Predicting and Correlating the Strength Properties of Wood Composite Process Parameters by Use of Boosted Regression Tree Models

Dillon M. Carty   Timothy M. Young   Russell L. Zaretzki
Frank M. Guess   Alexander Petutschnigg

Abstract

Predictive boosted regression tree (BRT) models were developed to predict modulus of rupture (MOR) and internal bond (IB) for a US particleboard manufacturer. The temporal process data consisted of 4,307 records and spanned the time frame from March 2009 to June 2010. This study builds on previous published research by developing BRT models across all product types of MOR and IB produced by the particleboard manufacturer. A total of 189 continuous variables from the process line were used as possible predictor variables. BRT model comparisons were made using the root mean squared error for prediction (RMSEP) and the RMSEP relative to the mean of the response variable as a percent (RMSEP%) for the validation data sets. For MOR, RMSEP values ranged from 1.051 to 1.443 MPa, and RMSEP% values ranged from 8.5 to 11.6 percent. For IB, RMSEP values ranged from 0.074 to 0.108 MPa, and RMSEP% values ranged from 12.7 to 18.6 percent. BRT models for MOR and IB predicted better than respective regression tree models without boosting. For MOR, key predictors in the BRT models were related to ''pressing temperature zones,'' ''thickness of pressing,'' and ''pressing pressure.'' For IB, key predictors in the BRT models were related to ''thickness of pressing.'' The BRT predictive models offer manufacturers an opportunity to improve the understanding of processes and be more predictive in the outcomes of product quality attributes. This may help manufacturers reduce rework and scrap and also improve production efficiencies by avoiding unnecessarily high operating targets.

The forest products industry is an important contributor to the US economy; its products account for approximately 4 percent of the total US manufacturing gross domestic product, placing it on par with the automotive and plastics industries (American Forest and Paper Association 2014). The industry generates more than $210 billion per year in sales and employs approximately 900,000 people earning $50 billion in annual payroll, making it one of the top 10 manufacturing employers across 42 states (American Forest and Paper Association 2014).

Sustaining business competitiveness by reducing costs and maintaining product quality is essential for this industry. A key challenge facing this industry is to continually improve its understanding of process variables and their relationship with final product quality attributes. Quantifying the relationships between process variables (line speed, press temperature, etc.) and final product quality attributes (internal bond [IB], modulus of rupture [MOR], etc.) while predicting strength properties was the goal of this study.

MOR in bending is the maximum fiber stress at failure. Tensile strength perpendicular to the surface is used as a measure of the IB of particleboard.

The delay between the time at which a test sample is taken at the output end of the production line and the time at

The authors are, respectively, former Graduate Research Assistant, Dept. of Statistics, Operations, and Management Sci. (dcarty1@utk.edu), Professor and Graduate Director, Inst. of Agric., Center for Renewable Carbon, Dept. of Forestry, Wildlife and Fisheries (tmyoung1@utk.edu [corresponding author]), and Associate Professor and Professor, Dept. of Statistics, Operations, and Management Sci. (rzaretzk@utk.edu, fguess@utk.edu), Univ. of Tennessee, Knoxville; and Professor and Head, Salzburg Univ. of Applied Sci., Kuchl, Austria (alexander.petutschnigg@fh-salzburg.ac.at). This paper was received for publication in August 2012. Article no. 12-00085.

which the strength characteristics of this sample (e.g., IB) have been determined in a testing laboratory is also important. This delay can be as long as 1 to 3 hours in particleboard. In the absence of a real-time model that predicts mechanical properties, it is difficult to optimize production and correct for possible poor mechanical properties of the final manufactured product. Boosted regression tree (BRT) models in real-time settings may offer wood composite manufacturers a competitive advantage for improving production efficiency, avoiding waste, and avoiding higher-than-necessary operating targets.

A host of regression techniques offers a means for explorative correlation analyses and possible prediction of product properties to overcome this shortcoming of time gaps between destructive samples from the production line. The regression techniques are well documented in the published literature (Young 1996; Cook and Chiu 1997; Bernardy and Scherff 1998; Greubel 1999; Cook et al. 2000; Eriksson et al. 2000; Young and Guess 2002; Sjöblom et al. 2004; Young et al. 2004, 2008, 2014; Lei et al. 2005; Xing et al. 2007; André et al. 2008; Clapp et al. 2008; Mora and Schimleck 2010; André and Young 2013; Riegler et al. 2013).

