Developments in the Built Environment 27 (2026) 101018

Contents lists available at ScienceDirect

Developments in the Built Environment

journal homepage: www.sciencedirect.com/journal/developments-in-the-built-environment

Application of ensemble machine learning models for predicting ultimate tensile strength of trembling aspen (*Populus tremuloides*) dimension lumber

Dawei Wang *, Hoang D. Nguyen, Qipei Mei, Ying Hei Chui

Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, T6G 1H9, Canada

ARTICLE INFO

Keywords:
Machine learning
Ultimate tensile strength
Trembling aspen lumber
Ensemble models
GBRT
SHAP analysis

ABSTRACT

Trembling aspen (*Populus tremuloides*) dimension lumber is a promising product for light wood-frame construction, but its ultimate tensile strength (UTS) remains difficult to predict due to high variability from natural defects. This study developed four ensemble machine learning models, random forest, extreme gradient boosting, gradient boosting regression tree (GBRT), and adaptive boosting, to predict UTS of aspen lumber. The dataset covered 321 specimens with a 38 mm × 89 mm cross section and 12 input features, including grade, geometry, specific gravity, moisture content, knot characteristics, and modulus of elasticity (MOE) from machine stress rating (MSR) and longitudinal stress wave tests, all measurable in sawmills. GBRT achieved the best performance ($R^2 = 0.718$), with the minimum MSR-MOE as the most influential predictor. Against empirical models ($R^2 = 0.239$-0.487) and multivariable linear regression ($R^2 = 0.585$), GBRT improved R^2 by 0.133-0.479, supporting industry-oriented AI for individual-specimen UTS prediction and economical material use.

1. Introduction

Wood is a natural and renewable material that has been widely used in construction. Before steel and concrete became common, it was the main building material in North America (Brashaw and Bergman, 2021). Over time, wood has remained an important part of the forest products industry and continues to hold a strong position in the North American market. In Canada, the softwood lumber industry is a key part of the national forest economy, supporting harvesting, processing and forest management, making Canada the world's second-largest producer and the main supplier to the U.S., with exports valued at about 8.4 billion USD in 2020 (Natural Resources Canada, 2025).

Most lumber produced in North America is softwood dimension lumber, usually between 38 mm and 89 mm in depth, 38 mm and 286 mm in width, and up to around 5 m in length (Canadian Wood Council, 2019; Gong, 2019). Because of its high strength-to-weight ratio and workability, it remains the primary material for light wood-frame construction (Panshin and de Zeeuw, 1981). However, since wood is anisotropic, its mechanical properties vary with direction, which reveals wide variability in its anatomical and physical characteristics. For this reason, studies on wood performance often require large specimen sizes to capture this natural variation and to produce statistically reliable results (Zobel and Van Buijtenen, 2012; Gong, 2021). The mechanical properties of wood are affected by many factors, such as growing conditions, soil nutrients, species, specific gravity (SG), age, visible and hidden defects, harvesting methods and drying processes (Burdon and Thulin, 1966; Barrett and Lau, 1994; Boyle, 1998; Arnold and Mauseth, 1999; Yamamoto et al., 2013). In some cases, trees exposed to storms or insect attack develop internal micro-damage that is not visible but can lower lumber quality through crack propagation (Smith et al., 2003; Qiu et al., 2014). Such variations are mostly random and cannot be fully controlled in industrial production. Understanding the strength properties of lumber is therefore essential for ensuring the structural safety of buildings. In addition, if wood strength can be accurately predicted using practical methods, it would allow for more precise grading and better economic value in industrial applications.

Research on predicting the strength of wood and wood-based materials has developed over the past few decades from empirical observation, through multi-parameter fitting, to data-driven modelling, with steady improvements in reliability for grading and structural design. Early studies focused on understanding how local defects affect strength properties. Cramer and Goodman (1983, 1986) developed the KMESH1 finite element model, which simulated the effects of knots and slope of grain on the mechanical behaviour of lumber. By accounting for the redistribution of stress around knots and the local grain deviation, the

* Corresponding author.
E-mail address: dawei8@ualberta.ca (D. Wang).

https://doi.org/10.1016/j.dibe.2026.101018
Received 28 May 2026; Received in revised form 6 August 2026; Accepted 13 August 2026
Available online 14 August 2026

model captured the bending behaviour of edge-knotted lumber more closely than the strength reduction approach in ASTM D245 (American Society for Testing and Materials, 2025), which tended to underestimate the actual strength. This line of work clarified the mechanisms by which defects govern lumber strength and motivated later efforts to incorporate defect information into prediction models.

The development of non-destructive evaluation (NDE) introduced a new stage in predictive research. Kaiserlik and Pellerin (1977) first linked stress-wave velocity to wood strength, establishing the basis for using elastic wave propagation as an indirect indicator of mechanical performance. Subsequent studies extended this principle to full-size lumber through longitudinal vibration and ultrasonic methods, where the dynamic modulus of elasticity (MOE) derived from wave velocity and SG became a standard predictor in machine stress rating (MSR). Guntekin et al. (2013) reported that dynamic MOE provided a closer prediction of bending strength than static MOE for Turkish red pine (*Pinus brutia*), confirming the practical value of stress-wave measurements for rapid mechanical assessment of softwood lumber.

