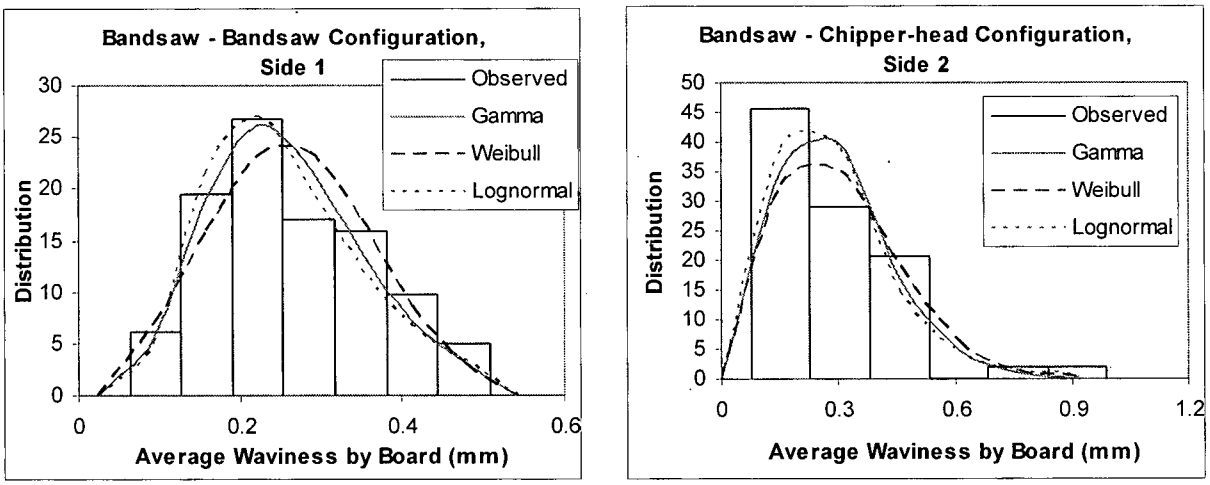



**Figure 5-20. Distribution of average waviness ($\overline{w_{ijk\cdot}^{(a)}}$) for Saw Configurations BB-Side 1 ($n$=41) and BC-Side 2 ($n$=24).**

As shown in Figure 5-21 and Figure 5-22, much the same pattern was observed for RMS waviness ($\overline{w_{ijk\cdot}^{(q)}}$) and peak-to-peak waviness ($\overline{w_{ijk\cdot}^{(p)}}$). The distribution of values from the bandsawn saw configuration-sides tended to be more Gaussian, whereas those from chipped or circular sawn boards tended to be more long-tailed. The gamma was therefore again chosen to find percentiles and compute control limits.

The three distributions for roughness were not as well-described by the fitted gamma, Weibull, or lognormal distributions (Figure 5-23 to 5–26). In particular, no distribution adequately fit the

long-tails of the three roughness distributions. However, the gamma was judgmentally selected to find percentiles and compute control limits for each of the three distributions based on visual assessment and because of its characteristic ability to take on many shapes.

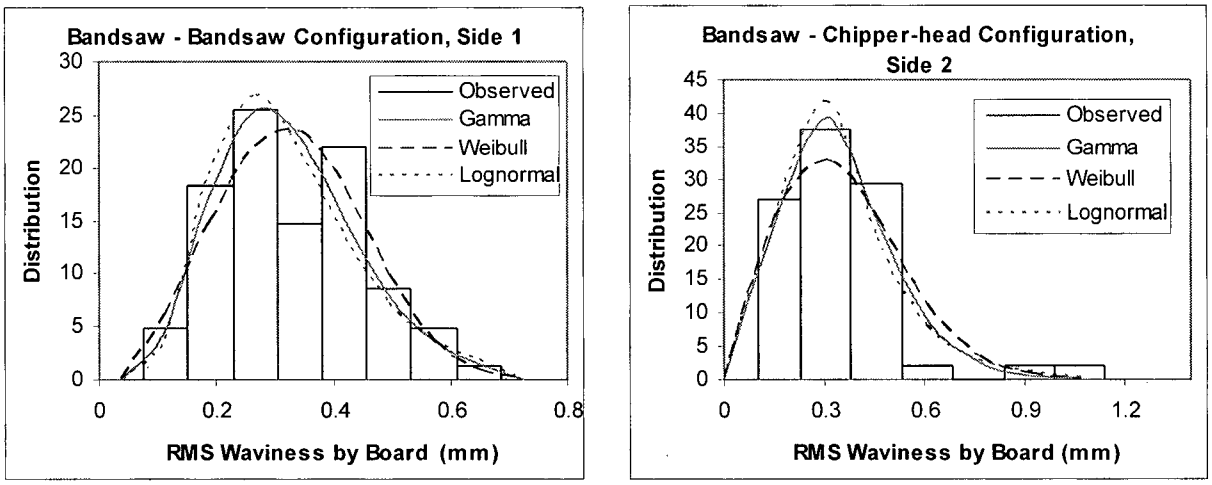



**Figure 5-21. Distribution of RMS waviness ($\overline{w_{ijk\cdot}^{(q)}}$) for Saw Configurations BB-Side 1 ($n$=41) and BC-Side 2 ($n$=24).**

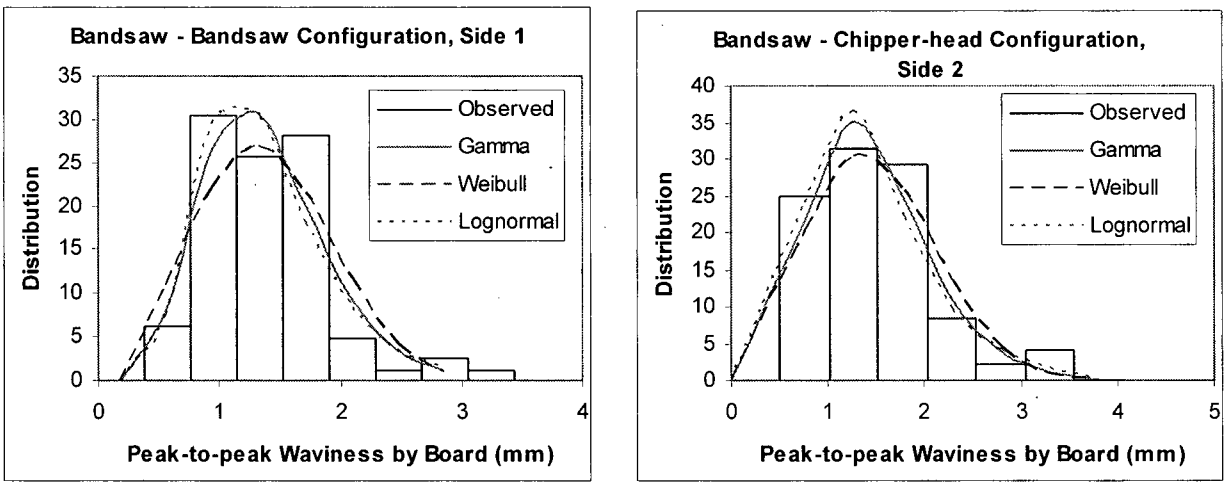



**Figure 5-22. Distribution of peak-to peak waviness ($\overline{w_{ijk\cdot}^{(p)}}$) for Saw Configurations BB-Side 1 ($n$=41) and BC-Side 2 ($n$=24).**

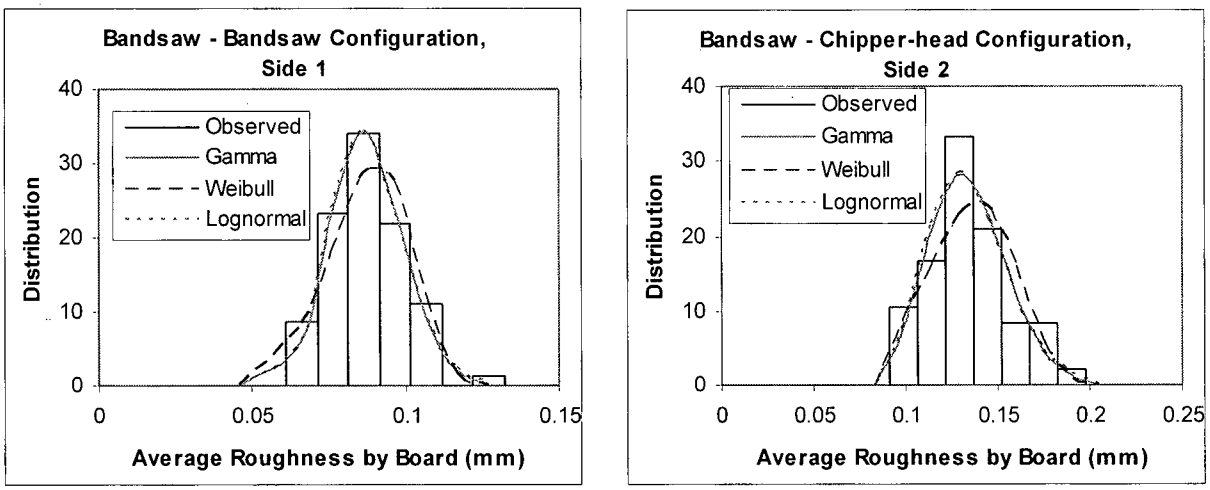



**Figure 5-23. Distribution of average roughness ($\overline{r_{ijk\cdot}^{(a)}}$) for Saw Configurations BB-Side 1 ($n$=41) and BC-Side 2 ($n$=24).**

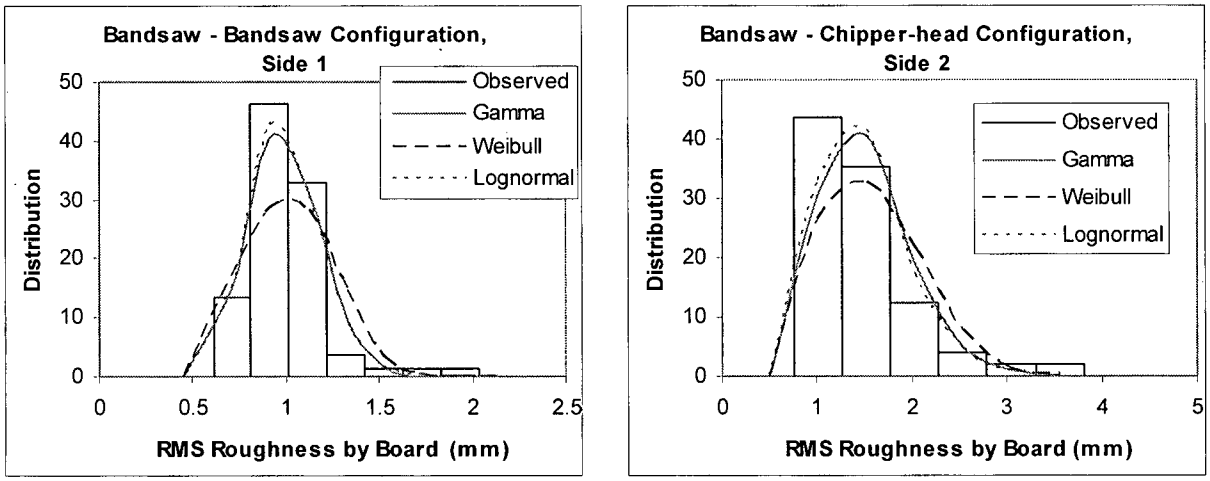



**Figure 5-24. Distribution of RMS roughness ($\overline{r_{ijk\cdot}^{(q)}}$) for Saw Configurations BB-Side 1 ($n$=41) and BC-Side 2 ($n$=24).**

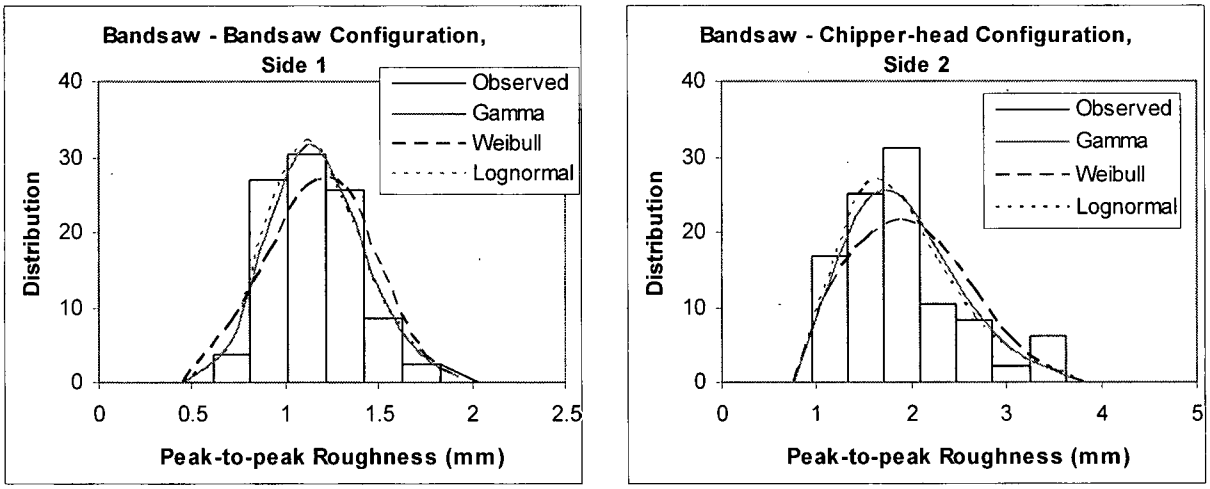



**Figure 5-25. Distribution of peak-to peak roughness ($\overline{r_{ijk\cdot}^{(p)}}$) for Saw Configurations BB-Side 1 ($n$=41) and BC-Side 2 ($n$=24).**

Using the gamma quantiles, the in-control performances of the three charts for snake based on waviness measures (Equations 5-38 – 5-40) were much better than those based on roughness measures (Equations 5-41 – 5-43). Overall, the $Q_{w_q}$ chart performed the best, with the most even rate of out-of-control signals over all saw configurations and sides (Table 5-6). The charts that monitored roughness were particularly poor, with about 100 times more out-of-controls than expected. Further investigation revealed that the number of out-of-control signals was balanced on the right and left tails; that is, the simulated data were both rougher and smoother than the lab data used to generate it. This increased variability in the roughness of the simulated data, however, could be an artifact of the method used to generate it.

**Table 5-6. Percent out of control for proposed charts for snake by saw configuration and side.**

| | | Percent Out of Control | | | | | |
|---|---|---|---|---|---|---|---|
| Saw Configuration | Side | $Q_{w_a}$ | $Q_{w_q}$ | $Q_{w_p}$ | $Q_{r_a}$ | $Q_{r_q}$ | $Q_{r_p}$ |
| Bandsaw–Bandsaw | 1 | 0.2% | 0.1% | 0.2% | 17.4% | 15.5% | 5.8% |
| | 2 | 0.1% | 0.0% | 0.0% | 24.9% | 23.3% | 16.6% |
| Bandsaw–Chipper-head | 1 | 0.7% | 0.3% | 0.0% | 33.3% | 29.5% | 14.7% |
| | 2 | 0.2% | 0.2% | 1.1% | 32.0% | 30.3% | 8.9% |
| Chipper-head–Bandsaw | 1 | 0.0% | 0.0% | 0.2% | 39.1% | 35.3% | 8.4% |
| | 2 | 1.0% | 0.8% | 0.9% | 7.1% | 5.8% | 9.0% |
| Circular saw–Circular saw | 1 | 0.0% | 0.0% | 0.1% | 61.0% | 54.9% | 12.6% |
| | 2 | 0.0% | 0.0% | 0.0% | 26.2% | 28.6% | 24.7% |
| Average | | 0.3% | 0.2% | 0.3% | 30.1% | 27.9% | 12.6% |

### 5.5.4.2 Out-of-control Evaluation

The six charts proposed for monitoring snake were evaluated with simulated deviations for snipe/flare and snake. (Since data are detrended prior to measuring roughness and waviness, taper did not affect these charts.) All three waviness charts responded very similarly to the addition of simulated snake defects, and thus only one chart ($Q_{w_q}$) is shown. As the amplitude of snake increased, the charts signaled more often (Figure 5-26). On the other hand, the period of snake had a smaller and opposite impact on the charts; with smaller periods, the $Q_{w_q}$ chart was more responsive, signaling on average 50% when the amplitude was 1 mm (0.04 inch).