BRTs constitute a data mining technique that has had considerable success in predictive modeling. As Schonlan (2005) noted, boosting is a highly flexible regression tree method that allows the researcher to specify the independent variables without specifying the functional relationship to the response. Schonlan (2005) further notes that this flexibility of BRT models will tend to fit better than a linear model, resulting in improved inference and creditability. The BRT technique draws on insights and techniques from both statistical and machine learning traditions to enhance the predictive power of regression trees (Hastie et al. 2009). Improved prediction using BRT models can help minimize the risk of producing hours of defective or off-grade product or hours of production that are unnecessarily overengineered and of higher cost. Predictive models using BRT can also reduce the costs associated with rework (i.e., remanufactured panels due to poor strength properties), reduce feedstock costs (e.g., resin and wood), reduce energy usage, and improve wood utilization from the valuable forest resource. Improving production efficiency and overall business competitiveness for the wood composites industry was the rationale and motivation to support this work. The authors are not aware of literature documenting the use of BRT in wood composites manufacture.

''Trees''[1] or regression trees are the fundamental precursor to BRTs, and the methodologies have roots in both statistics and computer science. A precursor to current tree methodology was CHAID, developed by Morgan and Sonquist (1963).[2] Breiman et al. (1984) first introduced the

main ideas of tree methodology to statistics. Hastie et al. (2009) also described decision trees from a statistical perspective. In their most fundamental form, tree-based methods partition the predictor space into rectangular regions using a series of rules to identify regions having the most homogeneous responses to the predictors and then fit a constant or a relatively simple regression model to the data in each partition (Bishop 2006, Loh 2008). The growing of a tree involves recursive binary splits, meaning that a binary split is continually applied to its own output until a stopping specification is obtained. Decision trees are well liked because they represent information that is innate and easy to visualize and identify prioritized interactions.

A regression tree is a piecewise linear estimate of a regression function that is constructed by the recursive partitioning of the data and the sample space (Loh 2002). The construction of a regression tree generally consists of the following four steps performed iteratively: (1) partition the data, (2) fit a model to the data after each partition, (3) stop when the residuals of the model are approximately zero or when there are only a few observations left, and (4) prune the tree (i.e., if the tree overfits). Even though the two-dimensional hierarchical interactions displayed by regression trees provide very good explanatory value, a limitation of regression trees without any type of boosting is poor predictive power (Hastie et al. 2009).

Boosting is a technique used to enhance the predictive performance of regression trees. As Elith et al. (2008) noted in citing Schapire (2003), ''Succinctly, boosting is a method for improving the accuracy of a model, based on the simple idea that it is easier to find and average many rough rules of thumb, than to find a single, highly accurate prediction rule.'' A weak learner (also known as a base classifier or weak classifier) is one whose error rate is only slightly better than random guessing. The main point of boosting is to sequentially apply the weak learning algorithm to repeatedly modified versions of the data, hence creating a sequence of weak learners. In boosting, models are fit iteratively to the training data using methods to increase emphasis on observations that are modeled poorly by the existing collection of models (Elith et al. 2008).

The original design for boosting made it specific to classification problems, but it can be ''profitably extended to regression'' as well (Hastie et al. 2009). As related to regression problems, boosting is a form of functional gradient descent (Elith et al. 2008). Take a loss function that represents the loss in predictive performance owing to a suboptimal model. Boosting is a numerical optimization technique for minimizing the loss function by adding, at each step, a new model (e.g., a regression tree) that best reduces, or steps down, the gradient of the loss function (Elith et al. 2008). According to Elith et al. (2008), in a BRT, the initial regression tree is the one that reduces the loss function the most. At each iteration, the focus is on the residuals and root mean square error of prediction (RMSEP) reduction.[3]

---

[1] ''Tree'' or ''trees'' in this article refers to the methodology of decision trees, of which regression trees are a subset when modeling measurement data. A regression tree is a piecewise constant or piecewise linear estimate of a regression function, constructed by recursively partitioning the data and sample space. Its name derives from the practice of displaying the partitions as a decision tree, from which the significance of the regressors is inferred.

[2] Refer to Quinlan (1993) for an overview of how tree methodology was developed in machine learning.

---

[3] $\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$, where y_i is the observed value, $\hat{y}_i$ is the predicted value, and n is the total number of records in the validation data set. $\text{RMSEP\%} = (\text{RMSEP}/\hat{\mu}_y) \times 100\%$, where $\hat{\mu}_y$ is the mean of the observed response values in the validation data set. The use of $\hat{\mu}_y$ allowed for easier comparison of each BRT model to a similar baseline.

Also according to Elith et al. (2008), in the second step, a regression tree, which can contain different variables and split points from the first tree, is fit to the prediction residuals of the first tree. The overall model now contains two trees, and the residuals from this two-term model are estimated. The process is stagewise, i.e., existing trees are left unchanged as the model grows increasingly larger. Only the fitted value for each observation is reestimated at each step to reflect the contribution of the newly added tree. In the end, the final BRT model is a linear combination of numerous trees and can be thought of as a regression model with each term being a tree.

