Diagnostic analysis of board foot per ton of chips ratio for a small sawmill

Brian K. Via*
Todd F. Shupe*

Abstract

The board foot to ton of chips (BFTC) ratio is a common lumber quality index used by small sawmills to gauge sawing efficiency. Producing too many chips, relative to board footage at any given time, is considered inefficient since a unit conversion into lumber is clearly more valuable than a unit conversion into ton of chips. Sawyers and managers thus assume normality and independence of the BFTC ratio when they attempt to identify outlying performance. The objectives of this study were to determine: a) the distributional properties of BFTC ratio; b) if a negative autocorrelation exists between board footage and ton of chips; c) the appropriate diagnostics for BFTC ratio to determine abnormal deviation from mean process performance; and d) if differences in BFTC models between species exist. The results of this study demonstrate a skewed distribution. Also, the independence assumption was violated as a result of autocorrelation between consecutive months. A residual plot versus time series chart from an auto regression model provided a better estimate of outlier boundaries, which was needed to detect process errors or changes. All processes appeared to follow a linear autocorrelation model instead of the traditional least squares regression. Absence of significant auto regression for poplar was an indication of a process change at 18 months, which masked the ability to detect the autoregressive correlation. Identification of the process change and the creation of two models for poplar confirmed the autoregressive nature for all three species.

Large sawmills usually have the financial resources available to measure variables necessary to monitor real time sawing efficiency. Such ability enables them to make immediate adjustments to the process. The most common index measured is the lumber recovery factor (LRF), which is simply the nominal board feet per cubic foot of volume of the log. Scanners can be used to estimate the volume of input log material and to measure the board footage of sawn lumber. Some factors that increase LRF are increased log diameter, decreased kerf, decreased sawing variation, and better log positioning (Savsar and Kersavage 1982, Steele and Risbrudt 1985, Wade et al. 1992, Keegan et al. 1998).

Smaller sawmills have shown to yield lower LRF values and have been classified as being mills who produce less than 5 million board feet (MMBF) per year (Steele and Risbrudt 1985). One disadvantage of smaller sawmills is they usually do not possess the capital necessary to monitor real time process indices. Quite often, these sawmills only measure and record what is needed for accounting purposes; i.e., board footage per species, tons of chips per species (or species mix), log costs, and lumber sales. A company accountant may not formally organize these values, commonly measured during the point of sale, until the end of the month.

Small sawmills often cannot afford to measure real time quality control indices like LRF. Instead, they may use the lumber board foot to ton of chips ratio (BFTC) as measured during a given period of time. The chips are usually sold to local paper mills for a lower price than if the same fiber could have been converted into lumber. As a result, the goal is usually to produce more board footage of lumber per ton of chips. Likewise, identification of low ratios may help to identify an adverse process change.

One assumption made about the BFTC ratio is that the data should be normally distributed and independent of one another. With this assumption, a basic t-distribution can be utilized to determine if a sample point exceeds upper

The authors are, respectively, Graduate Research Assistant and Assistant Professor, School of Forestry, Wildlife, and Fisheries, Louisiana State Univ. Agricultural Center, Baton Rouge, LA 70803. This paper (No. 03-40-1364) is published with the approval of the Director of the Louisiana Agri. Expt. Sta. This paper was received for publication in January 2003. Article No. 9613.
*Forest Products Society Member.

Forest Prod. J. 54(4):88-94.

and lower confidence limits. If any BFTC exceeds the confidence limits, then an abnormal event may have occurred. Too low of a ratio may justify an investigation for problems in the manufacturing process. Quite often, the Shewhart control chart is used to plot and identify a change in the manufacturing process when data are independent of one another (Hunter 1986, Young and Winistorfer 2001).

The assumption of independence between BFTC measurements is likely to be violated if the sale of chips occurs at a different time than the sale of lumber. In such an event, a negative autocorrelation pattern would be expected since products not sold one month, are likely to be sold on the consecutive month, thus adding an error component at time *t*. Negative autocorrelation is when the residuals of a model tend to follow a positive, negative, positive pattern over time assuming the proper regression model is used. Most forest products processes follow a positive autocorrelation since measurements are made in real-time (Cook 1992). When autocorrelation exists between consecutive samples, one has an additional error term not accounted for by least squares regression. As a result, false confidence limits may be derived, thus leading to inappropriate investigations into process behavior (Cook and Massey 1990, Cryer and Ryan 1990, Montgomery and Mastrangelo 1991, Cook 1992, Atienza et al. 1998). Auto regression is a viable method to account for time series correlation between subsequent samples. Such models, demonstrated to be linear in this work, can provide appropriate confidence limits adjusted for the additional error term present. Exceeding these limits can indicate a change or error in the manufacturing process.