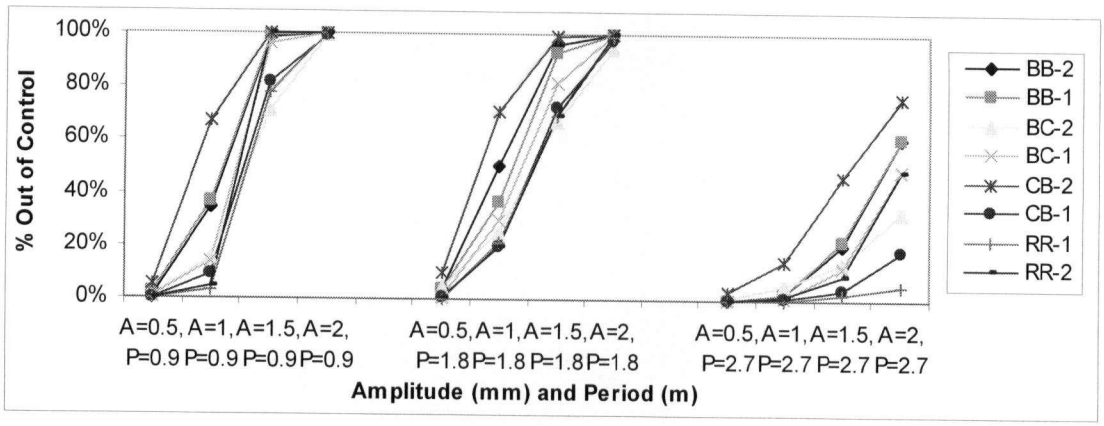



**Figure 5-26. Percent out of control for $Q_{w_q}$ chart by size of simulated snake deviations (*P*=period and *A*=amplitude), saw configuration, and side.**

Snipe and flare had a very small impact on the $Q_{w_a}$ and $Q_{w_q}$ charts. Since their responses were very similar, only one ($Q_{w_q}$) is shown (Figure 5-27). The $Q_{w_p}$ chart was more sensitive, with more out-of-control signals at 1.5 mm snipe/flare deviation (Figure 5-28). Positive and negative values of $\Delta_f$ gave similar results, and therefore only positive values are shown.

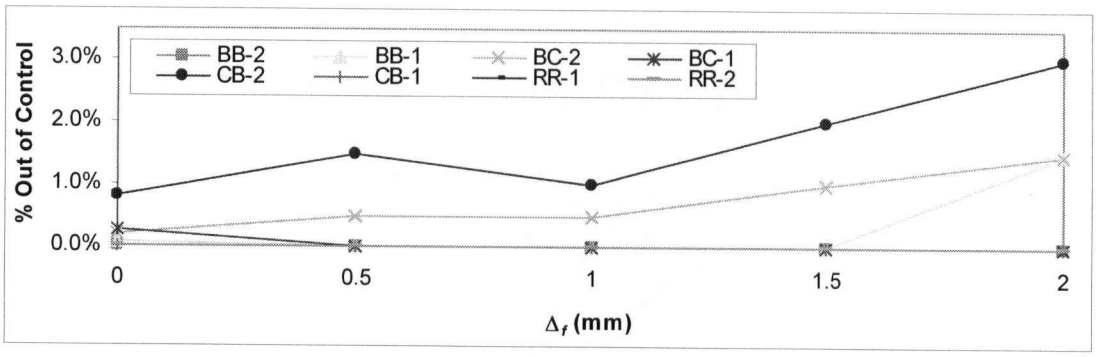



**Figure 5-27. Percent out of control for $Q_{w_q}$ chart by size of simulated snipe deviations ($\Delta_f$), saw configuration, and side.**

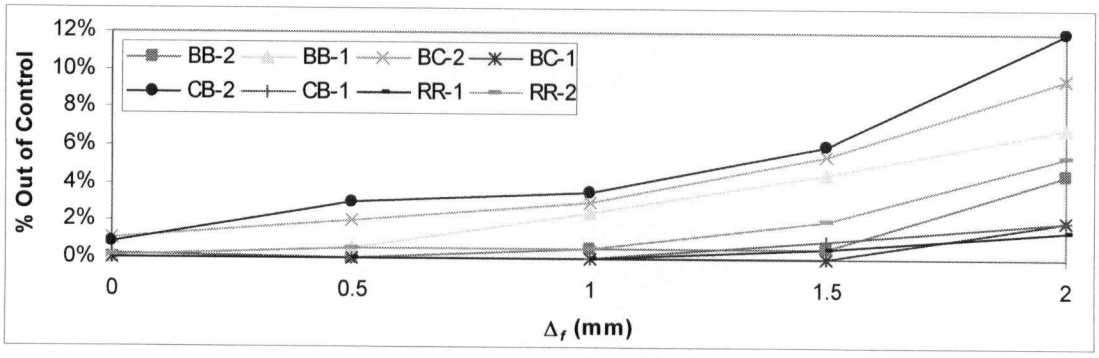



**Figure 5-28. Percent out of control for $Q_{w_p}$ chart by size of simulated snipe deviations ($\Delta_f$), saw configuration, and side.**

As expected, the $Q_{r_a}$, $Q_{r_q}$, $Q_{r_q}$ charts showed virtually no change with the addition of snake and flare/snipe defects, and thus, they are not shown.

## 5.6 Discussion

### 5.6.1 Machine Positioning Problems

Five charts were evaluated for their adequacy in detecting machine positioning problems. The X-bar$_{ind}$ and X-bar$_{grp}$ charts based on individuals and on subgroups, respectively, performed equally well under in-control conditions. However, the out-of-control response to specific defects varied by size of subgroup. The X-bar$_{grp}$ chart was more likely to signal when larger subgroups were used, but there is a trade-off in the amount of time necessary to accumulate larger subgroups for sampling. Moreover, with larger subgroup sizes, even the smallest shift in machine positioning ($\Delta_m$=0.25 mm) caused the charts to signal over 50% of the time. Given normal mill operating conditions, charts constructed with large subgroups may be too sensitive. On the other hand, this sensitivity may be advantageous, as $\Delta_m$ must be sustained throughout the subgroup to achieve the reported results.

Although the $MR_\beta$ and $R_{\beta_{grp}}$ charts performed well during in-control conditions, a shift in machine positioning was indicated by a single out-of-control signal only, and this signal was only noted if the change occurred mid-subgroup. This is in line with findings from Woodall *et al.* (2000) and others, who reported that the standard moving range chart (which is applied to the same data as the X-bar chart) is not effective in detecting sustained changes in a process. In this type of situation, a CUSUM chart may be more appropriate.

The $S_\beta$ chart for the between-board variation had similar issues for out-of control signals. Moreover, its in-control performance was poor and appeared to be affected by the size of subgroup chosen. This result was not unexpected, given that the Satterthwaite procedure works

well only when $MS_{\beta_{ij}}/MS_{\beta\lambda_{ij}} > F_{(G-1,\, G-1;\, 0.99865)} * F_{(G-1,\, G-1;\, 0.5)}$ (Gaylor and Hopper 1969); however, this condition is only met for large values of $G$, and with only a few types of saw configurations (e.g., CB-side 1). A similar result was reported by Maness *et al.* (2004) in simulation studies using between-board variation values in the same range. This indicates that the $S_\beta$ chart is inappropriate for the variance components found in typical mill data.

### 5.6.2 Wedge

The $S_\lambda$ chart met Gaylor and Hopper's condition only when monitoring charts for saw configuration RR with larger values of $G$. On the other hand, the $S_{\beta\lambda}$ chart performed more to expectation on average, as it met the conditions for every saw configuration and subgroup size. The $S_{\beta\lambda}$ chart, however, did not respond well for out-of-control conditions, remaining virtually unchanged with the addition of wedge defects.

Under out-of-control conditions, the $R_{\lambda_{ind}}$ and $R_{\lambda_{grp}}$ chart were less sensitive than their board × laser counterparts (the $MR_{\beta\lambda}$ and $R_{\beta\lambda_{grp}}$ charts), signaling only when $\Delta_w$ was > 0.5 mm. This response is more reasonable, given the range of normal mill conditions. On the other hand, the rate at which out-of-controls were signaled while the process was in control was slightly lower than expected for these charts, which may be caused by non-normality in the profile data (Burr 1967). Since a slight departure from normality appears to be a characteristic of these data, adjustments to chart limits to account for non-normality should be made if these charts are used operationally.

### 5.6.3 Trend-related Defects

The in-control performance of the $Q_\tau$ chart tended to be better than that of the $Q_{\tau'}$ chart. Although both charts were derived using the same regression method, more observations were used in the derivation of the parameters in the $Q_\tau$ chart, which would lead to more precise results.

The somewhat higher number of out-of-controls varied by saw configuration, and further investigation revealed that the data underlying the chart for circular-sawn data (which signaled at a rate of 4% when the process was in control) was platykurtotic. Because of the fatter tails in this distribution, more out-of-control signals resulted. Snipe and flare, therefore, may be better controlled with a chart based on an alternate distribution, such as the Pareto distribution.

When snipe/flare deviations were introduced, the $Q_{\tau'}$ chart was not very sensitive, owing to its wider control limits. The $Q_{\tau}$ chart was very sensitive to taper deviations for some saw configurations (e.g., RR), indicating ~40% out-of-controls for $\Delta_f$=0.5 mm. Both the $Q_{\tau}$ and $Q_{\tau'}$ charts responded to snake deviations, but only at its extremes. Therefore, it may be necessary to examine multiple charts to determine the exact cause of out-of-control signals in these charts.

### 5.6.4 Snake

Charts for snake are based on measures that have never been used in lumber size control, and their distributions are not well understood. While the charts based on waviness gave in-control results close to expected, those based on roughness performed less consistently than other charts. However, roughness may be of little concern for some sawmilling applications, e.g., when boards are sent to the planer mill as part of processing. Where applicable, further study should be devoted to understanding how these charts perform with normal sawing conditions. Moreover, these charts could be based on economic limits or capability analyses. Based on the specifications of the product's end-user, limits could be developed that better describe how the product compares to consumer's expectations.

### 5.6.5 Establishing, maintaining and updating charts

Like traditional SPC charts, the proposed charts introduced here will require an initial sample to establish control limits. Mill-specific or machine-specific target sizes, as well as estimates of the components of variance, will be necessary for the introduction of these new charts. Although

many of these variance components are unfamiliar to mill staff, the derivation of such components can be performed with statistical software. Periodic maintenance of the charts will also be necessary to ensure optimal chart performance. As mills become more familiar with the new SPC system, chart limits will necessarily be updated to account for process improvements and/or changes to product specifications.

## 5.7 Conclusions and Recommendations

In total, nineteen proposed charts were evaluated for use with real-time LRS data in lumber manufacturing. Both the in-control performance and out-of-control response was evaluated with respect to the specific sawing defects of machine positioning problems, wedge, taper, snipe, flare, and snake.

Of the five charts presented for detecting machine positioning problems, the best-performing charts were the X-bar charts, using limits based on the components of variance of the statistical model. The X-bar$_{ind}$ chart provided adequate in-control performance, and was not overly sensitive to minor changes in machine positioning deviations. Therefore, it is recommended for use in real-time SPC.

Six charts were presented for detecting wedge. Although the $S_{\beta\lambda}$ chart performed adequately for in-control boards and met the conditions necessary for Satterthwaite's procedure, it was not sensitive to wedge deviations. The range charts for laser position averages ($R_{\lambda_{ind}}$ and $R_{\lambda_{grp}}$) are recommended over the other charts, as their out-of-control performance was not overly sensitive, and they gave consistent out-of-control rates over all subgroup sizes. However, these charts tended to signal at a rate slightly lower than the expected 0.2% rate, which was likely due to non-normality in the data. Further study should be made to quantify this difference and adjust the values of $D_{0.001}$ and $D_{0.999}$ that are used with this chart accordingly.

While the $Q_\tau$ chart performed adequately under in-control conditions, its out-of control performance was overly sensitive when using 3-sigma limits based on the normal distribution. The $Q_{\tau'}$ chart also suffered from over-sensitivity and also signaled at a high rate during in-control conditions. Both of these charts are good candidates for economically based limits. Taper, snipe, and flare are quantities that are easy to measure accurately with laser equipment and limits could be derived that are tied to customer specifications or machinery limitations, such as planer settings. Field research involving mill staff could lead to more practically based limits based on economic constraints, as well as statistics. On the other hand, the $Q_\tau$ and $Q_{\tau'}$ charts were prone to signaling in the presence of snake. Thus, as designed, these charts will not be useful in isolation for suggesting definitive causes for out-of-control signals.

The snake charts presented are based on the concepts of roughness and waviness, which are new to the field of SPC in lumber manufacturing. Although snake has been studied extensively, measurements of snake have not been monitored as a quality characteristic. Development of limits for roughness and waviness should be based on the end uses of the product. Economic limits could be derived that are custom-tailored to the specifications desired by the consumer. For instance, for laminate stock, boards must have 0% planer skip. With this specification limit and an estimate of the percent shrinkage that will occur in kiln-drying, a green spec-limit can be obtained and a routine for monitoring the absolute deviation from this green specification limit can be calculated. The use of a peak-to-peak waviness chart with economic limits is suggested as a starting point for mills to gain an understanding of how snake affects their final product.

The use of real-time LRS data is a reality for many mills today, and will likely be more common in the near future. Systems developers must update statistical algorithms to take into account the vastly different data acquired by these devices. Moreover, systems should be designed with non-traditional control charts to take advantage of the opportunity for better sawing defect

recognition. Recognition of sawing defects will relate out-of-control signals to specific causes and help mills to more efficiently find the source of quality problems.

## 5.8 Literature Cited


Baseville, M. 1988. Detecting changes in signals and systems - a survey. Automatica 24: 309-326.

Burr, I. W. 1967. The effect of non-normality on constants for X-bar and R charts. Industrial Quality Control 24: 563-569.

Does, R. J. M. M., K. C. B. Roes and A. Trip. 1999. Handling multivariate problems with univariate control charts. Journal of Chemometrics 13: 353-369.

Fujiwara, Y., Y. Fujii, Y. Sawada and S. Okumura. 2001. Development of a parameter to reflect the roughness of a wood surface that corresponds to tactile roughness. Holz als Roh- und Werkstoff 59: 351-355.

Funck, J. W., J. B. Forrer, D. A. Butler, C. C. Brunner and A. G. Maristany. 1993. Measuring surface roughness on wood: a comparison of laser scatter and stylus tracing approaches. *In* (1992) (Ed.). SPIE - The International Society for Optical Engineering, Bellingham, WA. SPIE Vol. 1821: pp. 173-184.

Gaylor, D. W. and F. N. Hopper. 1969. Estimating the degrees of freedom for linear combinations of mean squares by Satterthwaite's formula. Technometrics 11(4): 691-706.

Grimshaw, S. D. and F. B. Alt. 1997. Control charts for quantile function values. Journal of Quality Technology 29(1): 1-7.

Harter, H. L. 1960. Tables of range and studentized range. The Annals of Mathematical Statistics 31(4): 1122-1147.

Lemaster, R. L. and J. B. Taylor. 1999. High speed surface assessment of wood and wood-based composites. Presented at 14th International Wood Machining Seminar. September 12-19, Epinal, France.

Levinson, W. A. 1997. Watch out for non-normal distributions of impurities. Chemical Engineering Progress 93(5): 70-76.

Maness, T. C. 1993. Real-time quality control system for automated lumbermills. Forest Products Journal 43(7/8): 17-22.

Maness, T. C., R. A. Kozak and C. L. Staudhammer. 2003. Applying real-time statistical process control to manufacturing processes exhibiting between and within part size variability in the wood products industry. Quality Engineering 16(1): 113-125.

Maness, T. C., R. A. Kozak and C. L. Staudhammer. 2004. Reliability testing of statistical process control procedures for manufacturing with multiple sources of variation. Wood and Fiber Science 36(3): 443-458.

Mastrangelo, C. M., J. M. Porter and R. V. Baxley. 2001. Multivariate process monitoring for nylon fiber production. *In* Frontiers in Statistical Quality Control. T. Wilrich (Ed.). Springer-Verlag: pp. 228-246.

Montgomery, D. C. 2001. Introduction to Statistical Quality Control. 4th ed. New York, John Wiley & Sons.

Montgomery, D. C. and C. M. Mastrangelo. 1991. Some statistical process control methods for autocorrelated data. Journal of Quality Technology 23: 179-193.

Negiz, A., E. S. Lagergren and A. Cinar. 1994. Statistical quality control of multivariable continuous processes. *In* Proceedings of American Control Conference, June 29-July 1, Baltimore, MD. American Automatic Control Council, Evanston, IL. pp. 1289-1293.

Padgett, W. J. and J. D. Spurrier. 1990. Shewhart-type charts for percentiles of strength distributions. Journal of Quality Technology 22(4): 283-290.

Raja, J., B. Muralikrishnan and S. Fu. 2002. Recent advances in separation of roughness, waviness, and form. Precision Engineering 26(2): 222-235.

Rasmussen, H. K., R. A. Kozak and T. C. Maness. 2004. An analysis of machine caused lumber shape defects in British Columbia sawmills. Forest Products Journal 54(6): 47-56.