This study is aligned with Gleser's (1996) "First Law of Applied Statistics" principle that two individuals using the same statistical method on the same data should arrive at the same conclusion. The objectives of this study were (1) to quantify the correlations of strength properties of particleboard and process parameters from a manufacturing data set by use of BRTs and (2) to assess the predictions of strength properties of particleboard by use of BRTs. The specific method of stochastic gradient boosting was used in this study to model the strength properties of particleboard.

Methods

Stochastic gradient boosting

Stochastic gradient boosting is one loss function algorithm for BRT and was used in this study. Friedman (2002) stated that "gradient boosting constructs additive regression models by sequentially fitting a simple parameterized function (i.e., base learner) to current 'pseudo'-residuals by least-squares at each iteration."

In the function estimation problem, one has a response variable y and a set of random explanatory values. $X = \{x_1, \ldots, x_n\}$ Friedman (2002) noted that given a training sample $\{y_i, X_i\}_1^N$ of known (y, X) values, the objective is to find a function $F^*(x)$ that maps X onto y such that over the joint distribution of all (y, X) values, the expected value of some loss function $\Psi(y, F(x))$ is minimized:

$$F^*(X) = \underset{F(X)}{\arg \min} \, E_{y,X} \Psi\left(y, F(X)\right) \qquad (1)$$

Boosting approximates $F^*(x)$ by an additive expansion of the form

$$F(X) = \sum_{m=0}^{M} \beta_m h(X; a_m) \qquad (2)$$

where functions $h(X;a)$ (i.e., "base learner") are generally simple functions of X with parameters $a = \{a_1, a_2, \ldots\}$ In a forward stagewise manner, the expansion coefficients β_m and the parameters a_m are jointly fit to the training data. According to Friedman (2002), one starts with a preliminary guess $F_0(X)$, and then for $m = 1, 2, \ldots, M$ iterations,

$$(\beta_m, a_m) = \underset{\beta, a}{\arg \min} \sum_{i=1}^{N} \Psi\left[y_i, F_{m-1}\left(X_i + \beta h(X; a)\right)\right] \quad (3)$$

and

$$F_m(X) = F_{m-1}(X) + \beta_m(X; a_m) \qquad (4)$$

where y_i is the response of interest and $F_m(X)$ is the final BRT.

Software and learning parameters used for the study

Statistica 10 (http://www.statsoft.com) software was used in this study to estimate the BRT models. The BRT algorithm of Statistica 10 is a "full featured implementation of the stochastic gradient boosting method." Five key parameters used to control the stochastic gradient boosting algorithm were manipulated in the BRT analysis. These criteria represented the optimization criteria for the BRT models:

- First, the "learning rate," or the shrinkage parameter (lr), specified the specific weight with which consecutive simple regression trees are added into the prediction equation; that is, lr specified the shrinkage applied to each tree in the final BRT model (Elith et al. 2008). For example, a BRT model with 500 trees fitted and with $lr = 0.01$ will produce predictions that are the sum of predictions from each of the 500 trees multiplied by 0.01 (compare, e.g., Figs. 1 and 2, where lr varies from 0.005 to 0.5).

- Second, the "number of additive terms" (nat) specified the number of simple regression trees (i.e., additive terms) to be computed in successive boosting steps. According to Elith et al. (2008), a smaller lr and larger nat are preferable. Because smaller values for lr (i.e., more shrinkage) result in larger training risk for the same nat, both lr and nat control the prediction risk on the training data. We used nat values of 100, 200, 300, 400, 500, 600, and 1,000, respectively, given prior experience from preliminary runs using the stochastic gradient boosting. Training risk minimizes the average value of the loss function on the training set. Training risk is necessary for BRT models to avoid model overfitting, which is a consequence of regression trees or BRT, which are known to be "greedy" algorithms.

- Third, the "maximum number of nodes" (mnn) specified the maximum number of nodes allowed for each individual tree in the boosting sequence. This is a stopping parameter in the sense that each time a parent node is split, the total number of nodes in the tree is examined, and the splitting is stopped if this number exceeds the number specified by mnn. Stopping parameters are a key element of regression trees in general in that tree algorithms will tend to overfit the model. The stopping parameter defines the ending point of tree. Setting $mnn = 3$ produces BRFT models with only main effects. Setting $mnn = 5$ produced models with main effects and two-variable interactions. In this study, $mnn = 3$ and 5 was used.[4]

- Fourth, the "subsample proportion" (sp) was used for selecting the random learning sample for consecutive boosting steps. Given two prior works documented in the literature (Elith et al. 2008, Hastie et al. 2009), $sp = 0.5$ was used in this study. This literature suggested a balanced $sp = 0.5$ to avoid overfitting of the fraction subsample, which may occur if sp is too great.