Beyond softwood bending applications, research on the ultimate tensile strength (UTS) of structural-size lumber has progressed along a similar trajectory, with multivariable statistical models gradually replacing single-predictor relationships. Building on dynamic testing principles, several studies have examined how combinations of stiffness, SG and defect-related parameters explain the variation in UTS of full-size specimens. Senalik et al. (2020) tested southern yellow pine dimension lumber and reported a coefficient of determination (R^2) of about 0.70 for UTS when SG and dynamic test results were combined in a multivariable linear regression (MLR). Briggert et al. (2020) reached a similar level of accuracy for Norway spruce (*Picea abies*) by combining dynamic MOE with scanned fibre direction, a parameter that reflects local grain deviation around knots. These results indicate that, for full-size lumber loaded in tension, MLR based on a small number of physically meaningful predictors can already reach a stable performance range, against which more flexible models should be compared.

While these predictor-based models describe specimen-to-specimen variation, other studies have focused on within-member variability and on more efficient estimation of multiple mechanical properties. Lam and Varoğlu (1991) developed a two-stage stochastic model for spruce-pine-fir (SPF) lumber to simulate parallel-to-grain UTS variation, showing that within-member variability plays an important role in structural performance. Lei et al. (2005) applied the seemingly unrelated regression method, a system of regression equations whose error terms are correlated, to jointly estimate MOE and modulus of rupture (MOR) of black spruce (*Picea mariana*), reducing standard errors by up to 15% compared with ordinary least squares and improving prediction efficiency.

Advances in measurement and computation later led to multi-parameter and imaging-based NDE approaches. Divos and Tanaka (1997) integrated parameters such as MOE, shear modulus, SG, ultrasonic reading and knot diameter ratio for full-size lumber, achieving a correlation coefficient (Pearson's r value, hereafter denoted as r) of 0.74 for UTS prediction. Schajer (2001) used multi-channel X-ray scanning on industrial sawn specimens to measure SG distribution and identify knots, improving the R^2 of the strength prediction model to 0.65-0.75. Steele and Cooper (2003) used radio frequency scanning to estimate MOR, and after moisture correction the model reached an R^2 of 0.59. At the same time, probabilistic and computational approaches gained traction. Wang et al. (1995) developed an integrated method combining a gain-factor model, trend removal and multivariate normal simulation to represent the non-stationary spatial variability of wood properties such as MOE and compressive strength, with results in good agreement with experimental data. Fink and Kohler (2014) combined dynamic MOE with the total knot area ratio for full-size spruce dimension lumber, achieving high accuracy for clear-wood stiffness ($R^2 = 0.95$) and knot-area stiffness ($R^2 = 0.91$). Nazerian et al. (2020) applied an artificial neural network (ANN) to predict the MOR of laminated veneer lumber, capturing the nonlinear effects of resin composition and pressing conditions. Their optimal 4-6-1 ANN architecture, determined through trial and error and trained with the Levenberg-Marquardt backpropagation algorithm, achieved a testing set R^2 of 0.83. Taken together, this body of work has progressed from simple empirical relations to multi-scale frameworks combining experimental data, statistical modelling and computational analysis, deepening the understanding of variability in wood properties and improving the reliability of strength prediction in both research and industrial settings.

In recent years, machine learning (ML) has been applied to a growing range of problems in timber engineering, offering a way to represent nonlinear interactions among multiple predictors that are difficult to specify in advance through fixed functional forms. Its use has expanded across material characterization, property prediction and structural performance assessment, particularly for anisotropic materials where the relationships between inputs and outputs are not easily captured by closed-form expressions (Feng et al., 2024). Common ML algorithms include regression, classification, clustering and association analysis (Wang et al., 2023), with model performance typically assessed through accuracy, precision, recall or mean squared error (MSE) and verified by experimental or simulation results. At the same time, the performance of ML relative to classical multivariable regression depends strongly on specimen size, on the physical relevance of input variables and on the degree to which the training data span the variable space, so a meaningful evaluation requires direct comparison with conventional statistical models on the same dataset rather than reliance on results reported under different experimental conditions.