Sandak, J., C. Tanaka and T. Ohtani. 2003. Evaluation of surface smoothness by a laser displacement sensor 1: effect of wood species. Journal of Wood Science 49: 305-311.

SAS Institute. 2002. SAS/STAT User's Guide, Version 8. Cary, NC, SAS Publishing.

Shewhart, W. A. 1931. Economic Control of Quality of Manufactured Product. New York, NY, Van Nostrand.

Vermaat, M. B., R. A. Ion, R. J. M. M. Does and C. A. J. Klaassen. 2003. A comparison of Shewhart individuals control charts based on normal, non-parametric, and extreme-value theory. Quality and Reliability Engineering International 19: 337-353.

Wheeler, D. J. 1995. Advanced Topics in Statistical Process Control. Knoxville, TN, SPC Press.

Willemain, T. R. and G. C. Runger. 1996. Designing control charts using an empiricial reference distribution. Journal of Quality Technology 28: 31-38.

Woodall, W. H., R. W. Hoerl, A. C. Palm and D. J. Wheeler. 2000. Controversies and contraditions in statistical process control. Journal of Quality Technology 32(4): 341-350.

Young, T. M. and P. M. Winistorfer. 1999. SPC. Forest Products Journal 49(3): 10-17.

# Chapter 6 Conclusions and Directions for Future Research

Canada is the largest exporter and second largest producer of softwood lumber in the world. In 2002, wood products manufacturing provided over 177,000 Canadian jobs, and over $10 billion in softwood lumber products were produced (Anonymous 2004). Clearly, the industrial production of lumber and other wood products is an important part of the Canadian economy.

At the same time, environmental and economic pressures have led to scarcity of raw materials for lumber production. Local timber shortages and the Canada-US softwood lumber dispute have forced the closure of many mills and led to increased concerns about the costs of production (Random Lengths 2004). Sawmills must get increased value from every log processed in order to stay competitive.

Technologically advanced SPC tools, like laser range sensors (LRSs), can help Canadian sawmills to increase log recovery by improving size control programs. This thesis provides methods for processing of real-time LRS data, as well as a statistical model for a SPC system involving multiple LRSs. The major challenges for real-time SPC with LRSs involve filtering the LRS data for non-sawing defects and measurement error, and designing a system of control charts that uses an appropriate model to utilize the wealth of data that these systems provide.

## *6.1 Filtering for Non-sawing Defects and Measurement Error*

In the case of SPC data collection, the automatic processing of data containing measurement errors or non-sawing defects, such as wane, could lead to false signals of an out of control process. Thus, these measurements should be eliminated from the data prior to processing. Defects such as tear-out are sawing defects and thus, should be retained as a feature of the data.

Three methods were proposed that used the LRS data alone for filtering. A simple median method (Method 1) and Lee's sigma method (Method 2) reduced small-scale variation and

tended to smooth over defects. The MSD method (Method 3) retained the small-scale variation and tended to eliminate defects, regardless of their source. If the goal of filtering is to eliminate small-scale noise, such as saw marks, then filtering with a simple median method or Lee's sigma method is recommended; if the goal of filtering is to eliminate defects when preparing data for an SPC system, the MSD method is recommended.

However, the MSD method was not able to distinguish non-sawing defects, such as wane, from sawing defects, such as tear-out; under the MSD method, both types of defects were eliminated by setting values in those areas to missing. Because sawing defects are clues in detecting quality problems, this is unacceptable for most SPC applications. However, in detecting some defects (e.g., wedge or taper), removing tear-out would have little impact on SPC algorithms.

One crucial assumption for all methods was that the precision of the LRS data was constant. Because all measurements were taken within a limited range of approximately 40 cm, (15 ½ inches), this assumption was valid for the lab scan data. In a commercial sawmill, however, scanning would likely take place at the outfeed of the canterline. In addition to the practical considerations involved in placing lasers at a safe distance from the production line, the distance from the LRSs to cants of different sizes will vary. In a typical British Columbia sawmill, cants are produced in a range of sizes, usually from 10 to 30 cm to (4 to 12 inches). If the LRSs are placed at a distance of 40 cm (15 ¾ inches) from a 30 cm cant, the distance to a 10 cm cant will be 50 cm (19 ¾ inches). With the Hermary LRS-50, the difference in resolution is fairly minimal at these distances, changing from ~0.075 mm (0.003 inch) at 40 cm to 0.15 mm (0.006 inch) at 50 cm (Hermary Opto Electronics Inc. 2002). On the other hand, if the LRSs are set-up so that they are 100 cm (39 ½ inches) from the passing 30 cm cants, the difference in resolution is 0.15 mm. This difference is large enough to warrant additional filtering so that data from different sized cants will be comparable when used in size control algorithms.

In order to discriminate between sawing and non-sawing defects, a multi-sensor approach was employed, with a sheet-of-light profile imaging system incorporated into the measurement apparatus. Two methods were presented for classifying defects using a stratified dataset which consisted of 87% defect-free lumber. In order to justify the cost and computing time of such a system, the accuracy rate for defect classification should have been sufficiently higher than the accuracy rate achieved by assuming all lumber was defect-free (87%).

The two methods, a discriminant analysis method and a rule-based algorithm, performed very similarly. The two methods correctly classified about 93% of the lumber in the validation dataset, a 50% increase over assuming all lumber was defect-free. The rule-based method was preferred because the algorithm was developed as a series of logical steps with each rule serving a distinct purpose. The discriminant functions, on the other hand, were based on up to seven independent variables and were not easily interpretable. The rule-based method is recommended for use in removing gross defects from the LRS dataset prior to performing SPC.

The significant findings in the research on LRS data filtering methods were therefore:

1. Method 3 (the MSD method) was better than Methods 1 and 2 (the median method and Lee's sigma method) to filter LRS data for measurement errors;
2. Using a multi-sensor method represented a significant improvement over simple LRS-based filtering, when filtering for non-sawing defects; and
3. A rule-based method was equally good at filtering for non-sawing defects, when compared to a discriminant analysis method; however, the rule-based method provided easier interpretation.

Some improvements are recommended. Increasing the number of pictures per inch would provide better coverage of smaller defects, such as holes. Also, sheet-of-light profiling methods

were developed in isolation from the LRS data; including LRS measurements from the local sawing area in the algorithm may improve the defect detection system. Both of these improvements would require upgrades to the hardware used in the measurement apparatus. The data acquisition component of the system was only able to acquire one frame per 0.8 mm of lumber. Increasing the density of pictures taken would aid in both improvements, as under the current limitations, the density of the LRS data is ten times that of the camera data, making alignment of the two datasets inaccurate.

## 6.2 Statistical Model

The development of a model that adequately described the data from an LRS measurement apparatus was challenging because of its multiple sources of variation and complex correlative structure. Observations within board, side, and laser position were well-described by autoregressive integrated moving average models that accounted for seasonality and long-range memory. These models are not ideal because of their non-stationarity; however, this was an important and interesting finding, allowing for comparisons to simpler models and providing a model form that was used to generate simulated data in subsequent steps in this research.

For SPC, the autocorrelation can be ignored if there is not long-range dependence in the data. Quantification of the autocorrelative structure is only important for performing tests of significance and predicting profile values. For use in "traditional" SPC charts, which are based on averages, the autocorrelation in the model could be ignored with only a negligible change in accuracy. The recommended model for SPC data applications was an uncorrelated errors model; that is, a model ignoring autocorrelation, with parameters for the effects of laser position and sample boards and their interactions, and model parameters estimated for each saw configuration and side.

The significant findings in the research on the statistical model were therefore:

1. In comparison to data collected with periodic sampling, the model for LRS data requires additional components of variance;
2. The model parameters vary by saw;
3. The within-LRS errors are non-independent and autocorrelated; and
4. Autocorrelation is only a concern if tests of significance are to be performed, or individual measurements are to be predicted.

A major limitation in the development of the model was the lab scanned dataset. Of primary concern was that the information about the specific saws that were engaged during the cut and the order that the lumber was cut in were not available. Although the saw configuration was known, the particular saw (e.g., Bandsaw #1 versus Bandsaw #2) was not. Obtaining this information would require substantial effort in interfacing with the programmable logic controllers used to control the saws; however, it would tie data to specific saws, rather than specific saw configurations and increase the usefulness of the model. The order of cutting was important to the model, as it gave the autocorrelative structure between subsequent boards cut. This information was available in the field scan dataset, but autocorrelation between subsequent boards did not prove to be significantly different from zero. The pattern of correlation may be an important factor in identifying quality problems in the mill. This possibility should be further investigated with a sample taken during a time of saw malfunction.

Further, the general lack of defects in both the lab and field datasets was a concern in model development. An out-of-control sample with sawing defects would help to better describe lumber with these quality problems. Also, other kinds of sawing defects, such as washboarding, would be interesting to analyze for seasonality components, as cycle detection via spectral analysis may provide guidance for saw maintenance.

## 6.3 SPC System

Many lumber manufacturers have already implemented real-time scanning technologies in their modern mills; however, the SPC methods that are being applied to the LRS data have not been updated to reflect the sampling frequency or the capacity of the new technology. These systems should be updated so that control limits are correctly constructed, and also to take advantage of the wealth of data available.

The systems that were proposed and tested in this thesis were designed to specifically target five types of sawing defects that are frequent problems in modern sawmills and that may be identified with LRS technology. These defects were setworks problems, wedge, taper, snipe or flare, and snake.

Machine positioning problems and wedge can be identified using summary statistics and traditional Shewhart-type charts. The X-bar individuals chart (X-bar$_{ind}$ chart) using averages by board was recommended for monitoring machine positioning problems, and the range charts for laser position averages ($R_{\lambda_{ind}}$ and $R_{\lambda_{grp}}$ charts) were recommended for wedge. This latter chart was slightly biased due to deviations from normality in the profile data, and thus will require some further research to ensure satisfactory in-control behaviour of the chart.

For trend related defects (taper, snipe/flare), quantile charts for monitoring the estimated slope parameter of a linear regression through the profile data were recommended. While a chart that considered the entire length of the board performed well for monitoring taper ($Q_\tau$ chart), the chart for snipe and flare ($Q_{\tau'}$ chart) was overly sensitive. These charts were based on 3-sigma limits that implicitly assumed normally distributed data; that is, the distributions of the parameters for slope were assumed normal. This assumption may have to be re-assessed with a larger dataset.

Charts for snake were based on measures for roughness and waviness. These parameters have never been measured in the area of lumber size-control, and thus, there is no prior expectation about their distributions. Quantile charts were developed using fitted gamma distributions to account for the long tails seen in the sample data; however, even the recommended $Q_{w_q}$ chart for monitoring RMS waviness was very sensitive to small changes and its performance varied by saw configuration and side. Since the gamma distributions were fit with only a small dataset, a larger dataset may be the key to finding more appropriate control limits.

The significant results from the research on the SPC system were therefore:

1. The recommended control charts for each of the five sawing defects targeted were:
    a. Machine positioning problems: X-bar$_{ind}$ chart;
    b. Wedge: $R_{\lambda_{ind}}$ chart;
    c. Taper: $Q_\tau$ chart with 3-sigma limits;
    d. Snipe/flare: $Q_{\tau'}$ chart with limits derived from a fitted (possibly non-normal) distribution; and
    e. Snake: $Q_{w_q}$ chart with limits derived from a fitted gamma distribution.
2. Under in-control conditions, the recommended control charts signaled at a rate approximately equal to the expected rate (0.27% in most cases); and
3. Under most simulated out of control conditions, the recommended control charts signaled only when the defects they were intended to monitor were present in the data.

There are two main purposes of any SPC system: (1) to provide a signal when defective products are being produced and (2) to identify when processes are performing above expectations (Maness 1993). The SPC system developed in this thesis went far toward serving these

purposes. However, there is room for further research and refinements for these charts, especially for snake and trend-related charts, which were based on quantities not normally included in SPC systems. In particular, these defects are well-known to mill personnel and have definitive consequences at nominal levels. As an extreme example, a defect that increased thickness by 0.25" at any point along a board destined for the planer would create a jam at Weyerhaeuser's New Westminster (British Columbia, Canada) planer mill, as it would be too large to fit under the planer heads.

Given the consumer specifications for the product being produced and machinery limitations at the mill, control limits could be derived as economic tests rather than statistical tests. So-called "economic limits" have been in use since the 1950's in many industries (Keats *et al.* 1997). In particular, control limits for roughness and waviness could be a function of the specifications of the final lumber product, the baseline variability in these quantities, and the expertise of mill personnel.

More generally, judging if a process is "in-control" must be a decision tailored to the real-time process it is applied to. In the Shewhart sense, a process is in control when it is uneconomic to look for assignable causes (Tukey 1946, as cited in Nelson 1999). The implicit assumption in the research relating to this SPC system is that run times associated with 3-sigma limits are long enough for commercial mills. Not only is the occurrence of false alarms costly, but it also leads to a distrust of the SPC system. Using 3-sigma limits, may give an average run length (ARL) that is too short for many mills. Consider the pilot set-up at Canfor's Upper Fraser Mill (Upper Fraser, British Columbia, Canada). Typically, six logs are processed per minute at the quad-bandsaw. Assuming a Type I error rate of 0.27%, the ARL is $1/0.0027 \approx 370$, which roughly translates to one false alarm every hour. However, there are multiple control charts being monitored, each with roughly the same possibility of a false alarm. Assuming that the charts are

independent (which may be a poor assumption), using six charts increases the Type I error rate to 1.6%, which results in an ARL of 62, or just over ten minutes of processing.

On the other hand, if the process is truly out of control, then out of control behaviour of the charts should be sustained. Mill personnel should consider whether the process is totally out of control and should be stopped, or that the problem is sporadic and should be investigated further. The charts work very well if put in this context. This illustrates the need for mill staff to carefully consider the derivation of their control limits. All the charts that have been introduced in this research could be adjusted to target any ARL. The choice of an ARL will depend on balancing the costs of unnecessary shut down and the costs of producing faulty product.

## 6.4 Summary of Future Research Needs

Several enhancements can be made using this research as a base. The following is a list of suggestions for future research:

1. Enhance filtering algorithm to adjust for resolution of lasers at varying distances from target lumber;
2. Improve data acquisition system for digital camera data, enabling a higher density of pictures and the integration of simultaneous LRS and camera data;
3. Obtain and analyze an additional dataset, with information about the particular saw used during the cut, the order of cutting, and out of control samples representing the five sawing defects discussed. This dataset will be used to:
    a. Refine model parameters for particular saws (versus saw configurations);
    b. Find a stationary model that fits the autocorrelated data;
    c. Further study seasonality components in the model with various sawing defects;

d. Study the distribution of ranges for the $R_{l_{ind}}$ and $R_{l_{grp}}$ charts;

e. Better quantify the distribution of the fitted slopes monitored for flare and snipe; and

f. Study the distribution of waviness and roughness measures.

4. Investigate economically based control limits for a particular product, based on a particular mill's production costs and machine limitations.

This thesis provides a basis for operational real-time data acquisition and SPC methods for LRS data in automated lumber mills. Using the methods developed in this thesis will improve the ability of lumber mills to track sawing problems and improve the quality of the wood products they produce. Future research can enhance the research already completed, and bring the proposed real-time SPC methods closer to implementation in commercial sawmills.

## 6.5 References


Anonymous. 2004. The State of Canada's Forests: 2002-2003. Natural Resources Canada, Ottawa, Ontario.

Hermary Opto Electronics Inc. 2002. What is scanner technology [online]. Available from http://www.hermaryopto.com/scantech.html [cited January 14, 2004].

Keats, J. B., E. Del Castillo, E. Von Collani and E. M. Saniga. 1997. Economic modeling for statistical process control. Journal of Quality Technology 29(2): 144-147.

Maness, T. C. 1993. Real-time quality control system for automated lumbermills. Forest Products Journal 43(7/8): 17-22.

Nelson, L. S. 1999. Notes on the Shewhart control chart. Journal of Quality Technology 31(1): 124-126.

Random Lengths. 2004. Daily WoodWire > Curtailment Watch > Lumber Mills-Canada > November 2003 [online]. Available from http://www.randomlengths.com/base.asp?s1=Daily_WoodWire&s2=Curtailment_Watch&s3=Lumber_Mills-Canada&pub=list [cited July 28, 2004].

# Appendix I Within Sample Board Model

To properly account for the autocorrelation of the error terms ($\varepsilon_{ijklm}$) in model [4-1], time series models were investigated. Four models were fit, tested, and evaluated by board, sample, and laser position.

## *AI.1 Candidate Models*

To facilitate notation, consider a single series of laser range sensor (LRS) measurements from a single board, side, and laser position ($\varepsilon_m$).