One hundred forty BRT models were the product of testing 10 different levels of lr values (ranging from 0.005 to 0.5), two levels of mnn values (3 and 5), and seven different

---

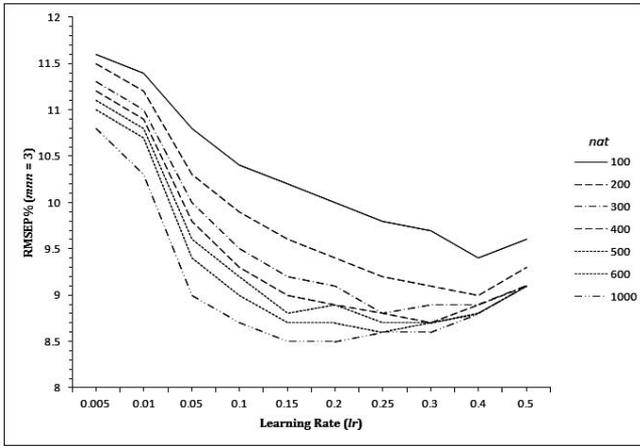[4] As noted by Elith et al. (2008), BRT regularization involved jointly optimizing nat, lr, and mnn.

Figure 1.—The relationship between learning rate (lr) and modulus of rupture (MOR) root mean squared error for prediction (RMSEP) and the RMSEP relative to the mean of the response variable as a percent (RMSEP%) for the 70 different boosted regression tree models with seven values chosen for the number of additive terms (nat) and a value of 3 chosen for the maximum number of nodes (mnn).
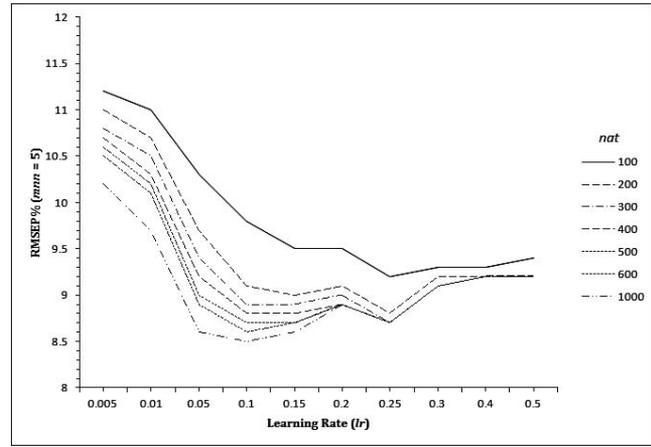

Figure 2.—The relationship between learning rate (lr) and modulus of rupture (MOR) root mean squared error for prediction (RMSEP) and the RMSEP relative to the mean of the response variable as a percent (RMSEP%) for the 70 different boosted regression tree models with seven values chosen for the number of additive terms (nat) and a value of 5 chosen for the maximum number of nodes (mnn).

levels of nat values (ranging from 100 to 1,000). The parameter settings for the best BRT model for MOR were $lr = 0.15$, $mnn = 3$, and $nat = 1,000$. Importantly, the optimal number of trees obtained for these 1,000 iterations was 943 three-node trees (i.e., the smallest average squared error for the validation sample was obtained at 943 trees for these 1,000 boosting steps).

In the BRT model development, 3,449 (80% of the entire data set) randomly selected records were used for the training models, and 858 (remaining 20%) were used for validation. To compare the predictive abilities of the BRT models, RMSEP and RMSEP% were used as performance measures in validation.

Data set

A time-ordered data set was obtained from a particleboard manufacturer with a continuous press in the United States. The key quality strength metrics, or response variables for this manufacturer's product, were MOR and IB. As is typical for particleboard manufacturing, destructive test samples were taken from the production at irregular time intervals that varied from 1 to 2 hours or as product type changed. The data set consisted of 4,307 records that spanned the time period from March 2009 to June 2010. There were an equal number of samples for MOR and IB as specified by the manufacturer's sampling protocol. There were 189 possible continuous predictor variables. The predictor variables in this study represented mat forming, mat weight, mat temperature, line speed, pressing temperatures, pressing pressures, etc. Specific detail concerning the predictor variables is not possible given the terms of a confidentiality agreement with the manufacturer. The data of the predictor variables represented a fused data set from a study by Young et al. (2014). The fused data set of strength properties and process parameters accounted for appropriate time lags in the process parameters as related to the time of destructive testing sampling, i.e., fiber passes under sensors of the process at different times relative to the time the panels exits the press. Process parameter data were obtained from the process data warehouse and represented a median value of 25 programmable logic controllers (PLC) data samples. The PLC data were sampled from the sensors every 5 seconds. Outliers that represented sensor failure or production line downtime were not included in the data samples.

There were 118 different particleboard product types manufactured by the producer within the 4,307 records. Product types were not differentiated in the overall BRT model predictions of MOR and IB. A holistic BRT model for all product types seems more plausible and practical for the manufacturer than attempting to develop distinct models for 118 different product types with different record lengths.