The objectives of this paper were to determine: a) the distributional properties of BFTC ratio; b) if a negative autocorrelation exists between board footage and ton of chips; c) the appropriate diagnostics for BFTC ratio to determine abnormal deviation from mean process performance; and, d) if differences in BFTC models between species exist.

Materials and methods

Processing

The selected sawmill was located in the north central portion of North Carolina. It produces an average of 3.5 MMBF per year of both hardwood and softwood lumber including ash, beech, birch, cedar, cherry, hickory, black gum, sweet gum, maple, pine, poplar, red oak, walnut, and white oak. The primary species used in the analysis were red and white oak combined, poplar, and pine, which represent the largest three sources of species groups processed at the mill. When either oak, poplar, or pine was processed, only that species was processed, making inventory tracking of these three species accurate.

A circle sawblade was used with one sawyer for the first 18 months. After 18 months, a less experienced sawyer was hired and trained to work on a part-time basis to relieve the primary sawyer. The secondary sawyer helped to process lower value species while the primary sawyer still cut most of the higher value species (oak, for example).

Species were organized in the log yard so that one species at a time could be cut. Typical products made by the sawmill were 5/4 and 4/4 lumber for furniture and flooring stock orders. Pallet stock of various sizes was also milled. Cants were also sometimes processed for special orders. A conventional system for cutting logs was used. Logs were manually positioned by the sawyer at the head saw on a dogging carriage with the goal of optimizing volume for pallet stock. For furniture or flooring stock, the goal was to maximize volume and lumber visual grade. The set works for the head saw were hydraulic and a back stand indicator was used to determine the proper positioning of the log needed to cut the thickness of the next piece of lumber.

After processing, the wet lumber was stacked on stickers. A kiln was not available at the mill and thus typical buyers were usually distributors who manage or have access to kilns. Chips were conveyed from the manufacturing process to a truck bin such that the weight of the wet chips could be measured via a truck scale and shipped to a local buyer. Sawdust was also collected and sold on a weight basis. Chip tons and board foot were recorded monthly.

Log, lumber, and chip inventory

Logs were organized by species in the log yard upon arrival for the three main species of interest: oak, poplar, and pine. After processing a batch of just one species at a time, the lumber was stacked by species, measured for board footage, and recorded on a daily basis. This daily figure was summed to yield a monthly figure. Also, each load of chips per species was loaded onto a truck and sent to a paper mill. The paper mill received the chips 1 to 5 days later and weighed the ton of chips. The sawmill then added this weight to their records. All pine chips were kept separate due to a higher dollar value paid by the paper mill to separate the pine chips. The total ton of chips was also summed up at the end of the month. An additional error occurred when measuring the ton of chips for oak or poplar. If a separate hardwood species was cut either before or after a run of oak or poplar, then the load of chips would get contaminated with other species for that particular load of chips.

Statistical diagnostics

BFTCs for each month and for each species group analyzed were determined by dividing the board footage by the ton of chips. Histograms were developed to determine if a normal distribution existed. If a distribution was normal and independently distributed, then a Shewhart chart would be used to determine outlying values. If either of the assumptions were violated, other diagnostics needed to be employed.