### AI.1.1 Autoregressive Moving Integrated Average (ARIMA) Models

An autoregressive moving average (ARMA) model with $p$ autoregressive (AR) parameters and $q$ moving average (MA) parameters is of the form:

$$\phi(B)\varepsilon_m = \alpha + \theta(B)\upsilon_m \qquad \text{[AI-1]}$$

where: $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$;

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$;

$B$ is the backshift operator, e.g., $(1 - \phi_1 B - \phi_2 B^2)\,\varepsilon_m = \varepsilon_m - \phi_1\,\varepsilon_{m-1} - \phi_2\,\varepsilon_{m-2}$;

$\phi_1, \phi_2, \ldots, \phi_p$ are the AR parameters;

$\theta_1, \theta_2, \ldots, \theta_q$ are the MA parameters;

$\alpha$ is the intercept; and

$\upsilon_m$ is a white noise error process, with $\upsilon_m \sim N(0,\sigma_\upsilon^2)$.

For example, an ARMA(1,1) model is of the form:

$$\varepsilon_m - \phi\,\varepsilon_{m-1} = \upsilon_m - \theta\,\upsilon_{m-1} \qquad \text{[AI-2]}$$

Some time series, especially those found in economics, are not adequately represented by ARMA models, but instead are non-stationary. Stationarity can be defined as a quality of a process in which the statistical parameters (mean and standard deviation) do not change with time (Box and Luceño 1997). A non-stationary process tends to drift away from its initial state. For example, if a sawing machine was left to run without maintenance, its saw blades would dull and over time would no longer function properly. Initial plots of the LRS data showed non-stationary characteristics, such as cycles and shifts in the mean.

Stationary models are desirable in that they ensure that early values of $\upsilon_m$ do not have influence over later values of $\varepsilon_m$ (Brocklebank and Dickey 2003). Stationarity may be achieved by considering differenced data, as is modeled in autoregressive integrated moving average (ARIMA) models. An ARIMA($p, \delta, q$), where $\delta$ is the degree of differencing, is of the form (Box and Jenkins 1970):

$$\phi(B)\nabla^{\delta}\varepsilon_m = \alpha + \theta(B)\upsilon_m \qquad \text{[AI-3]}$$

where: $\nabla$ is the difference operator, e.g., $\nabla^{\delta} = (1-B)^{\delta}$.

Including an intercept in this model corresponds to having a deterministic trend in the data series. This may be appropriate for a board that increases or decreases in profile along its length.

### AI.1.2 Seasonal ARIMA (SARIMA) Models

Although first differences are common, especially in SPC applications (e.g., Vander Wiel 1996; Noffsinger and Anderson 2002), lumber surface profile data may contain other patterns that are not unlike the seasonal differences found in monthly econometric data. Monthly econometric data are often de-trended with the twelfth difference, or the twelfth difference of the first-order differences (Brocklebank and Dickey 2003):

$$(1 - B^{12})\,\varepsilon_m = \upsilon_m - \theta\,\upsilon_{m\text{-}12}$$

$$\varepsilon_m - \varepsilon_{m\text{-}12} = \upsilon_m - \theta \upsilon_{m\text{-}12} \qquad \text{[AI-4]}$$

And,

$$(1 - B^{12})(1 - B)\,\varepsilon_m = (1 - \theta_1 B)(1 - \theta_2 B^{12})\,\upsilon_m$$

$$\varepsilon_m - \varepsilon_{m\text{-}1} - (\varepsilon_{m\text{-}12} - \varepsilon_{m\text{-}13}) = \upsilon_m - \theta_1 \upsilon_{m\text{-}1} - (\theta_2 \upsilon_{m\text{-}12} - \theta_1 \theta_2 \upsilon_{m\text{-}13}) \qquad \text{[AI-5]}$$

Equation [AI-4] is a seasonal ARIMA(0,1,1) model with seasonal period $s = 12$, or SARIMA$(0,1,1)_{12}$. Equation [AI-5] is a SARIMA(0,1,1)x$(0,1,1)_{12}$. The general form of the SARIMA$(p,\delta,q)$x$(P,\delta',Q)_s$ is (Ray 1993):

$$\phi(B)\Phi(B^s)(1-B)^{\delta}(1-B^s)^{\delta'}\varepsilon_m = \theta(B)\Theta(B^s)\upsilon_m \qquad \text{[AI-6]}$$

where: $\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}\,\Phi$;

$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \ldots - \Theta_Q B^{Qs}$;

$\Phi_1, \Phi_2, \ldots, \Phi_P$ are the seasonal AR parameters;

$\Theta_1, \Theta_2, \ldots, \Theta_Q$ are the seasonal MA parameters; and

$\delta'$ is the degree of seasonal differencing.

### AI.1.3 Autoregressive Fractionally Integrated Moving Average (ARFIMA) Models

Most time series theory assumes that observations separated by long time spans are independent (or nearly so). However, there are numerous empirical examples (e.g., from economics, meteorology, and hydrology) where distant observations exhibit non-negligible amounts of dependence (Beran 1992). These so-called "long memory models" are characterized by a slowly decaying autocorrelation function, and may have features of non-stationary time series, such as cycles and changes of levels (Hosking 1984). Autoregressive fractionally integrated moving average (ARFIMA) models describe this behaviour by permitting fractional differences (i.e., $\delta$ is

non-integer). Using a binomial expansion, the differenced term in [AI-3] becomes (Granger and Joyeux 1980; Hosking 1984):

$$\begin{aligned}\nabla^{\delta}\varepsilon_m &= (1-B)^{\delta}\varepsilon_m = \sum_{a=0}^{\infty}\binom{\delta}{a}(-B)^{\delta}\varepsilon_m \\ &= \varepsilon_m - \delta\varepsilon_{m-1} - \frac{1}{2}\delta(1-\delta)\varepsilon_{m-2} - \frac{1}{6}\delta(1-\delta)(2-\delta)\varepsilon_{m-3} - \ldots\end{aligned} \qquad \text{[AI-7]}$$

When $|\delta| < ½$, the infinite sum in [AI-7] converges in mean square (Hosking 1984), and therefore the fractional difference form of [AI-3] is stationary.

ARFIMA models have also been extended to include a fractional seasonal component (Carlin and Dempster 1989; Porter-Hudak 1990) by permitting both parameters $\delta$ and $\delta'$ from the SARIMA($p, \delta, q$)x($P, \delta', Q$)$_s$ to take fractional values. A seasonal fractionally differenced model (SARFIMA) is appropriate for data that exhibit both seasonal and non-seasonal short term dependence, as well as slowly decaying autocorrelation at periodic lags (Ray 1993).

## *AI.2 Model Selection*

Models were selected using LRS data from 110 sample boards[41]. First, models were tested for stationarity. Then, four model forms were investigated:

1. ARIMA models using first differences, i.e., ARIMA($p$,1,$q$);
2. SARIMA models using first differences and a seasonal component corresponding to cycles found in the data, i.e., SARIMA($p$,1,$q$)x($P$,1,$Q$)$_s$;
3. Long-memory models, or fractional values of $\delta$, i.e., ARFIMA($p, \delta, q$); and
4. Long-memory models with a seasonal component, i.e., SARFIMA($p, \delta, q$)x($P, \delta', Q$)$_s$.

[41] For data collection methods, see Sections 4.3 and 4.4 (Chapter 4).

The models were fit by saw configuration, side, board, and laser position ($110 \times 2 \times 2 = 440$ series), allowing the model orders ($p$, $q$, $P$, $Q$, $\delta$, $\delta^*$, and $s$) and model parameters ($\phi$, $\theta$, $\Phi$, and $\Theta$) to vary by saw configuration, side, board, and laser position. To facilitate notation, let $f=ijkl$, so that $f$ denotes the observations from the $i$th sawing configuration, $j$th side, $k$th sample board, and $l$th laser position. Thus, orders $p_f$, $q_f$, $P_f$, $Q_f$, $\delta_f$, $\delta^*_f$, and $s_f$ were selected, and the parameters $\phi_f$, $\theta_f$, $\Phi_f$, and $\Theta_f$ were estimated for each $f$th series.

### AI.2.1 Testing for Stationarity

Stationarity was visually evaluated by examining plots of raw data, as well as plots of the autocorrelation function (ACF). Non-stationary time series have slowly decaying ACFs, and raw plots often show cycles and a wandering mean.

The stationarity of each $f$th series was also evaluated formally with the Dickey-Fuller test (Dickey and Fuller 1979), which essentially tests the hypothesis that the data from the $f$th series form an AR($p_f$) process with $\phi_{t_f}=1$ for some $t=1, \ldots, p_f$. The Dickey-Fuller test was performed in the SAS procedure PROC ARIMA[42], assuming an AR(10) process with a non-zero mean and significant trend. Since the model form was unknown, but there was graphical evidence suggesting very slow decay, the large lag number ($p_f=10$) was used. The assumption of non-zero mean with trend was selected, as trends by series were unknown, and this results in a more conservative test (Brocklebank and Dickey 2003).

### AI.2.2 ARIMA($p$,1,$q$) Parameters

The autoregressive and moving average orders were selected based in part on the Box-Jenkins strategy for time series modeling (Box and Jenkins 1970). Under this strategy, models are identified using the ACF and partial autocorrelation function (PACF). For instance, the ACF of

[42] All SAS procedures referred to in this chapter were run with Version 8.2 (SAS Institute 2002).

an MA($q$) process cuts off after $q$ lags, whereas that of an AR process tails off. On the other hand, the PACF of an AR($p$) process cuts off after $p$ lags, whereas that of an MA process tails off. Estimates of the ACF and PACF were generated using the SAS procedure PROC ARIMA.

Using a range of initial values of $p_f$ and $q_f$, candidate models were also evaluated using the Minimum Information Criteria (MINIC) method in the SAS procedure PROC ARIMA. Given that the candidate models are stationary, the MINIC method selected the values of $p_f$ and $q_f$ that produce the best (lowest) value of Bayesian Information Criteria (BIC) for the series.

### AI.2.3 SARIMA($p$,1,$q$)x($P$,$\delta$,$Q$)$_s$ Parameters

Seasonality components were found by examining the spectrum of the first-differenced series. Estimates of the spectrum for each $f$th series were found via the SAS procedure PROC SPECTRA, and periodicities were found by identifying peaks in plots of the smoothed spectral density. Seasonal periods were restricted to be the same for the top and bottom position of each board × side combination.

After identifying an appropriate period for each $f$th series ($s_f$), the orders $\delta'_f$, $p_f$, $q_f$, $P_f$, and $Q_f$ were selected. Although the parameter $\delta'_f$ could take on any integer value, examples of this are highly uncommon. Furthermore, initial examination of ACF's with parameters of $\delta'_f$ =2 and 3 showed erratic behaviour. For these reasons, and in the interest of parsimony, $\delta'_f$ was assumed to be 1. While the MINIC method could be used to search for the optimal AR and MA orders, this method was extremely impractical, as lags of 20 or more could have to be considered (e.g., for $s_f$ >20). Instead, only values of 0 and 1 were considered for AR and MA orders, resulting in sixteen possible combinations of $p_f$, $q_f$, $P_f$, and $Q_f$. Akaike's Information Criteria (AIC) was computed for each $f$th series, and the model with the lowest AIC was considered the best fit.

### AI.2.4 ARFIMA(*p,d,q*) Parameters

The ARFIMA models were estimated in two stages. First, the order of fractional differencing, $\delta_f$, was estimated. Then, the series $\varepsilon'_{fm} = \nabla^{\delta_f} \varepsilon_{fm}$ was used as the dependent variable in PROC ARIMA to find estimates of the AR and MA parameters. Optimal values of $p_f$ and $q_f$, were again found using the MINIC method by saw type, board, side, and laser position.

Beran (1994) identified more than five ways to estimate the fractional differencing order, using estimates of the autocorrelation and partial autocorrelation functions, variogram, variance plots, and spectral density. The latter, due to Geweke and Porter-Hudak (1983), uses the shape of the spectral density near zero, and is widely recommended (Andel 1986; Beran 1994). Under this method, the periodogram of the ARFIMA process, $I(\omega)$, is written in terms of the spectral densities of the ARFIMA and ARMA processes. For example, consider a single series of measurements from a single board, side, and laser position. The log of the periodogram can be written as:

$$\ln I(\omega_m) = \ln \frac{\sigma_\upsilon^2 f_u(0)}{2\pi} - \delta \ln\left(4\sin^2 \frac{\omega_m}{2}\right) + \ln \frac{f_u(\omega_m)}{f_u(0)} + \ln \frac{I(\omega_m)}{g(\omega_m)} \qquad \text{[AI-8]}$$

where: $\omega_m$ is the frequency: $\omega_m = m\pi/n$, $m$=0, 1, ..., $n$;

$n$ is the number of observations in the series;

$f_u(\omega_m)$ is the spectral density of an ARMA($p,q$) process;

$g(\omega_m)$ is the spectral density of an ARFIMA process: $g(\omega_m) = f(\omega_m) f_u(\omega_m)$; and

$$f(\omega_m) = \frac{\sigma_\upsilon^2}{2\pi}\left(4\sin^2 \frac{\omega_m}{2}\right)^{-\delta}.$$

Because only the frequencies ($\omega_m$) near zero are considered, $\ln(f_u(\omega_m)/f_u(0))$ is negligible. It can also be shown that $\ln(I(\omega_m)/g(\omega_m)) \rightarrow 0.57721\ldots$, the Euler constant. Thus, $\delta$ can be estimated straightforwardly in an ordinary least squares (OLS) regression:

$$Y_m = c_0 + c_1 X_{1_m} \qquad \text{[AI-9]}$$

where: $Y_m = \ln(I(\omega_m))$; and

$$X_{1_m} = \ln(4\sin^2(\omega_m/2)).$$

Then, the estimate of $\delta$ is $\hat{\delta} = -c_1$. Butler (1999) notes that this procedure is valid only when the input series (in this case, $\varepsilon_m$) is stationary. If the stationary series $(1-B)\varepsilon_m$ is instead used as the input series in finding $\delta$, the estimate of $\delta$ is $\hat{\delta} = 1 - c_1$.

For each $f$th series, estimates of the smoothed periodogram of $(1-B)\varepsilon_{fm}$ were found using the SAS procedure PROC SPECTRA. Regression parameters were estimated using PROC REG.

### AI.2.5 SARFIMA($p,\delta,q$)x($P,\delta',Q$)$_s$ Parameters

As in the ARFIMA case, order selections and parameter estimates for the SARFIMA models were found in two stages. First, $\delta_f$ and $\delta'_f$ were estimated; then, the ARMA orders were selected using PROC ARIMA with $\varepsilon''_{fm} = (1-B)^{\delta_f}(1-B^{s_f})^{\delta'_f}\varepsilon_{fm}$ as the dependent variable.

In a method similar to Geweke and Porter-Hudak (1983), Andel (1986) extended the spectral density estimation method to seasonal ARFIMA processes. Using the same method as in AI.2.4, it can be shown that $\delta'$ can be estimated by OLS with $Y_m = \ln(I(\omega_m))$ and $X_m = \ln(4\sin^2(s\omega_m/2))$. Thus, PROC SPECTRA and PROC REG could be similarly used to estimate $\delta'$, and $\delta$ and $\delta'$ would be estimated separately in a two-stage process. Estimates obtained in this manner will be consistent (Porter-Hudak 1990); however, a different approach is suggested. Following Andel

(1986) in the derivation of [AI-8], the log periodogram of a single SARFIMA series can then be expressed as:

$$\ln I(\omega_m) = \ln \frac{\sigma_\upsilon^2 f_u(0)}{2\pi} - \delta \ln\left(4\sin^2 \frac{\omega_m}{2}\right) - \delta' \ln\left(4\sin^2 \frac{s\omega_m}{2}\right) + \ln \frac{f_u(\omega_m)}{f_u(0)} + \ln \frac{I(\omega_m)}{g(\omega_m)} \quad \text{[AI-10]}$$

Thus, $\delta$ and $\delta'$ can be simultaneously estimated in a multiple regression with $X_1$ and $X_{2_m}$=ln(4sin$^2$($s\omega_m$/2)). For each $f$th series, $\delta_f$ and $\delta'_f$ were then estimated via the SAS procedure PROC REG.

As in the SARIMA model, it would be extremely impractical to apply the MINIC method to find the AR and MA orders in the SARFIMA model. Assuming parsimony for the SARFIMA model form, values of 0 and 1 only were considered for the AR and MA orders. Sixteen possible combinations of $p_f$, $q_f$, $P_f$, and $Q_f$ were compared for each series, and the model with the lowest AIC was considered the best fit.

### AI.2.6 Model Evaluation and Comparison

Order selection and parameter estimation was performed separately for each model by saw configuration, side, board, and laser position ($f$=$ijkl$) series. For each of the four models (ARIMA, SARIMA, AFRIMA, and SARFIMA), the final model parameters, degrees of differencing, and seasonal periods were chosen by series, while AR and MA orders were restricted to be the same for all board × side × laser position combinations under each saw type (Bandsaw, Chipper-head, and Circular Saw). Thus, while parameter estimates for each series in each saw type were different, the AR and MA orders were the same by saw type. This simplification kept the number of model forms at a manageable size.