Results and Discussion

MOR predictions

Figures 1 and 2 illustrate the range of RMSEP% values for MOR resulting from BRT models with various combinations of BRT modeling parameters. The RMSEP for this BRT model of MOR was 1.051 MPa. The RMSEP% for this BRT model of MOR was 8.5 percent. The lowest RMSEP% occurred for the largest number of regression trees ($nat = 1,000$) using three- and five-node trees ($mnn = 3$, $mnn = 5$). For three-node trees, convergence occurred at an lr of 0.15 and diverged at $lr > 0.15$ for the regression trees of $nat = 1,000$. Five-node trees tended to converge at $lr = 0.5$. As with the three-node tree, $nat = 1,000$ had the lowest RMSEP%. The results illustrated in Figures 1 and 2 appear to support the BRT concept that "combining many weak learners forms a stronger one for prediction," as noted by Schapire (2003). Figures 1 and 2 for MOR also illustrate that larger trees outperform (lower RMSEP%) for $lr < 0.2$ where $mnn = 3$ and for $lr < 0.1$ where $mnn = 5$. The correlation between the observed values and the predicted values in the validation data set was $r = 0.91$, and the XY scatterplot does not reveal any inherent bias in the validation set (Fig. 3). Two distinct groupings in Figure 3 reveal the strength of MOR targets by manufacturer product types.
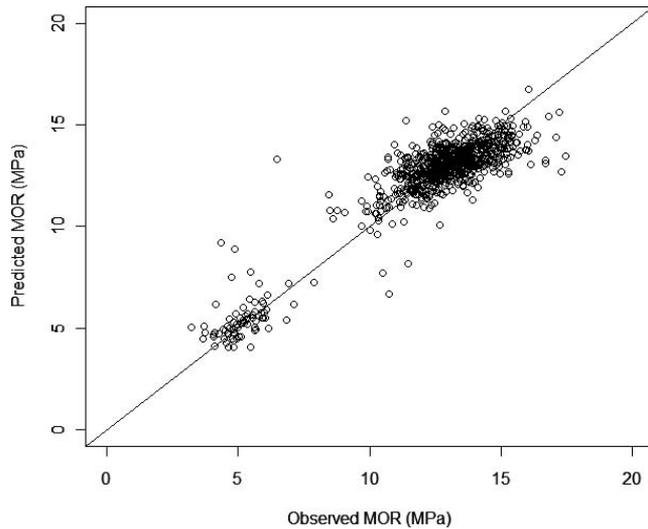
Figure 3.—Scatterplot of the observed values versus the predicted values for the boosted regression tree model that best predicts modulus of rupture (MOR).

More detail on product types is not possible given the confidentiality agreement of study.

For MOR, key predictors in the models related to particleboard were dominated by press-related parameters such as "pressing temperature zones within the press," "thickness of pressing by press zone," and "pressing pressure by press zone." This may be the result of the influence of a continuous press in manufacturing particleboard. There were multiple heating, pressure, and thickness zones in a series common to most continuous presses. Manufacturers with continuous presses have confidential pressing strategies as related to the type, brand, and length of the continuous press. More detail on predictor variables is not presented given the confidentiality agreement with the manufacturer.

For the sake of comparison, a regression tree model without boosting was fit to the same data. The regression
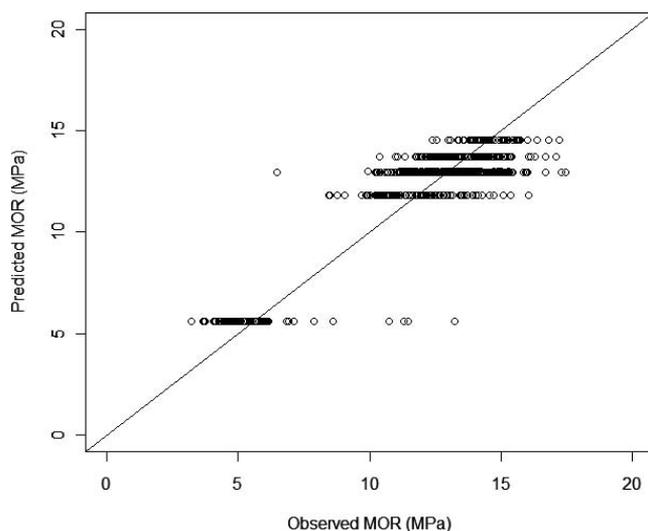