Ordinary least squares (OLS) regression and auto regression models were developed with board feet as the independent variable and ton of chips as the dependent variable. The Durbin-Watson (*d*) statistic Eq. [**1**] was calculated from the residuals of both models to determine if autocorrelation of the error terms was present. If the *d* statistic was not significantly greater than the *d* value with an alpha = 0.05 and the appropriate degrees of freedom, then an OLS model was determined to be adequate. If the *d* statistic was significant, then correlation between error terms was deemed to exist and an autoregressive model was evaluated. As a rule of thumb, values of *d* significantly less than 2 are positively correlated while those significantly greater than 2 are negatively correlated (Johnson et al. 1994). One-tailed tests of the *d* statistic were computed by SAS software, which used Eq. [**1**] and developed a *p*-value (Johnson et al. 1994); *p*-values less than 0.05 were determined to be significant and an autocorrelation was determined to exist. Proc AUTOREG was used to calculate *d*, autoregressive, and OLS models (SAS 2001). A lag time of only 1 month was tested since the likeli-

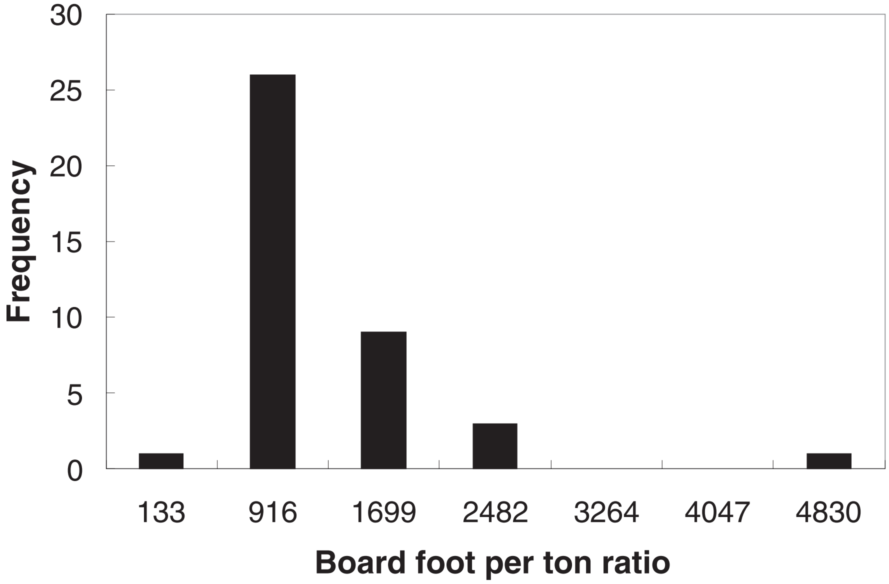

Figure 1. — BFTC ratio histogram for oak.

hood of lumber being produced and sold in 2 or more months was unlikely.

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^{\,2}} \qquad [\mathbf{1}]$$

where:

e = error of residual

t = time

The residual of a sample point is the actual dependent value (ton of chips) minus the predicted dependent value given by a regression model. Thus, when the model is properly used, the residual should be independent and normally distributed around zero. Any deviation from the zero mean is indicative that an improper model was likely used. For process manufacturing, a consistent deviation from zero was used to indicate a change in the process and hence the need for a new model. Residual plots against time were the control charts plotted to visually detect process shifts. Only upper confidence limits for predicted values were plotted to detect outlying values since it was undesirable to have a significantly greater ton of chips per board foot of lumber. Any value below the lower confidence limit would simply indicate superior performance in production and could have been plotted.

The first order autoregressive error model used is given in Eq. [**2**] (Neter et al. 1996). This model was identical to an OLS model with the exception of the structure of the error terms. The error terms consist of a fraction of the previous error term plus the addition of a new disturbance term.

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 X_t + e_t \\ e_t &= pe_{t-1} + \mu_t \end{aligned} \qquad [\mathbf{2}]$$

where:

p = some absolute value less than 1

μ_t = is a disturbance term and is independent and normally distributed around zero

X_t = board foot at time t

β_0 and β_1 = intercept and slope

Results and discussion

Distributive properties of BFTC

Histograms were charted to visually determine if normality existed for BFTC. Since this and other area sawmills use BFTC, a violation from normality and/or independence would mean that false decisions could be made via the use of traditional control charting. In other words, Type I or Type II errors would not occur in a predicted frequency as set by alpha and thus biasing the ability to detect outlying events of poor performance (too many chips per board feet produced). **Figure 1** shows a bar histogram for mixed oak. Similar histograms were plotted, but not reported, for poplar and pine. Just like the oak group, pine and poplar exhibited extreme left-handed skewness with the median being less than the mean.