Models were compared with the AIC and multivariate R-square. Because a comparison of the fit statistics for each board, side, and laser would be cumbersome, the overall AIC by saw type was

calculated as the sum of all AIC's for the saw type and the overall R-square was calculated as the average of all individual R-squares, weighted by the total sums of squares.

Lack of fit is indicated by patterns in the residual plots and ACF plots, or values of the ACF that are significantly non-zero. Tests for lack of fit are available, e.g., the Portmanteau test for autocorrelation of the residuals (Hosking 1980). However, the large sample sizes in this analysis make these tests very sensitive, and the white noise hypothesis (no autocorrelation) for the residuals would be rejected over an extremely small range. Thus, lack of fit was evaluated visually using plots of the model residuals and the ACF of the residuals.

## *AI.3 Results*

### AI.3.1 Tests of Stationarity

Plots of the raw data and the first differenced data series are shown for the bottom laser position for a single sample from each of the three saw types (Bandsaw, Chipper-head, and Circular Saw) in Figure AI-1 to Figure AI-3. For all three types of saws, the raw data series show signs of non-stationary behaviour, such as wandering means and irregular cycles. The first differenced data are considerably more stable.

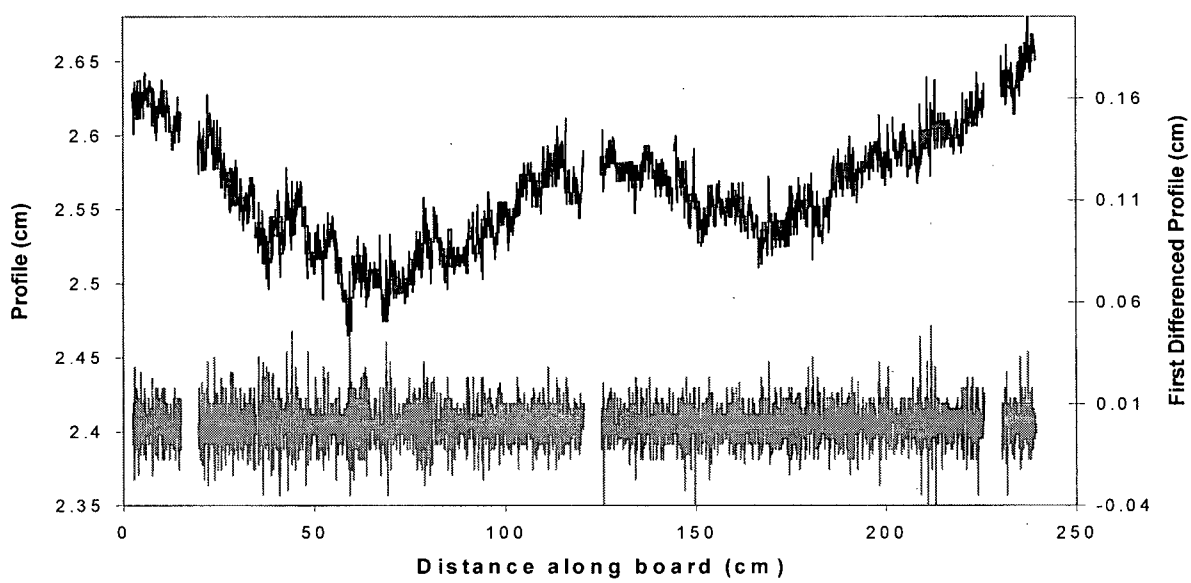


**Figure AI-1. Plot of original and first differenced series for Board 001-Side 1-Laser 1 (Bandsaw).**

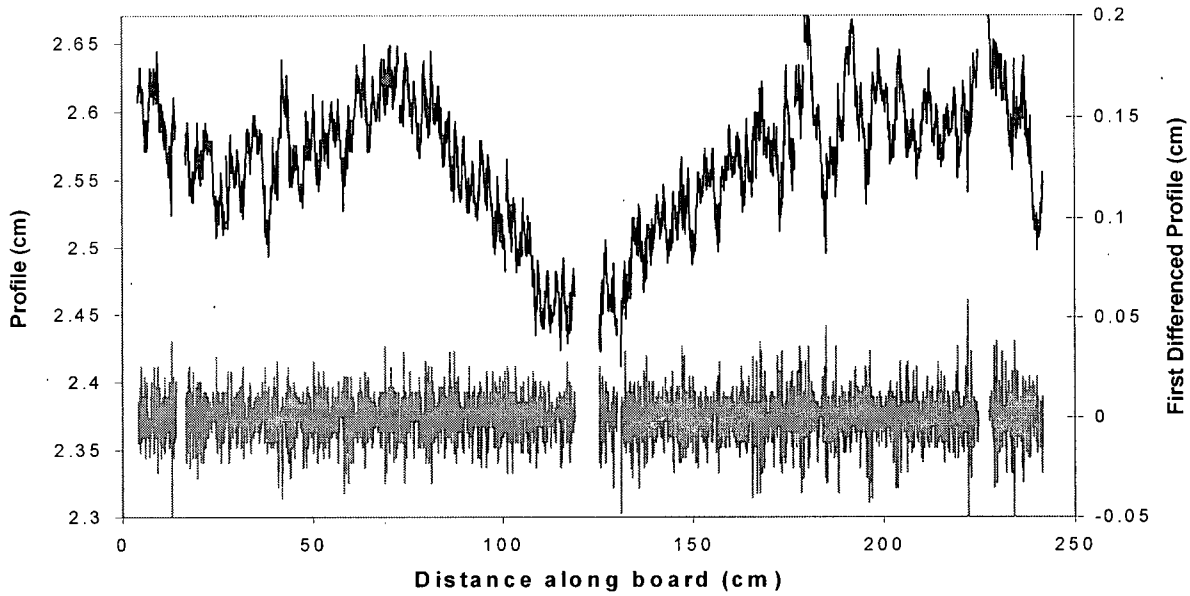



**Figure AI-2. Plot of original and first differenced series for Board 002-Side 1-Laser 1 (Chipper-head).**

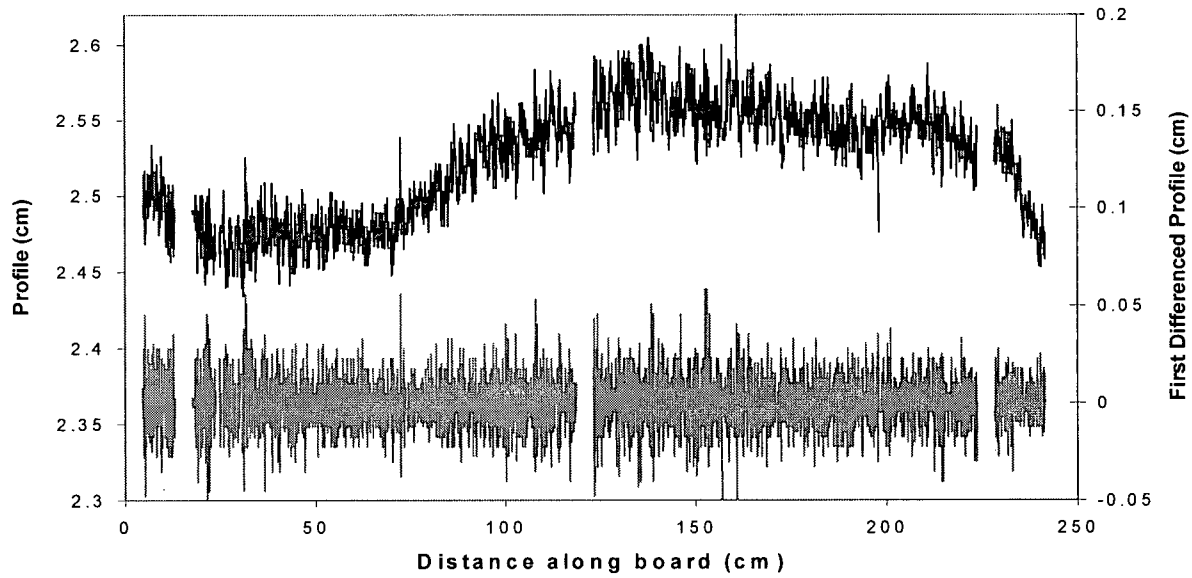



**Figure AI-3. Plot of original and first differenced series for Board 012-Side 1-Laser 1 (Circular Saw).**

Results of the Dickey-Fuller test by saw type and data series are shown in Table AI-1. Since more than half of these tests indicated that the series were significantly non-stationary ($\alpha = 0.05$), all series in the analysis were treated as non-stationary. This result is also supported by plots of the ACF and PACF of the raw data, shown in Figure 4-7 (Chapter 4).

**Table AI-1. Summary of non-stationary series, per the Dickey-Fuller unit root test.**

| Saw Type | Total Number of Series | Non-Stationary Series | | | | | |
|---|---|---|---|---|---|---|---|
| | | Side 1 | | Side 2 | | Total | |
| | | Number | % | Number | % | Number | % |
| Bandsaw | 110 | 7 | 75% | 82 | 73% | 89 | 63% |
| Chipper-head | 48 | 14 | 29% | 12 | 25% | 26 | 27% |
| Circular Saw | 42 | 21 | 50% | 19 | 43% | 40 | 48% |
| Total | 220 | 117 | 53% | 115 | 52% | 232 | 52% |

## AI.3.2 ARIMA(*p*,1,*q*) Model

Plots of the ACF and PACF for the first-differenced data from Figures AI-1 – AI-3[43] are shown in Figure AI-4 and Figure AI-5, respectively. In all three saw types, the ACF cut off sharply after lag one and the PACF tailed off after lag one, indicating that the data followed an MA(1) process. On the other hand, the ACF was significantly non-zero at some lags ($\alpha = 0.05$), indicating that a higher order model may be necessary.

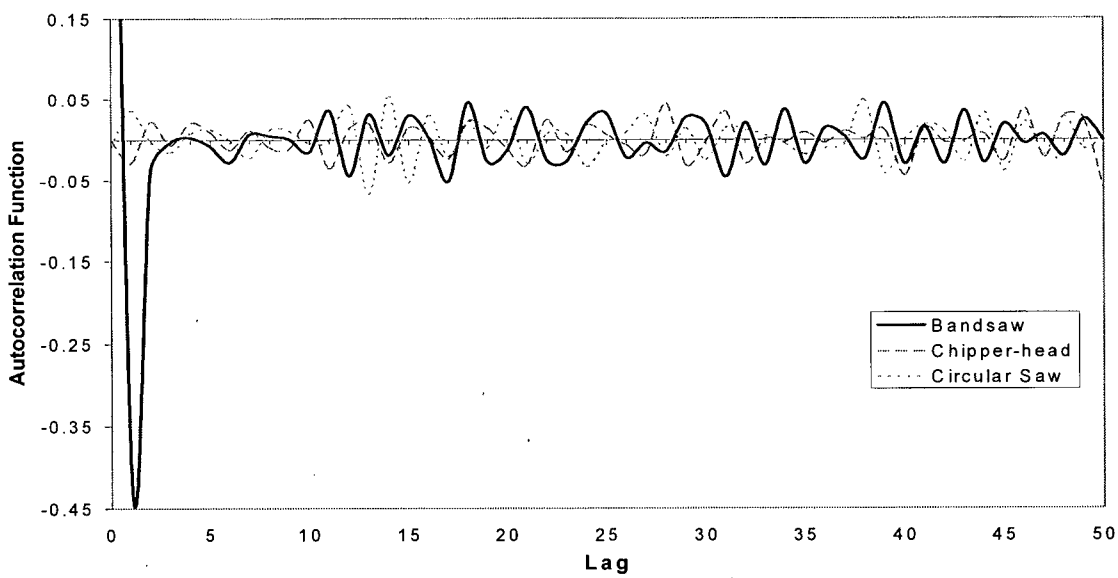



**Figure AI-4. ACF of first-differenced bottom laser data for three saw type samples.**

Results from the MINIC method are shown by saw type in Table AI-2. For bandsawn lumber, the method most often selected was an ARMA(1,1) model. For the chipped and circular-sawn lumber, ARMA(0,1) was most often selected.

[43] For these and subsequent graphs by saw type, the samples plotted are: Board 001-Side 1-Laser 1 (Bandsaw), Board 002-Side 1-Laser 1 (Chipper-head), Board 012-Side 1-Laser 1 (Circular Saw).

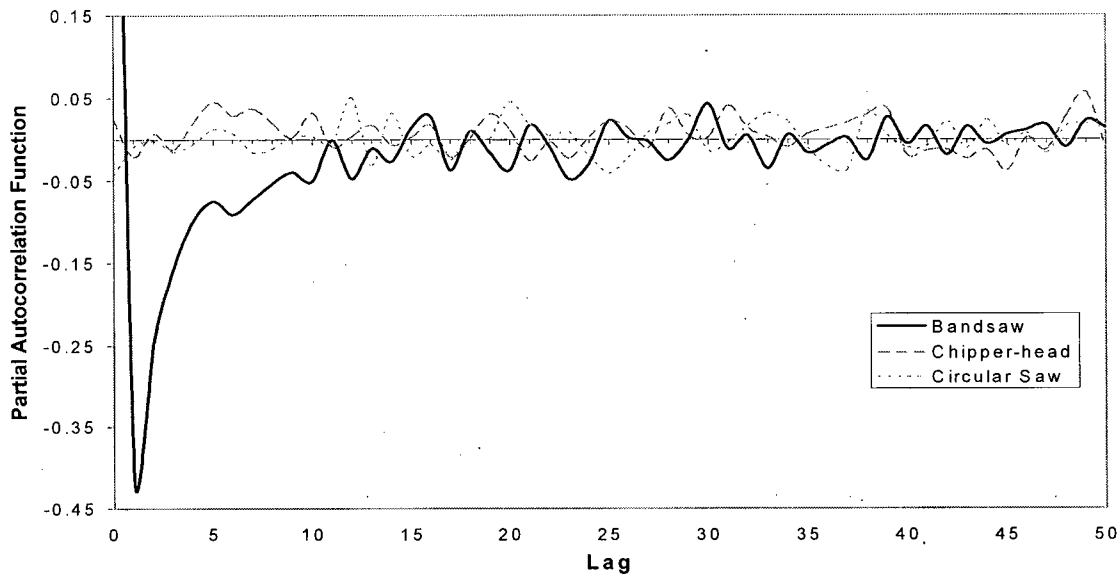



**Figure AI-5. PACF of first-differenced bottom laser data for three saw type samples.**

**Table AI-2. Summary of most often selected orders of *p* and *q* by series for ARIMA(*p*,1,*q*) [AI-3] model.**

| Bandsaw | | | Chipper-head | | | Circular Saw | | |
|---|---|---|---|---|---|---|---|---|
| $p_f$ | $q_f$ | % Samples | $p_f$ | $q_f$ | % Samples | $p_f$ | $q_f$ | % Samples |
| 1 | 1 | 49% | 0 | 1 | 43% | 0 | 1 | 38% |
| 0 | 1 | 24% | 1 | 1 | 14% | 1 | 1 | 32% |
| 2 | 2 | 4% | 1 | 2 | 13% | 1 | 2 | 6% |

The selected models forms were therefore:

Bandsaw (ARIMA(1,1,1)): $$(1-\phi_{1f}B)(\varepsilon_{f,m}-\varepsilon_{f,m-1}) = \alpha_f + (1-\theta_{1f}B)\upsilon_{f,m}$$

Chipper-head / Circular Saw (ARIMA(0,1,1)): $$\varepsilon_{f,m}-\varepsilon_{f,m-1} = \alpha_f + (1-\theta_{1f}B)\upsilon_{f,m}$$

Descriptive statistics for the model parameters are shown in Table AI-3. The percent significant is the proportion of estimates that were significantly non-zero using $\alpha = 0.05$. The estimate of the parameter $\alpha_f$ was significantly non-zero for only a few series, indicating few sample boards had significantly non-zero linear trend. The estimated values of $\theta_{1f}$ were positive in all cases, indicating a strong mixing process. Estimates of $\phi_{1f}$, on the other hand were slightly more varied; for three series, the estimate of $\phi_{1f}$ was negative, but only one of these was significantly so.