Figure 4.—Scatterplot of the observed values versus the predicted values for the regression tree model without boosting that predicts modulus of rupture (MOR).
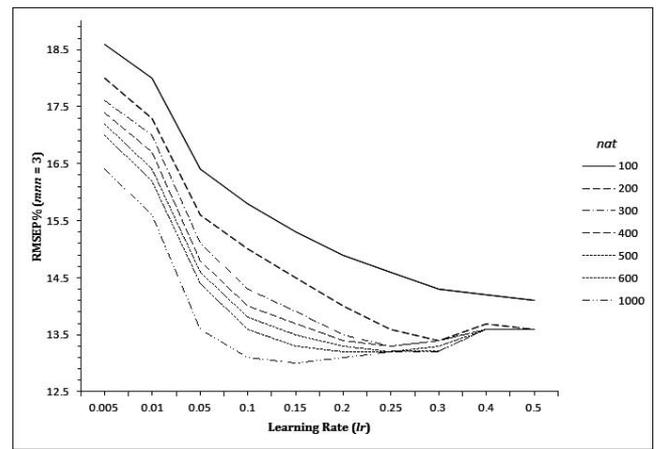

Figure 5.—The relationship between learning rate (lr) and internal bond root mean squared error for prediction relative to the mean of the response variable as a percent (RMSEP%) for the 70 different boosted regression tree models with seven values chosen for the number of additive terms (nat) and a value of 3 chosen for the maximum number of nodes (mnn).

tree without boosting had a higher RMSEP (1.263 MPa) and RMSEP% (10.2%).[5] The $r = 0.87$ between the observed values and the predicted values in validation was quite high but was misleading given the XY scatterplot of these values (Fig. 4). The predictive weakness of regression trees without boosting (i.e., fitting a mean, etc., to each binary split) is shown in Figure 4.

Predictions of IB

The loss function of the RMSEP% for IB using the same BRT modeling parameter combinations discussed above are given in Figures 5 and 6. The BRT parameter settings for the lowest RMSEP% were $lr = 0.1$, $mnn = 5$, and $nat = 1,000$. The optimal number of trees obtained for these 1,000 iterations was 957. This again strengthens the argument by Schapire (2003) that combining many weak learners has better predictability relative to one tree. Convergence for the IB strength property was similar to MOR and occurred for the largest number of regression trees ($nat = 1,000$) using three- and five-node trees ($mnn = 3$, $mnn = 5$). However, relative to MOR and IB, convergence occurred at a lower lr of 0.10 for $mnn = 3$ and an lr of 0.05 for $mnn = 5$. Figures 5 and 6 for MOR further illustrate that larger trees provide the best predictability.

IB was more difficult to predict than MOR. IB had an RMSEP = 0.074 MPa and an RMSEP% = 12.7 percent. The correlation between the observed values and the predicted values in validation was $r = 0.86$ and did not have bias (Fig. 7). The IB results using the BRT method have higher overall RMSEP in validation compared with the studies by Riegler et al. (2013) for high-density fiberboard (HDF) and André et al. (2008) for medium-density fiberboard (MDF), where principal components analysis and partial least squares methods were used. However, the overall data sets of these studies for HDF and MDF were approximately one-tenth the size of the data set of this study. The validation data set of

---

[5] The independent variables in the regression tree model in Figure 4 were nondescriptive given the confidentiality requests from the company supplying the data.
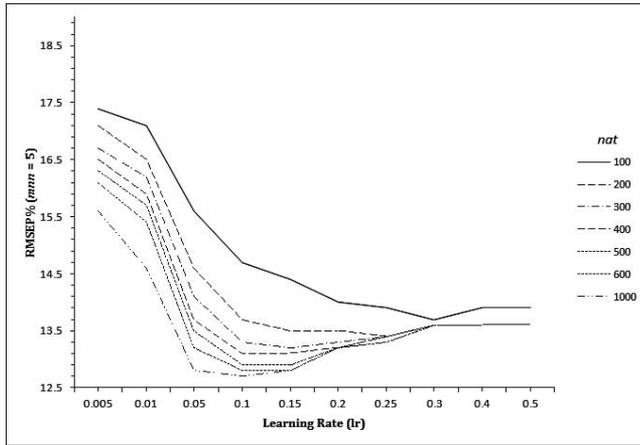
Figure 6.—The relationship between learning rate (lr) and internal bond root mean squared error for prediction relative to the mean of the response variable as a percent (RMSEP%) for the 70 different boosted regression tree models with seven values chosen for the number of additive terms (nat) and a value of 5 chosen for the maximum number of nodes (mnn).

Riegler et al. (2013) was extremely small, and this may affect the overall repeatability and robustness of his models. The coefficient of variation (CV; 7.1%) for the HDF IB data of Riegler et al. (2013) was also substantially smaller than the CV (24.6%) for the IB data from the manufacturer of this study, indicating that the particleboard process of this study had more inherent natural variation than the HDF mill used by Riegler et al. (2013). The validation data set of André et al. (2008) was also substantially smaller than that of this study.