The normal distribution was not appropriate to describe BFTC. As a result, traditional control charting (Shewhart control chart) would not adequately detect low outlying values. Such low values would indicate that too many chips were being produced relative to the amount of board feet produced. One reason for the non-normal distribution was the fact that the mill could never produce a BFTC below zero, thus skewing the histogram left.

Nonparametric models may accurately describe this distribution. However, most nonparametric models assume that each sample point is independent of each other. In reality, these points were not expected to be independent of one another. It is possible, for example, that chips could have been sold during the month of production while the lumber may not have been sold until the next month after production. If this occurred, one would expect some degree of negative autocorrelation. Such behavior would violate the independence assumption and make traditional control charting an improper tool to monitor performance.

Producing a histogram of all species also yielded a non-normal distribution. However, it did not show the same distributive pattern as when each species was plotted separately. It is thus suggested that one monitor each species separately to gain a better understanding of process distributive properties for each species. Given today's computer computational ability, one could easily monitor charts for different species.

Another concern is that the BFTC distributional properties will change from time to time, especially if a group of small- or large-diameter logs are run consecutively or nearly so. If this occurs, than one might wrongly deduce that an error in the manufacturing process has occurred based on the 95 percent confidence interval. However, most if not all sawyers would recognize when a run of trees is unusually large or small in diameter, or if the range of diameters was unusually different whether it be for a full day, week, or month. As a result, it is important for the sawmiller to utilize some common sense when relying on confidence intervals to determine if a process is out of control. A concern should only occur after one first tries to detect any

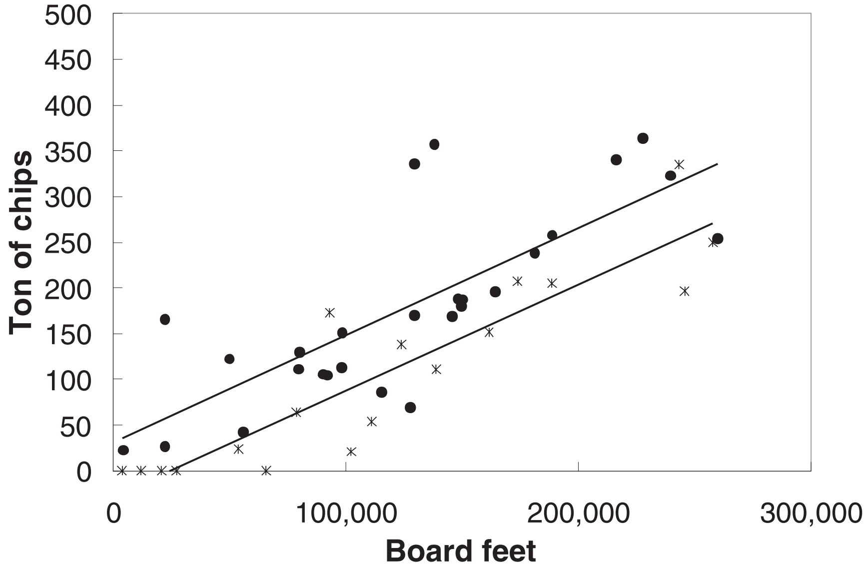

Figure 2. — Board feet versus ton of chips for poplar with the upper regression line (higher intercept) being the adjustment for a process change.

Table 1. — Models and test statistics for three major species produced with board feet (B) as the independent variable and tons (T) as the dependent variable.

Species	Durbin-Watson (OLS)	Durbin-Watson (AUTO)	p - value (OLS)	p - value (AUTO)	Model	r^2
Pine	2.58	2.02	0.0445	0.5009	T = 0.0012*(B) + 40.46	0.8147
Oak	2.58	2.12	0.0195	0.3434	T = 0.0015*(B) – 5.71	0.8376
Poplar[a]	2.00	1.99	0.5055	0.5164	T = 0.0012*(B) + 3.44	0.6855
Poplar ≤18[b]	2.26	--	0.2512	--	T = 0.0012*(B) – 27.72	0.8257
Poplar >18[c]	2.58	2.23	0.0631	0.3079	T = 0.0012*(B) + 30.56	0.6323

[a]An overall model encompassing all months.
[b]A separate model for up to 18 months.
[c]A separate model for greater than 18 months.

unusual changes in raw material diameter or quality.