**Table AI-3. Summary of estimated ARIMA($p$,1,$q$) model [AI-3] parameters by saw type (mm) ($\alpha$ = 0.05).**

| | Bandsaw | | | | Chipper-head | | | Circular Saw | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{\alpha}_f$ | $\hat{\theta}_{1f}$ | $\hat{\phi}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ | $\hat{\alpha}_f$ | $\hat{\theta}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ | $\hat{\alpha}_f$ | $\hat{\theta}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ |
| Average | 0.0000 | 0.0178 | 0.0053 | 0.0967 | 0.0000 | 0.0103 | 0.1087 | 0.0001 | 0.0140 | 0.1009 |
| Minimum | -0.0032 | 0.0092 | -0.0023 | 0.0619 | -0.0038 | 0.0050 | 0.0766 | -0.0013 | 0.0064 | 0.0688 |
| Maximum | 0.0031 | 0.0226 | 0.0149 | 0.1385 | 0.0035 | 0.0156 | 0.1616 | 0.0018 | 0.0191 | 0.1366 |
| % Significant | 5% | 100% | 94% | | 2% | 100% | | 0% | 100% | 5% |

### AI.3.3 SARIMA($p$,1,$q$)x($P$,1,$Q$)$_s$ Model

Estimates of the smoothed spectral density are shown for samples from each saw type in Figure AI-6. A strong periodicity is present in the chipped board, as evidenced by the pronounced peak at 0.22. This corresponded to a cycle of $2\pi/0.22 \approx 24$ observations (2 cm or 0.8 inch). A peak is also evident in the circular-sawn series at 0.57, indicating a cycle of 11 (9.4 cm or 0.37 inch). Although peaks were present in the bandsawn data, they were too frequent to indicate a consistent cyclical pattern.

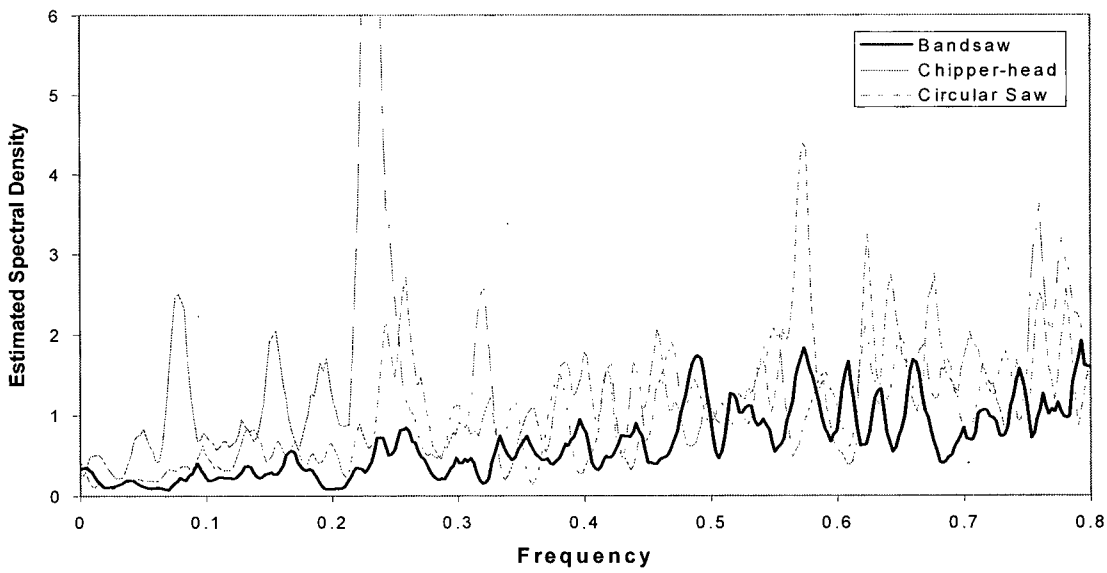



**Figure AI-6. Estimated smoothed spectral density for three saw type samples.**

Spectral density plots were examined for all 440 series. While the bandsawn data were consistently non-cyclical over all samples, the data from the chipped lumber were consistently cyclical. However, the length of cycle for the chipped lumber was not consistent over all sample boards. Most of the chipped lumber samples had a period of 22-24 observations, corresponding

to a distance of 2 cm (0.80 inch) along the board. There were cycles in the circular-sawn series only for boards with pronounced saw marks, and these cycles varied in length by board.

Table AI-4 summarizes the cycle lengths detected in the chipped and circular-sawn series for each board × side combination. Where different cycles were indicated for the top and bottom laser positions, the longer cycle length was selected. For instance, for sample 012, the top series indicated a stronger cycle of 24, whereas the bottom (shown in Figure AI-6) shows a cycle length of 11. Thus, the cycle length used was 24.

**Table AI-4. Cycles ($s_f$) found in data series for chipped and circular-sawn lumber.**

| | Chipper-head Samples | | Circular Saw Samples | |
|---|---|---|---|---|
| Cycle | Number | % | # samples | % |
| 44-50 | 10 | 21% | 4 | 10% |
| 31-34 | 5 | 10% | 0 | 0% |
| 25-27 | 6 | 13% | 0 | 0% |
| 22-24 | 27 | 56% | 15 | 36% |

Using the cycles from Table AI-4, SARIMA($p$,1,$q$)x($P$,1,$Q$)$_s$ models were fit to data series that exhibited cyclical patterns. This included all 48 Chipper-head series (counting each side as a separate series) and 19 of the 42 Circular-sawn series. Figure AI-7 shows the ACF for chipped and circular-sawn samples for data that was both first-differenced and seasonally differenced, i.e., the ACF of $(1-B^{s_f})(1-B)\varepsilon_{fm}$. The ACF's have large peaks around lag=$s_f$, indicating that the SARIMA models may include seasonal AR and MA parameters.

Table AI-5 shows the distribution of the best fitting (as defined by the lowest AIC) model forms for chipped and circular-sawn boards. For chipped boards, the model that best fit in most cases was a SARIMA(1,1,1)x(1,1,1)$_s$; for circular-sawn boards, a SARIMA(1,0,1)x(1,1,1)$_s$ fit better.

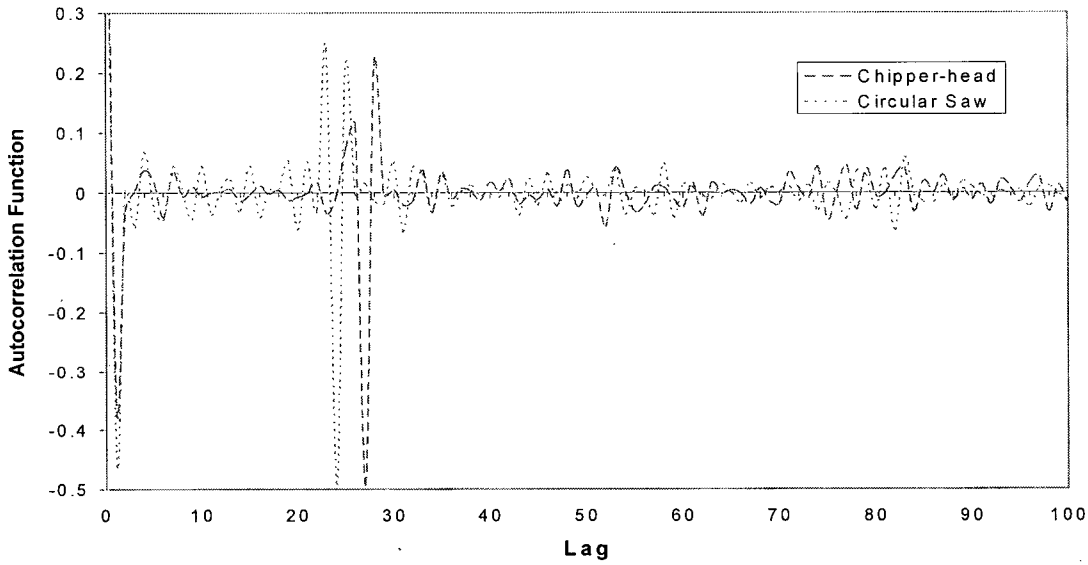



**Figure AI-7. ACF of first-differenced, seasonally-differenced data for two saw type samples.**

**Table AI-5. Summary of lowest AIC values by *p, q, P,* and *Q* by series for SARIMA(*p*,1,*q*)x(*P*,1,*Q*)$_s$ models [AI-6].**

| Chipper-head | | | | | Circular Saw | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $p_f$ | $q_f$ | $P_f$ | $Q_f$ | % Samples | $p_f$ | $q_f$ | $P_f$ | $Q_f$ | % Samples |
| 1 | 1 | 1 | 1 | 33% | 1 | 0 | 1 | 1 | 63% |
| 1 | 0 | 1 | 1 | 27% | 1 | 1 | 1 | 1 | 32% |
| 0 | 1 | 1 | 1 | 25% | 0 | 1 | 1 | 1 | 5% |
| 0 | 0 | 1 | 1 | 15% | 0 | 0 | 1 | 1 | 0% |

The selected model forms for the chipped lumber and circular-sawn lumber were therefore:

Chipper-head (SARIMA(1,1,1)x(1,1,1)$_s$):

$$(1-\phi_{1f}B)(1-\Phi_{1f}B^{s_f})(\varepsilon_{f,m}-\varepsilon_{f,m-1}) = \alpha_f + (1-\theta_{1f}B)(1-\Theta_{1f}B^{s_f})\upsilon_{f,m}$$

Circular Saw (SARIMA(1,0,1)x(1,1,1)$_s$):

$$(1-\phi_{1f}B)(1-\Phi_{1f}B^{s_f})(\varepsilon_{f,m}-\varepsilon_{f,m-1}) = \alpha_f + (1-B)(1-\Theta_{1f}B^{s_f})\upsilon_{f,m}$$

Descriptive statistics for the model parameters are shown in Table AI-6. The estimated value of $\alpha_f$ was significantly different from zero only for a small number of series, whereas the estimated values of $\theta_{1_f}$ and $\Theta_{1_f}$ were significantly non-zero for almost all series. The estimated values of $\theta_{1_f}$

were positive in all cases, indicating a strong mixing process. As in the ARIMA model, estimates of $\phi_{1_f}$ were more varied.

**Table AI-6. Summary of estimated SARIMA($p$,1,$q$)x($P$,1,$Q$)$_s$ model [AI-6] parameters by saw type (mm) ($\alpha$ = 0.05).**

| | Chipper-head | | | | | | Circular Saw | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{\alpha}_f$ | $\hat{\phi}_{1f}$ | $\hat{\Phi}_{1f}$ | $\hat{\theta}_{1f}$ | $\hat{\Theta}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ | $\hat{\alpha}_f$ | $\hat{\phi}_{1f}$ | $\hat{\Phi}_{1f}$ | $\hat{\Theta}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ |
| Average | 0.0000 | 0.0009 | -0.0006 | 0.0115 | 0.0208 | 0.1179 | 0.0000 | 0.0043 | 0.0170 | 0.0205 | 0.1125 |
| Minimum | -0.0005 | -0.0050 | -0.0044 | 0.0022 | 0.0179 | 0.0789 | -0.0005 | -0.0020 | 0.0075 | 0.0186 | 0.0737 |
| Maximum | 0.0009 | 0.0202 | 0.0031 | 0.0247 | 0.0230 | 0.1835 | 0.0007 | 0.0125 | 0.0225 | 0.0232 | 0.1496 |
| % Significant | 3% | 34% | 47% | 95% | 100% | | 3% | 79% | 100% | 100% | |

## AI.3.4 ARFIMA($p$,$\delta$,$q$) Model

Estimates of $\delta_f$ were obtained by fitting the regression in [AI-9]. Figure AI-8 plots the ($X_1$,$Y$) inputs to this regression for chipped and circular-sawn samples over the first 50 frequencies.

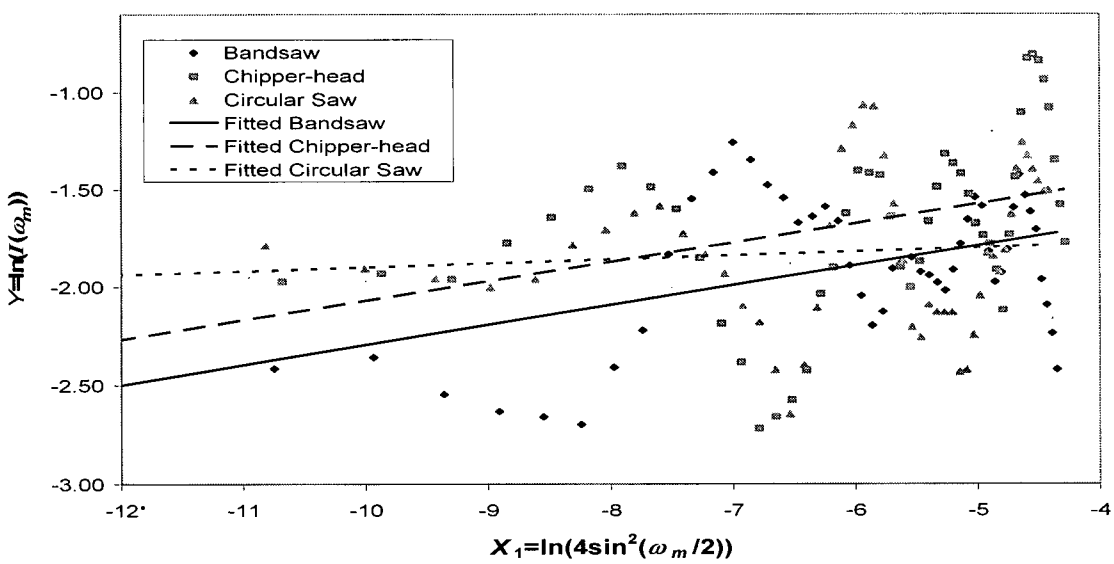



**Figure AI-8. Input to regression to estimate $\delta_f$ for three saw type samples.**

Descriptive statistics for the regression estimates of $\delta_f$ are listed in Table AI-7. Estimates of $\delta_f$ were approximately 0.1 higher for bandsawn and circular-sawn boards. Using a significance level of 0.05, most estimates were significantly less than one. The bound for stationarity (0.5) was contained in confidence interval around $\delta_f$ only for about 10% of chipped boards, and was not contained by the intervals for boards in either of the other saw types.

**Table AI-7. Summary of estimates of $\delta$ for ARFIMA($p,\delta,q$) model by saw type.**

| Saw Type | Average | Minimum | Maximum | Average Standard Error | % of Estimates Significantly < 1.0 ($\alpha$ = 0.05) | % of Confidence Limits Containing 0.5 |
|---|---|---|---|---|---|---|
| Bandsawn | 0.912 | 0.597 | 1.218 | 0.035 | 59% | 0% |
| Chipper-head | 0.818 | 0.429 | 1.223 | 0.046 | 76% | 11% |
| Circular Saw | 0.892 | 0.637 | 1.127 | 0.041 | 73% | 0% |

The ACF of the fractionally differenced series, $\varepsilon'_{fm} = \nabla^{\delta_f} \varepsilon_{fm}$, is shown in Figure AI-9 for samples from the three saw types. For bandsawn and circular-sawn boards, this figure is similar to that of the first differenced series because the $\delta_f$ values were close to 1.

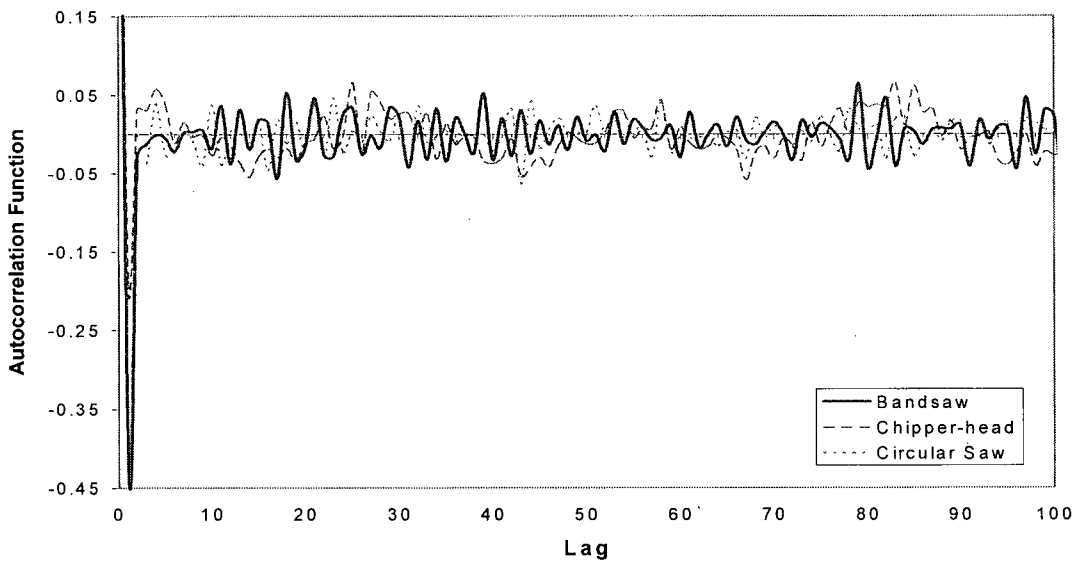



**Figure AI-9. ACF of fractionally differenced series for three saw type samples.**

Using $\varepsilon'_{fm} = \nabla^{\delta_f} \varepsilon_{fm}$, the optimal values of $p_f$ and $q_f$ were found for the ARFIMA($p,\delta,q$) model. The results of the MINIC procedure were slightly different than that of the ARIMA($p,1,q$) (Table AI-8). For both bandsawn and circular-sawn lumber, the model most often selected was ARIMA($1,\delta,1$). For the chipped lumber, ARIMA($0,\delta,1$) was most often selected.