For IB, the predictors occurring most often in the BRT models were related to ''thickness of pressing'' in the zones of continuous press. As with MOR, more detail on predictor variables is not possible given the confidentiality agreement with the manufacturer.

BRT models predicted MOR more accurately than IB, which had more inherent process variation (CV = 24.6%)
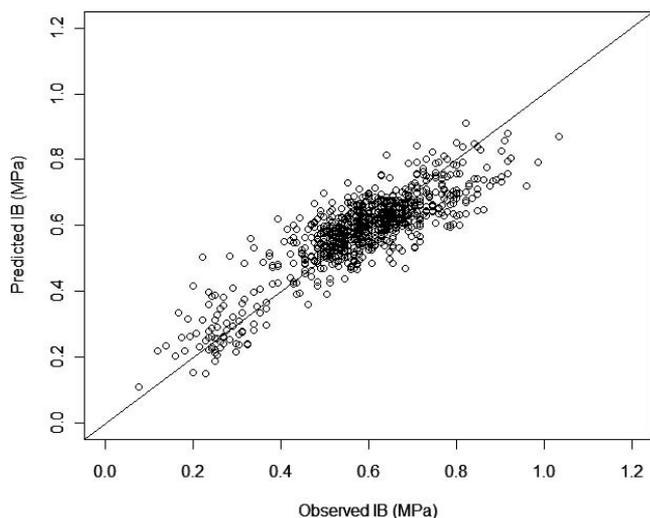

Figure 7.—Scatterplot of the observed values versus the predicted values for the boosted regression tree model that best predicts internal bond (IB).

for the particleboard manufacturer than did MOR (CV = 19.8%). This may represent one explanation for the difficulty in predicting IB using BRT models. The higher natural variation of IB relative to MOR may also be inherent to the higher potential for error associated with the standard test method for IB (ASTM International 2013); great care must be taken with regard to proper IB test sample preparation, the conditions of the blocks, the bonding quality between the blocks, etc. However, sample error data were not available in the data set, and there was nothing from personal observation of the destructive testing process to indicate that the particleboard manufacturer had an atypical IB testing methodology.

A key result of this study was the ability to use BRT models for predictive models for all 118 product types produced by the manufacturer across 16 months of production data ($n = 4,307$). Earlier modeling studies conducted for strength quality metrics of wood composites used much smaller training data sets and did not model multiple product types (André et al. 2008, Clapp et al. 2008). As computational power inevitably increases in the manufacturing sector, it is highly likely that future, expansive applications of BRT modeling will increase in application.

Conclusions

BRTs are a relatively new predictive modeling technique that draw on insights and techniques from both statistical and machine learning traditions by combining boosting algorithms with regression trees. BRT models have the ability to select pertinent variables, fit accurate functions, and model interactions. In this study, the boosting regression tree approach had better predictability in validation compared with regression tree methods without boosting.

As documented in this study, BRT predictive models may offer manufacturers a valuable tool for predicting product quality in real time, and this may in turn lower manufacturing costs by avoiding scrap and improving operational efficiency. BRT models in real-time settings may further help manufacturers avoid higher-than-necessary operating targets given improved predictability of final strength properties.

Literature Cited

American Forest and Paper Association. 2014. 2014 AF&PA sustainability report. http://www.afandpa.org/docs/default-source/one-pagers/-2014-sustainability-report.pdf. Accessed March 2, 2015.

André, N., H.-W. Cho, S. H. Baek, M.-K. Jeong, and T. M. Young. 2008. Enhanced prediction of internal bond strength in a medium density fiberboard process using multivariate statistical methods and variable selection. *Wood Sci. Technol.* 42(7):521–534.

André, N. and T. M. Young. 2013. Real-time process modeling of particleboard manufacture using variable selection and regression methods ensemble. *Eur. J. Wood Wood Prod.* 71(3):361–370.

ASTM International. 2013. Standard test method for surface bond strength of wood-base fiber and particle panel materials. ASTM D5651-13. ASTM International, West Conshohocken, Pennsylvania.

Bernardy, G. and B. Scherff. 1998. Saving costs with process control engineering and statistical process optimization: Uses for production managers, technologists and operators. *In:* Proceedings of the Second European Panel Products Symposium (EPPS), October 21–22, 1998, Llandudno, Wales, UK; BioComposites Centre, UWB, Bangor, Gwynedd, UK. pp. 95–106.

Bishop, C. M. 2006. Pattern Recognition and Machine Learning. Springer, New York. 729 pp.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984.

Classification and Regression Trees. Wadsworth International Group, Belmont, California. 276 pp.

Clapp, N. E., Jr., T. M. Young, and F. M. Guess. 2008. Predictive modeling the internal bond of medium density fiberboard using a modified principal components analysis. *Forest Prod. J.* 58(4):49–55.