Since mean log diameter can be dynamic, one could benefit by monitoring the diameter of every log coming in to account for the influence of log diameter on BFTC ratio. Such an addition would significantly improve the usefulness of the BFTC ratio in quality control. However, this would require log scanners, which are a substantial investment for most small sawmills. As a result, many have no choice but to rely on the seemingly primitive BFTC ratio to determine when the process is out of control. Confounding the problem, such troubles are detected only after weeks or months have past, causing a considerable reduction in profits. To reduce this time lag, a sawmill needs to consider measuring BFTC on as frequent a basis as possible.

Autocorrelation

It is possible that chips could be sold during the month of production while some lumber, produced at the same time, may have been sold the following month. If so, one would expect a negative autocorrelation. To test this hypothesis ($Pr > d$), the d statistic and corresponding p-value were computed for both OLS and auto regression (AUTO) models (**Table 1**). If the d statistic for the OLS model had a p-value greater than 0.05 then the OLS model was retained. If the p-value was less than 0.05, then an AUTO model was derived.

Pine and oak both had p-values less than 0.05, confirming the hypothesis that autocorrelation existed. When autocorrelation models were computed, the d statistic dropped close to 2, indicating that the extra error term had been properly accounted for. **Table 1** gives the resulting models with r^2 values above 0.80. The slopes for oak were not significantly greater than pine. Prior to the analysis, it was hypothesized that the oak may have a higher slope since oak is usually of higher density than poplar or pine.

Poplar had a lower r^2 value of 0.6855 and did not exhibit any autocorrelation. This was surprising given the strong autoregressive nature of the other two groups. However, as will be further discussed in the diagnostics section, there was a process change around month 18 due to the addition of a new sawyer. Since poplar was the lowest dollar value product, replacement and training occurred most frequently for the poplar group. As a result of this process change, two separate models were developed for months 1 to 18 and 19 + (**Table 1**). The first 18 months had only the experienced sawyer processing the logs, which would explain the lower intercept. The lower intercept means that the sawyer tends to cut less tons of chips per board feet of lumber. For the first 18 months, the board footage also accounted for almost 20 percent more of the variation in tons of chips than the latter model. This would make sense because a second sawyer might be more variable in efficiency. **Figure 2** demonstrates the shift in intercept with the experienced sawyer (first 18 months) having the higher intercept. It is interesting to note that no shift in slope occurred. For quality control, one could use an indicator variable to adjust the intercept for different sawyers.

Diagnostics

Residual plots with upper confidence limits were plotted to determine when a deviation from the process mean occurred. Exceeding the upper confidence limit may suggest that too many chips were produced with respect to board feet. This was particularly useful in the case of oak and poplar where additional error was expected to occur due to occasional contamination of oak or poplar chips when a different species was scheduled to run either before or after the oak or poplar run. This can be illustrated in **Figures 3** and **5**, where poplar had a 95 percent confidence interval residual of almost 150 tons; while pine, which was not allowed to mix with other species, had a lower 95 percent confidence interval residual of 100 tons.

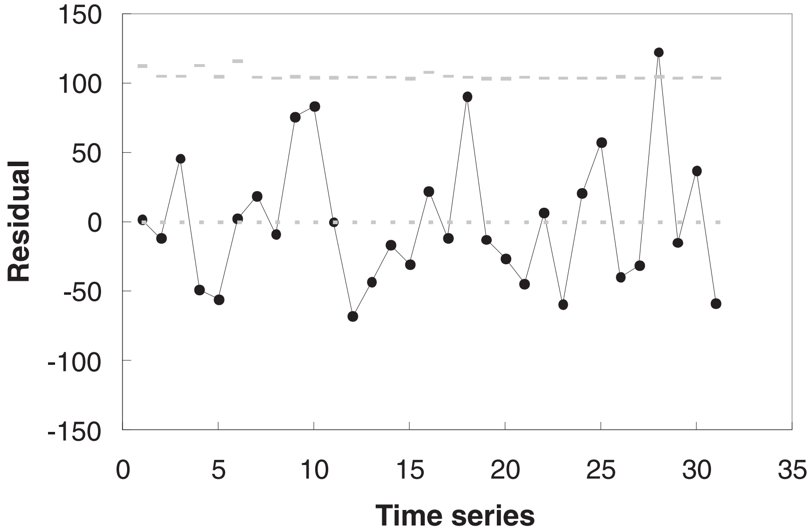

Figure 3. — Residual plots from an auto regression model with 95 percent upper confidence prediction limits for pine.