**Table AI-8. Summary of most often selected orders of $p$ and $q$ by series for ARFIMA($p,\delta,q$) model.**

| Bandsaw | | | Chipper-head | | | Circular Saw | | |
|---|---|---|---|---|---|---|---|---|
| $p_f$ | $q_f$ | % Samples | $p_f$ | $q_f$ | % Samples | $p_f$ | $q_f$ | % Samples |
| 1 | 1 | 46% | 0 | 1 | 23% | 1 | 1 | 33% |
| 0 | 1 | 25% | 1 | 2 | 19% | 0 | 1 | 23% |
| 2 | 1 | 6% | 1 | 1 | 17% | 2 | 1 | 12% |
| 2 | 2 | 5% | 2 | 1 | 17% | 0 | 2 | 7% |

The selected model forms were therefore:

Bandsaw/Circular Saw (ARIMA(1,$\delta$,1)): $(1-\phi_{1f}B)\nabla^{\delta_f}\varepsilon_{fm} = \alpha_f + (1-\theta_{1f}B)\upsilon_{fm}$

Chipper-head (ARIMA(0,$\delta$,1)): $\nabla^{\delta_f}\varepsilon_{fm} = \alpha_f + (1-\theta_{1f}B)\upsilon_{fm}$

The estimates of $\alpha_f$ were not significantly different from zero for any series (Table AI-9). The estimates of $\theta_{1f}$ and $\phi_{1f}$ were significantly non-zero for almost all series, but varied strongly in comparison to the ARIMA($p$,1,$q$) estimates.

**Table AI-9. Summary of estimated ARFIMA($p,\delta,q$) model parameters by saw type (mm) ($\alpha$ = 0.05).**

| | Bandsaw | | | | Chipper-head | | | Circular Saw | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\hat{\alpha}_f$ | $\hat{\theta}_{1f}$ | $\hat{\phi}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ | $\hat{\alpha}_f$ | $\hat{\theta}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ | $\hat{\alpha}_f$ | $\hat{\theta}_{1f}$ | $\hat{\phi}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ |
| Average | 0.0000 | 0.0164 | 0.0068 | 0.1007 | -0.0001 | 0.0049 | 0.1144 | 0.0000 | 0.0155 | 0.0058 | 0.1026 |
| Minimum | -0.0015 | -0.0105 | -0.0120 | 0.0626 | -0.0025 | -0.0094 | 0.0841 | -0.0010 | -0.0196 | -0.0204 | 0.0698 |
| Maximum | 0.0017 | 0.0240 | 0.0198 | 0.1534 | 0.0029 | 0.0172 | 0.1697 | 0.0008 | 0.0242 | 0.0191 | 0.1414 |
| % Significant | 0.4% | 96% | 88% | | 0.0% | 92% | | 0.0% | 96% | 81% | |

### AI.3.5 SARFIMA($p,\delta,q$)x($P,\delta',Q$)$_s$ Model

Estimates of $\delta_f$ and $\delta'_f$ were found via multiple regression with $Y$=ln($I(\omega_m)$), $X_1$=ln(4sin$^2$($\omega_m$/2)), and $X_2$=ln(4sin$^2$($s\omega_m$/2)). The length of the seasonal cycle, $s_f$, was taken from the SARIMA models found in AI.3.3. These regressions were performed for chipped and circular-sawn boards only, since no bandsawn boards had obvious seasonal components. Figure AI-10 shows the plot of $Y$ versus $X_2$ for the two saw types. (Plots of $Y$ versus $X_1$ are shown in Figure AI-8.)

Descriptive statistics for the regression estimates of $\delta_f$ and $\delta'_f$ are listed in Table AI-10.

Whereas the range of estimates for $\delta_f$ was similar for chipped and circular-sawn series, the range of estimates for $\delta'_f$ was much larger for chipped versus circular-sawn series. Using a one-sided test with $\alpha$=0.05, most of the estimates of $\delta_f$ and $\delta'_f$ were significantly less than one. The bound for stationarity (0.5) was contained in confidence intervals around $\delta_f$ and $\delta'_f$ for only 15-50% of

series. However, 95% confidence intervals overlapped with the interval [0, 0.5] for all but one series. Thus, almost all of the estimated SARFIMA models were not significantly non-stationary.

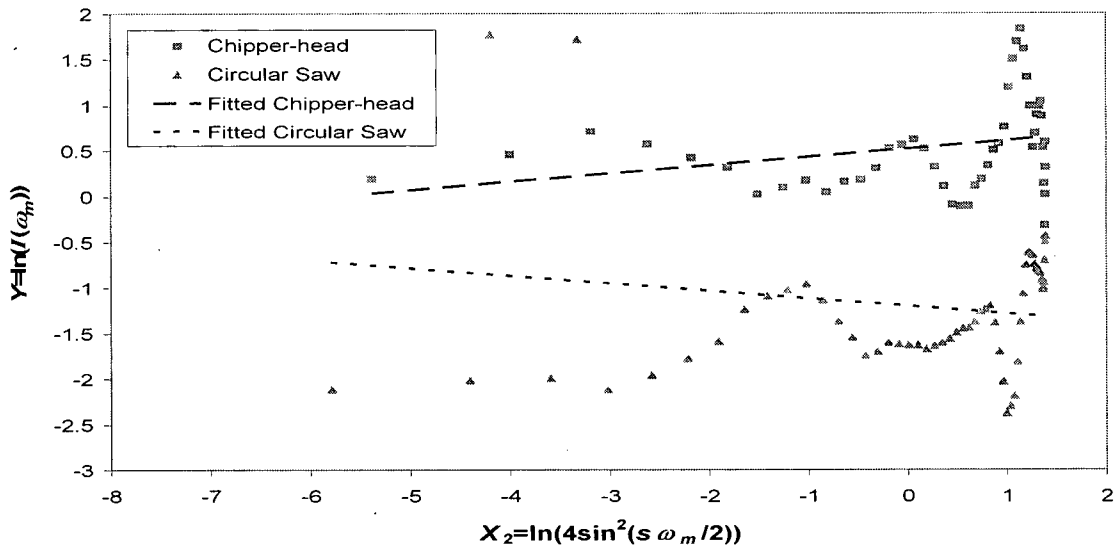



**Figure AI-10. Input to regression to estimate $\delta$ and $\delta'$ for two saw type samples.**

**Table AI-10. Summary of estimates of $\delta$ and $\delta'$ for SARFIMA($p,\delta,q$)x($P,\delta',Q$)$_s$ model by saw type.**

| Estimate | Saw Type | Average | Minimum | Maximum | Average Standard Error | % of Estimates Significantly < 1.0 ($\alpha$ = 0.05) | % of Confidence Limits Containing 0.5 |
|---|---|---|---|---|---|---|---|
| $\delta$ | Chipper-head | 0.677 | 0.304 | 1.000 | 0.039 | 97% | 21% |
| | Circular Saw | 0.587 | 0.308 | 0.975 | 0.028 | 97% | 40% |
| $\delta'$ | Chipper-head | 0.207 | -0.227 | 0.941 | 0.075 | 99% | 15% |
| | Circular Saw | 0.353 | 0.081 | 0.555 | 0.052 | 100% | 50% |

The ACF of the seasonally and fractionally differenced series, $\varepsilon''_{fm} = (1-B)^{\delta_f}(1-B^s)^{\delta'_f}\varepsilon_{fm}$, is shown in Figure AI-11 for chipped and circular-sawn samples. This figure is somewhat less stable than that of the first differenced series or the fractionally differenced series.

Table AI-11 shows the distribution of the best fitting (lowest AIC) model form by series for chipped and circular-sawn boards. For both types of sawing, the model that best fit in most cases was SARFIMA(1,$\delta$,1)x(1,$\delta'$,1)$_s$.

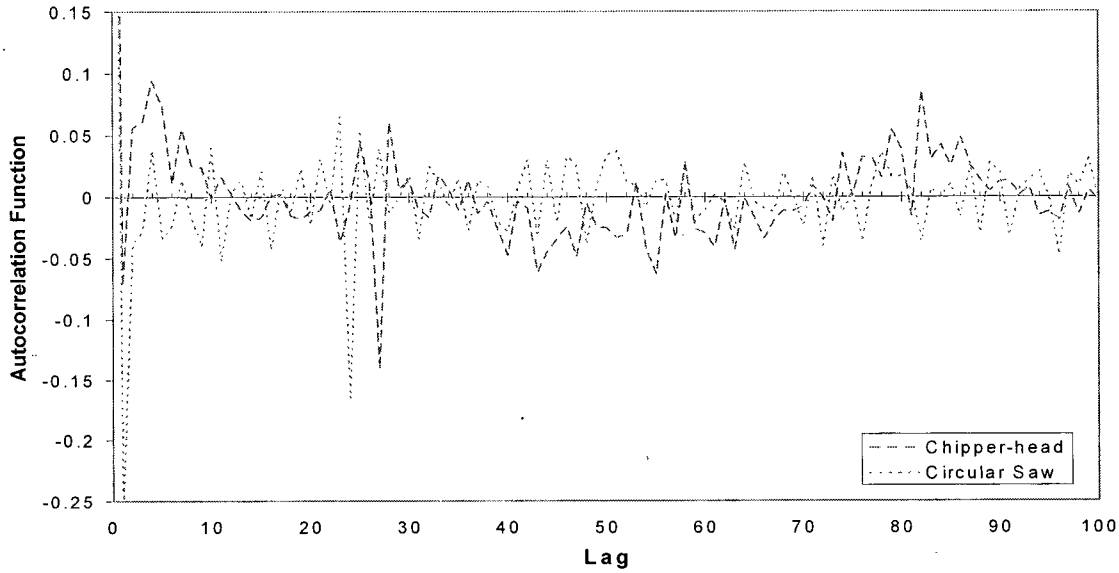



**Figure AI-11. ACF of SARFIMA($p,\delta,q$)x($P,\delta',Q$)$_s$ model for two saw type samples.**

**Table AI-11. Summary of lowest AIC values by $p$, $q$, $P$, and $Q$ by series for SARFIMA($p,\delta,q$)x($P,\delta',Q$)$_s$ models.**

| Chipper Head | | | | | Circular Saw | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $p_f$ | $q_f$ | $P_f$ | $Q_f$ | % Samples | $p_f$ | $q_f$ | $P_f$ | $Q_f$ | % Samples |
| 1 | 1 | 1 | 1 | 90% | 1 | 1 | 1 | 1 | 100% |
| 1 | 1 | 0 | 1 | 8% | | | | | |

The selected model form for both the chipped and circular-sawn lumber was therefore:

$$(1-\Phi_{1_f}B^{s_f})(1-\phi_{1_f}B)(1-B^{s_f})^{\delta'_f}(1-B)^{\delta_f}\varepsilon_{fm} = \alpha_f + (1-\theta_{1_f}B)(1-\Theta_{1_f}B^{s_f})\upsilon_{fm}$$

Descriptive statistics for the model parameters are shown in Table AI-12. The estimated values of $\theta_{1_f}$, $\Theta_{1_f}$, $\phi_{1_f}$, and $\Phi_{1_f}$ were significantly non-zero for almost all series. Estimates of $\theta_{1_f}$ and $\Theta_{1_f}$ were positive in all cases, indicating a strong mixing process; estimates of $\phi_{1_f}$ and $\Phi_{1_f}$ were slightly more varied. Unlike previous models, the estimate of $\alpha_f$ was significantly non-zero for 11-13% of series.

## AI.3.6 Evaluation and Comparison of Final Model Forms

Fit statistics are summarized in Table AI-13. The residual sums of squares and the AIC were summed over all sample series by saw type. For cases where a seasonal model was not used (e.g., all bandsawn boards), the numbers reported for the SARIMA model are those of the

ARIMA, and those reported for the SARFIMA are those of the ARFIMA. The R-square values for all models were very high, indicating that more than 90% of the within sample board variation was explained by the models. For all saw types, the residual sums of squares was lowest when considering the simple ARIMA($p$,1,$q$) model. The ARIMA model fit the bandsawn data the best, as indicated with the lowest AIC and the highest R-squared. For chipped and circular-sawn boards, the lowest AIC was achieved with the SARIMA model.

**Table AI-12. Summary of estimated SARFIMA(1,$\delta$,1)x(0, $\delta$,1)$_s$ model parameters by saw type (mm) ($\alpha$ = 0.05).**

| Saw Type | Statistic | $\hat{\alpha}_f$ | $\hat{\phi}_{1f}$ | $\hat{\Phi}_{1f}$ | $\hat{\theta}_{1f}$ | $\hat{\Theta}_{1f}$ | $\hat{\sigma}_{\upsilon f}$ |
|---|---|---|---|---|---|---|---|
| Chipper-head | Average | 0.0001 | -0.0040 | -0.0025 | 0.0179 | 0.0200 | 0.1425 |
| | Minimum | -0.0006 | -0.0082 | -0.0089 | 0.0109 | 0.0130 | 0.0891 |
| | Maximum | 0.0010 | 0.0020 | 0.0033 | 0.0233 | 0.0226 | 0.2315 |
| | % Significant | 13% | 90% | 73% | 100% | 100% | |
| Circular Saw | Average | 0.0000 | -0.0032 | -0.0036 | 0.0194 | 0.0200 | 0.1381 |
| | Minimum | -0.0006 | -0.0074 | -0.0060 | 0.0139 | 0.0145 | 0.0847 |
| | Maximum | 0.0004 | 0.0029 | -0.0013 | 0.0236 | 0.0227 | 0.1788 |
| | % Significant | 11% | 95% | 95% | 100% | 100% | |

**Table AI-13. Summary of corrected residual sums of squares (*SSres*), AIC, and $R^2$ by model and saw type.**

| Saw Type | Fit Statistic | ARIMA ($p$,1,$q$) | SARIMA (1,1,$q$)x(1,1,1)$_s$ | ARFIMA ($p$,$\delta$,1) | SARFIMA (1,$\delta$,1)x(1,$\delta$,1)$_s$ |
|---|---|---|---|---|---|
| Bandsaw | *SSres* | 10,303,619 | 10,303,619 | 11,219,229 | 11,219,229 |
| | AIC | 3,898,657 | 3,898,657 | 3,972,825 | 3,972,825 |
| | R-square | 94.5% | 94.5% | 94.0% | 94.0% |
| Chipper-head | *SSres* | 4,682,481 | 5,277,605 | 5,191,842 | 7,336,611 |
| | AIC | 1,461,869 | 1,443,002 | 1,490,898 | 1,461,715 |
| | R-square | 95.2% | 94.6% | 94.7% | 92.6% |
| Circular Saw | *SSres* | 3,636,654 | 3,829,988 | 3,762,760 | 4,760,074 |
| | AIC | 1,277,496 | 1,269,729 | 1,286,845 | 1,285,917 |
| | R-square | 93.0% | 92.7% | 92.8% | 90.9% |

The ACFs of the residuals from the final forms of each model are plotted in Figures AI-12 – AI.15. All plots are shown to lag 200 to emphasize cyclical patterns where they are present. In particular, chipped lumber and circular-sawn lumber showed strong patterns for all models except the SARFIMA model, which appeared to be free of any cyclical pattern.