Cook, D. F. and C. C. Chiu. 1997. Predicting the internal bond strength of particleboard utilizing a radial basis function neural network. *Eng. Appl. Artif. Intell.* 10(2):171–177.

Cook, D. F., C. T. Ragsdale, and R. L. Major. 2000. Combining a neural network with a genetic algorithm for process parameter optimization. *Eng. Appl. Artif. Intell.* 13:391–396

Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77:802–813.

Eriksson, L., P. Hagberg, E. Johansson, S. Rannar, O. Whelehan, A. Astrom, and T. Lindgren. 2000. Multivariate process monitoring of a newsprint mill: Application to modeling and predicting COD load resulting from de-inking of recycled paper. *J. Chemom.* 15:337–352.

Friedman, J. H. 2002. Stochastic gradient boosting. *Comput. Stat. Data Anal.* 38(4):367–378.

Gleser, L. J. 1996. Comment on ''bootstrap confidence intervals'' by T. J. DiCiccio and B. Enfron. *Stat. Sci.* 11:219–221.

Greubel, D. 1999. Practical experiences with a process simulation model in particleboard and MDF production. *In:* Proceedings of the Second European Wood-Based Panel Symposium, September 1999, Hannover, Germany. pp. 8–10.

Hastie, T., R. Tibshirani, and J. H. Friedman. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York. 319 pp.

Lei, Y. C., S. Y. Zhang, and Z. H. Jiang. 2005. Models for predicting lumber bending MOR and MOE based on tree and stand characteristics in black spruce. *Wood Sci. Technol.* 39(1):37–47.

Loh, W.-Y. 2002. Regression trees with unbiased variable selection and interaction detection. *Stat. Sin.* 12(2):361–386.

Loh, W.-Y. 2008. Classification and regression tree methods. *In:* Encyclopedia of Statistics in Quality and Reliability. F. Ruggeri, R. Kenett, and F. W. Faltin (Eds.). Wiley, New York. pp. 315–323.

Mora, C. R. and L. R. Schimleck. 2010. Kernel regression methods for the prediction of wood properties of *Pinus taeda* using near infrared spectroscopy. *Wood Sci. Technol.* 44(4):561–578.

Morgan, J. and J. Sonquist. 1963. Problems in the analysis of survey data, and a proposal. *J. Am. Stat. Assoc.* 58(302):415–434.

Quinlan, J. R. 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco. ISBN:1-55860-238-0. 241 pp.

Riegler, M., B. Spangl, M. Weigl, R. Wimmer, and U. Müller. 2013. Simulation of a real-time process adaptation in the manufacture of high-density fiberboards using multivariate regression analysis and feed forward control. *Wood Sci. Technol.* 47:1243–1259.

Schapire, R. E. 2003. The boosting approach to machine learning: An overview. *In:* MSRI Workshop on Nonlinear Estimation and Classification, 2002. D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, and B. Yu (Eds.). Springer, New York. pp. 113–141.

Schonlan, M. 2005. Boosted regression (boosting): An introductory tutorial and a Stata plugin. *Stata J.* 5(3):330–354.

Sjöblom, E., B. Johnsson, and H. Sundström. 2004. Optimization of particleboard production using NIR spectroscopy and multivariate techniques. *Forest Prod. J.* 54(6):71–75.

Xing, C., S. Y. Zhang, J. Deng, and S. Q. Wang. 2007. Investigation of the effects of bark fiber as core material and its resin content on three-layer MDF performance by response surface methodology. *Wood Sci. Technol.* 41(7):585–595.

Young, T. M. 1996. Process improvement through ''real-time'' statistical process control in MDF manufacture. *In:* Proceedings of Process and Business Technologies for the Forest Products Industry, July 29–31, 1996, Madison, Wisconsin; Forest Products Society, Madison, Wisconsin. pp. 50–51.

Young, T. M., N. André, and C. W. Huber. 2004. Predictive modeling of the internal bond of MDF using genetic algorithms with distributed data fusion. *In:* Proceedings of the Eighth European Panel Products Symposium, October 13–15, 2004, Llandudno, Wales, UK; BioComposites Centre, UWB, Bangor, Gwynedd, UK. pp. 45–59.

Young, T. M., N. E. Clapp, Jr., F. M. Guess, and C.-H. Chen. 2014. Predicting key reliability response with limited response data. *Qual. Eng.* 26(2):223–232.

Young, T. M. and F. M. Guess. 2002. Developing and mining higher quality information in automated relational databases for forest products manufacture. *Int. J. Reliab. Appl.* 3(4):155–164.

Young, T. M., L. B. Shaffer, F. M. Guess, H. Bensmail, and R. V. León. 2008. A comparison of multiple linear regression and quantile regression for modeling the internal bond of medium density fiberboard. *Forest Prod. J.* 58(4):39–48.