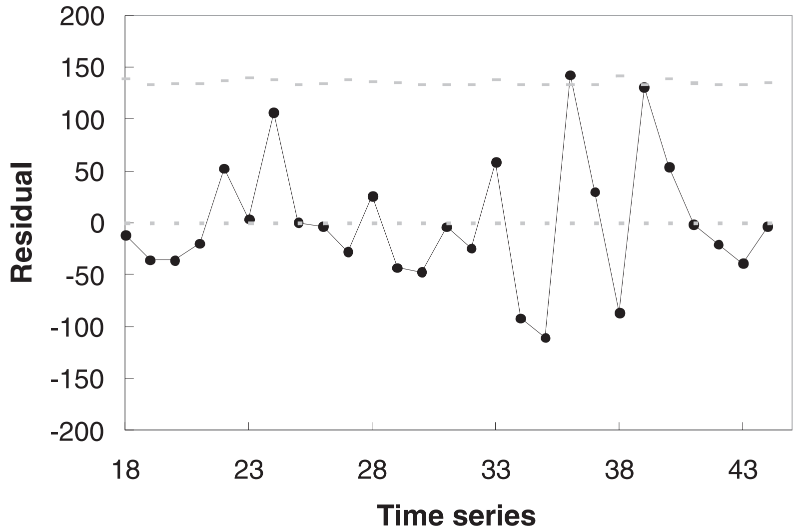

Figure 4. — Residual plots from an auto regression model with 95 percent upper confidence prediction limits for poplar > 18 months.

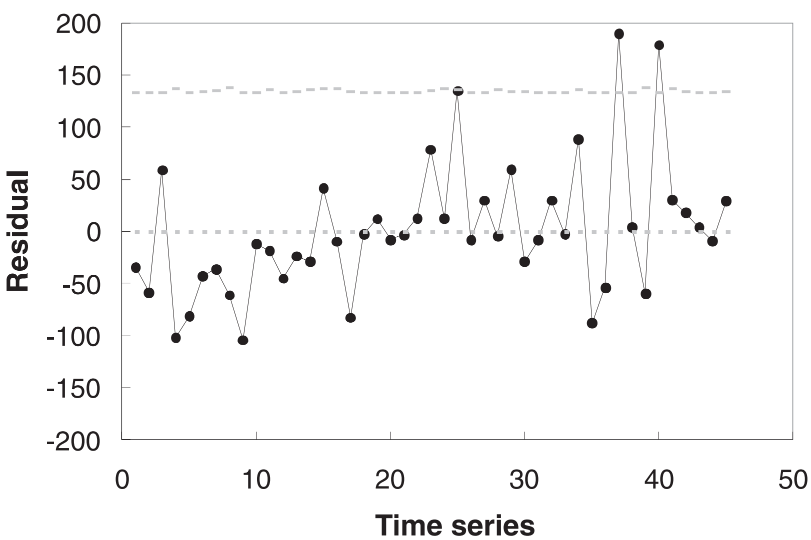

Figure 5. — Residual plots from an auto regression model with 95 percent upper confidence prediction limits for poplar, all months.

What was surprising was that oak had a smaller residual (80 tons) despite the expected increase in error due to chip contamination. However, after some investigation, this decreased variability was probably attributable to the fact that the mill consistently saw higher mean log diameters for oak than for poplar or pine, resulting in a decrease in BFTC error. In short, these control charts account for typical error associated with mixing of chips, which is important if a mill expects to distinguish between an outlying point attributable to contamination versus some outlier attributable to a process change.

What was encouraging to see was that oak and poplar models possessed strong r^2 values equal to pine, despite the added error that probably occurred for poplar and oak as a result of alternative species contamination. As a result, this supports that mixing of chips was probably occasional and had relatively little impact on the predictive ability of the models (**Table 1**). This error just punishes the sawmill by increasing the 95 percent confidence interval residual, making it more difficult to detect process changes.