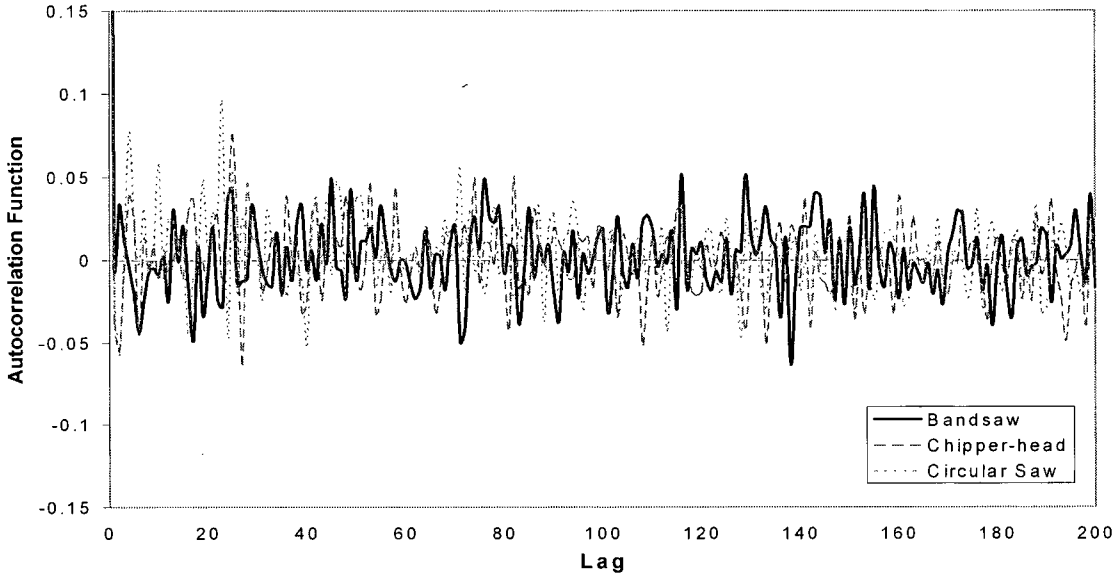



**Figure AI-12. ACF of ARIMA($p$,1,$q$) model residuals for three saw type samples.**

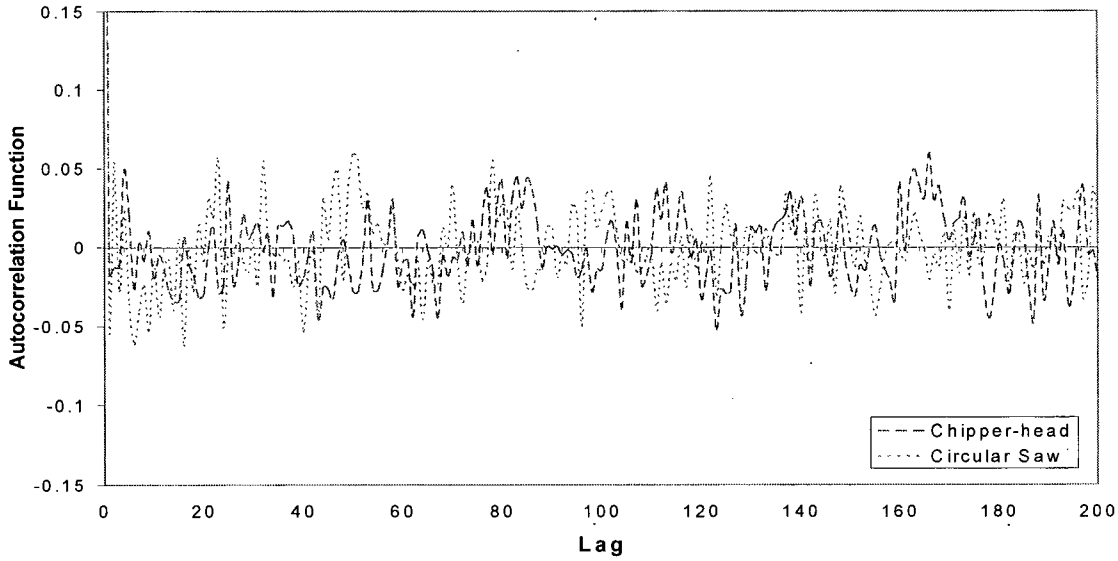



**Figure AI-13. ACF of SARIMA(1,1,1)x(0,1,1)$_s$ model residuals for two saw type samples.**

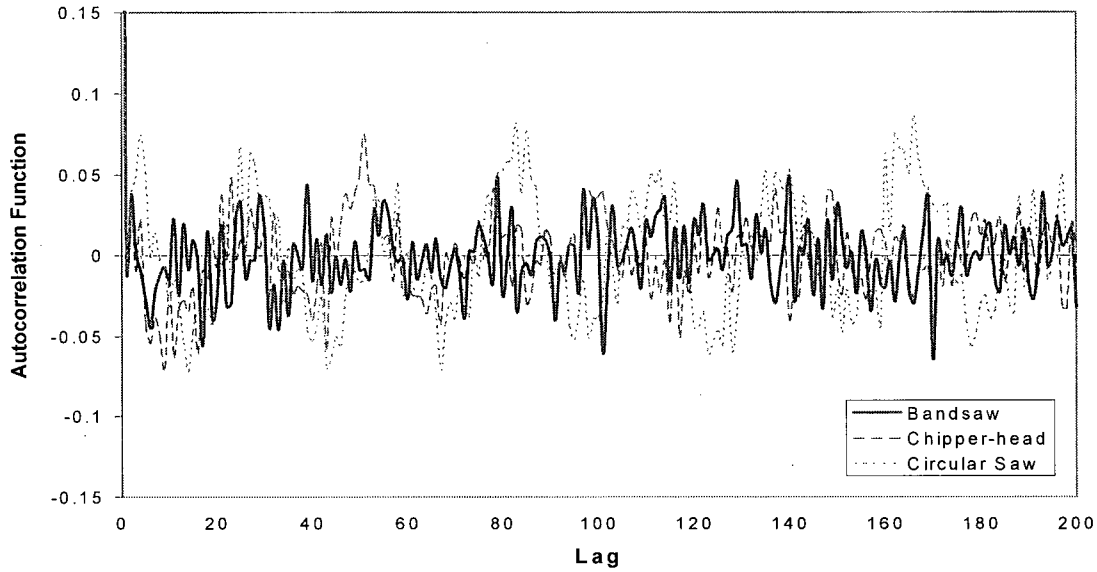



**Figure AI-14. ACF of ARFIMA($p,\delta,q$) model residuals for three saw type samples.**

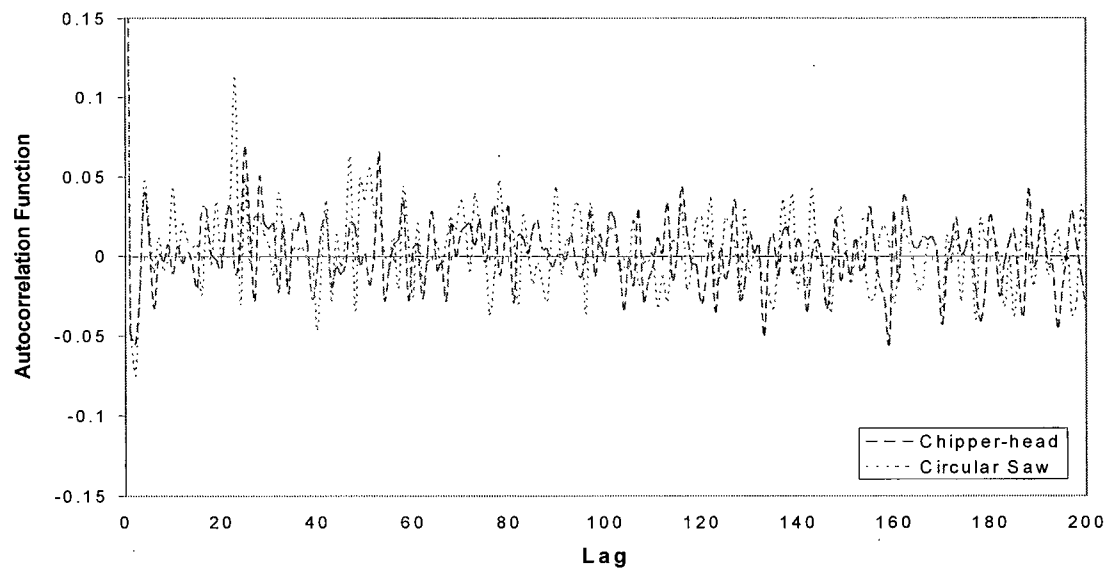



**Figure AI-15. ACF of SARFIMA(1,$\delta$,1)x(0,$\delta$,1)$_s$ model residuals for two saw type samples.**

## *AI.4 Discussion and Conclusion*

Given the results of the Dickey-Fuller tests, all series were assumed non-stationary. Although this decision was conservative, it led to several good-fitting candidate models and was appropriate for the sawing process. Non-stationary models describe processes that drift over time, and the mechanical process of sawing is compatible with this description.

Seasonal components were important in model selection, as cyclical patterns were present in both the circular-sawn and chipped data. The mechanical process that produces chipped surfaces produces cyclical patterns by virtue of its comparatively low cutting drum speed and coarseness of cut. Patterns on the chipped lumber are quite easily identifiable by the layman without magnification or other visual aid. The most frequent cycle length (22 observations, or 2 cm) roughly corresponds to one rotation of the cutting drum under usual cutting speeds used by the mill. Although visible cycles are less frequent in circular sawing, they can also be linked to the different mechanical processes. Dishing of the saw can cause a washboarding pattern. A bent tooth on a circular saw may also cause the saw to wobble, producing similar patterns on the lumber surface. Anything that causes the saw to heat up and lose its stiffness may lead to these kinds of surface defects, such as improper tensioning and poor guides (Schajer 1989).

The bandsawn data, on the other hand was notable in its lack of obvious seasonal components. The process of bandsawing, however, could produce a cyclical pattern if, for instance, saw tension was low or a tooth was severely bent.

For bandsawn boards and circular-sawn boards without saw marks, the ARIMA models had better fit statistics than the more complicated ARFIMA model and model residuals were free of patterns indicating lack of fit. Moreover, the AR and MA parameter estimates were in a consistent range, indicating a strong mixing process compatible with the mechanical processes of these types of sawing. For these types of sawing, the ARIMA is recommended.

The SARFIMA(1,$\delta$,1)x(1, $\delta$,1)$_s$ model provided a good fit for chipped lumber and for circular-sawn lumber with saw marks. Estimates of $\delta_f$ and $\delta_f$ in the [0, 0.5] range suggested stable long-memory components existed in both the seasonal and non-seasonal parts of the model. The presence of AR and MA parameters suggested that the model also included significant seasonal and non-seasonal short-term autocorrelation. Further, this model did not suffer from any of the

lack of fit problems obvious in other models. The appropriateness of this model may again be a consequence of the different mechanical processes involved in the three types of sawing. Long-memory models are appropriate for detecting low frequency events, whereas the process of bandsawing involves high speed, and therefore high frequency patterns. For chipped lumber and circular-sawn lumber with saw marks, the SARFIMA model is recommended.

Using the SARFIMA model, there were a substantial number of series with a significant trend component, implying that there was a consistent thickening (or thinning) along the length of the board. Plots of the LRS data series verified that the boards in question did have an obvious trend; however, this trend was on the order of 1.3 mm over 2.4 m. This is a good example of how a large number of observations can make statistical significance different from practical significance.

It is recommended that the ARIMA and SARFIMA models be used to describe the errors within LRS as part of the autocorrelated errors model. These model forms were chosen based on superior fit statistics, but more importantly, their residuals showed no evidence of lack of fit.

## *AI.5 Literature Cited*


Andel, J. 1986. Long memory time series models. Kybernetika 22(2): 105-123.

Beran, J. 1992. Statistical methods for data with long-range dependence. Statistical Science 7(4): 404-416.

Beran, J. 1994. Statistics for Long Memory Processes. New York, Chapman Hall.

Box, G. E. P. and G. M. Jenkins. 1970. Time Series Analysis: Forecasting and Control. 1st ed. San Francisco, Holden-Day.

Brocklebank, J. C. and D. A. Dickey. 2003. SAS for Forecasting Time Series. 2nd ed. Cary NC, SAS Institute, Inc.

Butler, J. 1999. Simulation and estimation of fractionally integrated time series. *In* Proceedings of Twenty-fourth Annual SAS Users Group International Conference, April 11-14, Miami Beach, Florida. SAS Institute. pp. 271-275.

Carlin, J. B. and A. P. Dempster. 1989. Sensitivity analysis of seasonal adjustments: empirical case studies. Journal of the American Statistical Association 84(405): 6-20.

Dickey, D. A. and W. A. Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 79: 355-367.

Geweke, J. and S. Porter-Hudak. 1983. The estimation and application of long memory time series models. Journal of Time Series Analysis 4(4): 221-238.

Granger, W. J. and R. Joyeux. 1980. An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1(1): 15-29.

Hosking, J. R. M. 1980. The multivariate portmanteau statistic. Journal of the American Statistical Association 75: 602 -608.

Hosking, J. R. M. 1984. Modeling persistence in hydrological time series using fractional differencing. Water Resources Research 20(12): 1898-1908.

Noffsinger, J. R. and R. B. Anderson. 2002. Effect of autocorrelation on composite panel production monitoring and control: a comparison of SPC techniques. Forest Products Journal 52(3): 60-67.

Porter-Hudak, S. 1990. An application of the seasonal fractionally differenced model to the monetary aggregates. Journal of the American Statistical Association 85(410): 338-344.

Ray, B. K. 1993. Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model. International Journal of Forecasting 9: 255-269.

SAS Institute. 2002. SAS/STAT User's Guide, Version 8. Cary, NC, SAS Publishing.

Schajer, G. S. 1989. The importance of circular saw tensioning. Forest Industries/World Wood Technical Report (Sawing Technology): T14-T16.

Vander Wiel, S. A. 1996. Monitoring processes that wander using integrated moving average models. Technometrics 38(2): 139-151.

# Appendix II Computation of Mean Squares

## *AII.1 Non-grouped Data*

Mean Squares Due to Board Effects:

$$MS_{\beta_{ij}} = \frac{2\bar{n}_{ij\cdots}}{b_i - 1}\sum_{k=1}^{b_i}(\bar{y}_{ijk\cdot\cdot} - \bar{y}_{ij\cdots})^2$$

Mean Squares Due to Laser Position Effects:

$$MS_{\lambda_{ij}} = \frac{b_i\bar{n}_{ij\cdots}}{2-1}\sum_{l=1}^{2}(\bar{y}_{ij\cdot l\cdot} - \bar{y}_{ij\cdots})^2$$

Mean Squares Due to Board × Laser Position Effects:

$$MS_{\beta\lambda_{ij}} = \frac{\bar{n}_{ij\cdots}}{(2-1)(b_i - 1)}\sum_{k=1}^{b_i}\sum_{l=1}^{2}(\bar{y}_{ijkl\cdot} - \bar{y}_{ijk\cdot\cdot} - \bar{y}_{ij\cdot l\cdot} + \bar{y}_{ij\cdots})^2$$

where: $b_i$ is the number of boards sampled in the $i$th saw configuration;

$\bar{n}_{ij\cdot\cdot}$ is the average number of observations per board and laser position, for the $i$th saw configuration and $j$th side;

$\bar{y}_{ijk\cdot\cdot}$ is the average profile value for the $i$th saw configuration, $j$th side, and $k$th board;

$\bar{y}_{ij\cdots}$ .is the average profile value for the $i$th saw configuration and $j$th side;

$\bar{y}_{ij\cdot l\cdot}$ .is the average profile value for the $i$th saw configuration, $j$th side, and $l$th laser position; and

$\bar{y}_{ijkl\cdot}$ .is the average profile value for the $i$th saw configuration, $j$th side, $k$th board, and $l$th laser position.

## *AII.2 Subgrouped Data*

Mean Squares Due to Board Effects in *g*th Group:

$$MS_{\beta_{ijg}} = \frac{2\bar{n}_{ijg\cdots}}{G-1}\sum_{k=1}^{G}(\bar{y}_{ijgk\cdot\cdot} - \bar{y}_{ijg\cdots})^2$$

Mean Squares Due to Laser Position Effects in *g*th Group:

$$MS_{\lambda_{ijg}} = \frac{G\bar{n}_{ijg\cdots}}{2-1}\sum_{l=1}^{2}(\bar{y}_{ijg\cdot l\cdot} - \bar{y}_{ijg\cdots})^2$$

Mean Squares Due to Board × Laser Position Effects in *g*th Group:

$$MS_{\beta\lambda_{ijg}} = \frac{\bar{n}_{ijg\cdots}}{(2-1)(G-1)}\sum_{k=1}^{G}\sum_{l=1}^{2}(\bar{y}_{ijgkl\cdot} - \bar{y}_{ijgk\cdot\cdot} - \bar{y}_{ijg\cdot l\cdot} + \bar{y}_{ijg\cdots})^2$$

where: $G$ is the number of boards per subgroup;

$\bar{n}_{ijg\cdot\cdot}$ is the average number of observations per board and laser position in the *g*th group, for the *i*th saw configuration and *j*th side;

$\bar{y}_{ijgk\cdot\cdot}$ is the average profile value for the *i*th saw configuration, *j*th side, *g*th group, and *k*th board;

$\bar{y}_{ijg\cdots}$ .is the average profile value for the *i*th saw configuration, *j*th side, *g*th group;

$\bar{y}_{ijg\cdot l\cdot}$ .is the average profile value for the *i*th saw configuration, *j*th side, *g*th group, and *l*th laser position; and

$\bar{y}_{ijgkl\cdot}$ .is the average profile value for the *i*th saw configuration, *j*th side, *g*th group, *k*th board, and *l*th laser position.