As discussed earlier, OLS models may yield false confidence limits when autocorrelation is present. As a result, the models from **Table 1** were used to generate residual plots to account for the presence of autocorrelation. These models give better representative confidence limits to evaluate when the process mean deviates from the normal or when a process change occurs.

Figure 3 shows the residual plot versus time series for pine where each point was the next month in which production of pine occurred. One outlier occurred and represented the month of January 2002. When residual plots from an OLS model were used, two outliers appeared to occur and thus one possible false alarm was avoided. **Figure 4** demonstrates the residual plot versus time series for poplar where residuals from the AUTO model greater than 18 months was used. One outlier was observed with another very near outlying value. The outlier occurred during the month of February 2002, the month after the outlier occurred for pine. It is probable that the same process error influenced both data points for two different species. The month that the other "near outlier" occurred did not exhibit any unusual data points for pine or oak.

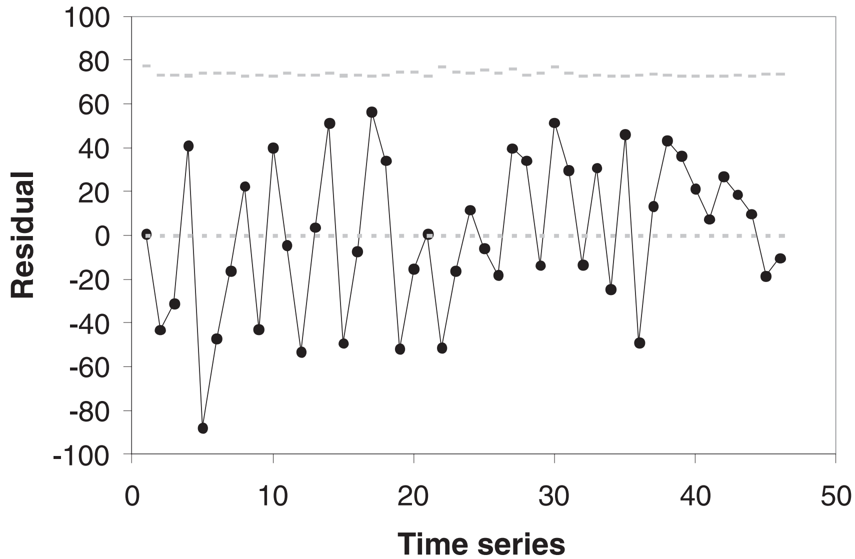

Figure 6. — Residual plots from an auto regression model with 95 percent upper confidence prediction limits for mixed oak.

Figure 5 shows the residual plot for all months of poplar production and demonstrates how a change in process makes more samples appear as outliers. The shift in process was established to occur around the 18th month since most of the residuals before the 18th month fell below zero, while many residuals were above zero for months greater than 18. The other notable occurrence was that three outliers occurred with the inadequate model (**Fig. 5**), while only two of the three points fell near the confidence limits in **Figure 4**. For oak, a distinct negative autocorrelation could be seen, given the "sawblade" pattern of the residuals around zero (**Fig. 6**). Around time series 38, this pattern changed to many residuals consecutively falling above zero. This suggests that a process change occurred and a new model would have been useful. A model was not developed for this segment since only a few data points were available.

It should be noted that these data were presented on a monthly basis but could have been performed on a weekly basis since it takes less than a week for the pulpmill to report weight data. However, should a small sawmill be able to afford a weight scale, then a daily account of BFTC could be made. It is probable that analyzing the data on a monthly basis would not be quick enough to detect and repair process problems or detect process changes. However, success at detecting process changes at the monthly level should encourage mills to collect data on a more frequent basis.

Process problems were considered temporary outliers attributable to some inefficiency in the manufacturing process. Process changes were considered to occur when the negative autocorrelation switched to a false positive autocorrelation, i.e., the residuals begin to consecutively fall either above or below zero. During a process change, the reason needs to be identified and remedied. If the process change is reasonable and acceptable, then a new model is needed to generate more accurate residuals.

Conclusions

BFTC ratio did not follow a normal distribution for any of the three major species processed at this sawmill. Furthermore, due to the autocorrelative nature between consecutive months, the assumption of independence was violated. Violation of this assumption suggests that traditional control charting such as the Stewhart chart may be invalid. As a result, a different measure was needed to assess when an outlier, and hence a potential problem in manufacturing, occurred.

As an alternative to traditional control charting, board foot was plotted against chip tons for each species and upper 95 percent confidence limits were determined and plotted in a residual plot. Any residual to exceed the upper confidence limit suggests that the process was out of control and a remedy was needed. Alternatively, consecutive positive or negative residuals suggest a process change and may justify creating a new model for the new process if correction to the process was not possible.

When board foot versus ton of chips was plotted, all three species exhibited similar slopes. Differences did occur in the intercept, with the higher intercepts indicating a more inefficient process. This was demonstrated with the introduction of a new sawyer who tended to focus on poplar production. The intercept, with the addition of the new sawyer, significantly rose while the slope stayed the same. Also, the variation of the residuals for the new model increased with the introduction of a second sawyer.

Residuals from AUTO models did not exceed the upper confidence boundaries as often as residuals plotted from OLS models, suggesting that false effort could be placed in repairing a manufacturing process with OLS-based control charting. Two residuals from separate processes, pine versus poplar production, were outliers at almost the exact same time period (January and February), which suggests that a process difference independent of species occurred. This justifies separation of residual plots by species and/or product, which provide more information about when or where the errors in the process might have occurred. Oak did not exhibit any outliers, although a slight process change appeared to occur toward the end of the time series, as indicated by a string of consecutive positive residuals.

Given the success of control charting on a monthly basis, these results would justify a mill acquiring data on a daily or weekly basis. If this decision was made, one would have to investigate the time of lag of the error term, i.e., how many equal time periods must occur before two samples can be considered independent of one another. Accurate measurement of board foot and ton of chips at the point of production would reduce or even eliminate the autocorrelation attributable to selling chips and lumber in two separate months.

Finally, it should be noted that the BFTC ratio should not supercede the profitability of the mill. Instead, it should be a supplemental tool for quality control personnel. Clearly, if the BFTC ratio is undesirable, but if profits are acceptable, then perhaps a temporary change in raw material quality occurred. As a result, it would be advisable to consider the selling value of the lumber and chips in conjunction with the BFTC ratio.

Literature cited

Atienza, O.O., L.C. Tang, and B.W. Ang. 1998. A SPC procedure for detecting level shifts of autocorrelated processes. J. of Quality Tech. 30(4):340-351.

Cook, D.F. 1992. Statistical process control for continuous forest products manufacturing operations. Forest Prod. J. 42(7/8):47-53.

__________ and J.G. Massey. 1990. Exponentially weighted moving average control chart. Tappi J. 73(6):247-248.

Cryer, J.D. and T.P. Ryan. 1990. The estimation of sigma for an X chart: MR/d2 or S/c4? J. of Quality Tech. 22(3):187-192.

Hunter, J.S. 1986. The exponentially weighted moving average. J. of Quality Tech. 18(4): 203-210.

Johnson, M., D. Latour, E. Routten, and J. Brocklebank. 1994. Principles of regression analysis. SAS Institute Inc., Cary, NC. 655 pp.

Keegan, C.E., D.P. Wichman, K.A. Blatner, D.D. Van Hooser, and S.A. Willits. 1998. Mill residue volume factor changes in Idaho and Montana. Forest Prod. J. 48(3):73-75.

Montgomery, D.C. and C.M. Mastrangelo. 1991. Some statistical process control methods for autocorrelated data. J. of Quality Tech. 23(3):179-193.

Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman. 1996. Applied Linear Statistical Models. WCB McGraw-Hill, Boston, MA. 1,408 pp.

SAS Institute. 2001. SAS/STAT software. Version 8. SAS Institute Inc., Cary, NC.

Savsar, M. and P.C. Kersavage. 1982. Mathematical model for determining the quantity of materials produced in sawmilling. Forest Prod. J. 32(11/12):35-38.

Steele, P.H. and C.D. Risbrudt. 1985. Efficiency of softwood sawmills in the southern United States in relation to capacity. Forest Prod. J. 35(7):51-56.

Wade, M.W., S.H. Bullard, P.H. Steele, and P.A. Araman. 1992. Estimating hardwood sawmill conversion efficiency based on sawing machine and log characteristics. Forest Prod. J. 42(11/12):21-26.

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 Prod. J. 51(11/12):70-77.