Within this broader trend, ML has been applied to a variety of wood materials and property combinations. Xin et al. (2022) combined resistance drilling, ultrasonic velocity and colour parameters in a relevance vector machine to predict the MOR and parallel-to-grain compressive strength of clear-wood specimens from ancient timber, achieving R^2 values of 0.79 and 0.77, respectively. Haftkhani et al. (2021) predicted the MOR and MOE of heat-treated fir (*Abies* spp.) specimens using ANN and regression models based on total colour difference, lightness difference, contact angle and mass loss, finding that ANN outperformed regression models, with lightness difference being the best single predictor for MOR (mean absolute percentage error, MAPE = 2.21%, $R^2 = 0.986$) and contact angle the best for MOE (MAPE = 0.61%, $R^2 = 0.974$). Nasir et al. (2021) used a decision tree model with CIELAB colour parameters to predict the MOE and MOR of artificially weathered poplar (*Populus euroamericana*), alder (*Alnus glutinosa*), oak (*Quercus* spp.) and fir (*Abies alba*), reaching R^2 values up to 0.87 for MOE and 0.88 for MOR. Fathi et al. (2020) applied the group method of data handling (GMDH) neural network to predict the MOE and MOR of green poplar (*Populus euroamericana*) using Lamb wave velocity, SG and moisture content (MC) as inputs, achieving r values of 0.95 for MOE and 0.89 for MOR. Rahimi and Avramidis (2022) used multilayer perceptron, radial basis function and GMDH models to predict the final MC of kiln-dried hem-fir specimens from 2,304 specimens, with the GMDH model giving the best performance (MSE = 5.2%) and identifying initial MC as the most influential factor. More recently, Esteghamati et al. (2023) used explainable ML to predict the fire resistance of timber columns from geometric and material properties. Among nine ML algorithms, the random forest (RF) model performed best ($R^2 = 0.84$, root mean square error, RMSE = 5.55 minutes), identifying column capacity and cross-section size as key predictors, while SG, MOE and strength had minor effects. Their model performed better than prescriptive methods such as Lie's equation and the National Design Specifications equations ($R^2 \leq 0.64$) and provided interpretable results through Shapley additive explanations (SHAP) and partial-dependence analysis.

The literature reviewed above shows that ML has been applied with reasonable success to predicting the mechanical properties of various wood and wood-based products. Many previous studies aimed to examine predictive relationships under research or laboratory conditions and therefore used a broad range of input variables to study bending properties, clear-wood or composite specimens. In contrast, the present study focuses on the industrial benefit of predicting the UTS of full-size dimension lumber using measurements that can be obtained easily with existing or readily available production-line equipment. Research on this application remains limited, particularly for hardwood species. This is important because UTS can only be measured directly by testing a specimen to failure, which makes specimen-by-specimen strength assessment impractical in production. Therefore, this study examines whether ML can capture the complex relationships between practical, non-destructive input measurements and UTS and provide rapid predictions for individual specimens with acceptable accuracy. Once trained, the models can estimate strength without complex testing systems or destructive testing of every specimen. These estimates may support more efficient evaluation and quality control of trembling aspen dimension lumber and may also inform the future development of machine-grading approaches. In addition, the proposed approach may provide a basis for applying similar measurements and modelling methods to other hardwood species. In this study, four ensemble models, including RF, extreme gradient boosting (XGBoost), gradient boosting regression tree (GBRT), and adaptive boosting (AdaBoost), are trained and compared using a dataset of 321 specimens. The contribution of each input variable is then examined using the SHAP approach to identify the factors most strongly associated with UTS. The best-performing ensemble model is also compared with MOE-based empirical relations and an MLR model fitted using the same predictors and dataset. Finally, a graphical user interface (GUI) is implemented to support the use of the model in subsequent research and industrial trials.

2. Dataset

2.1. Overview

Hardwoods are rarely used as structural dimension lumber in North America; instead, they are predominantly used for furniture and interior products (Wang et al., 2004). This is mainly because the fibre arrangement is less uniform, vessels are different and defects are inconsistent which makes their mechanical properties less predictable than those of softwoods (Panshin and de Zeeuw, 1981). Aspen has short fibres and a smooth texture, and large dead knots, which frequently occur in this species, often govern its mechanical strength (Dickmann et al., 2001). These relatively consistent features make it suitable for ML analysis, as strength-related defects offer informative signals for algorithms to predict mechanical properties. In the province of Alberta, Canada, aspen makes up approximately 80% of the hardwood resource but remains under-utilized (Alberta Agriculture and Forestry, 2021). Evaluating its mechanical performance contributes to the development of higher-value wood products, mitigates softwood supply shortages, enhances economic benefits and supports Canada's carbon-neutral objectives (Natural Resources Canada, 2023).

In ML models, the selection of input parameters plays a critical role in predicting the accuracy and reliability of the output variables. In this study, a series of NDE tests were conducted to generate input variables from dimension to physical and mechanical properties. The data collection workflow is illustrated in Fig. 1. The target variable was the UTS, which is a key strength parameter commonly considered in light wood-frame construction, particularly when lumber is used as tension members such as trusses. Since the intended application of this work is in industrial practice, the input features were selected based on their relevance to industry and accessibility, ensuring that the study can be replicated and implemented in practice in the near future. Details of the data collection process are provided in following Section 2.2 to Section 2.5.

2.2. Lumber collection

Aspen lumber dimension of 38 mm × 89 mm (nominal 2 in. × 4 in.) was obtained from a sawmill in Hines Creek, Alberta. A total of 321 specimens were collected and classified into two visual grades, select structural (SS) and No. 2, with lengths of 3,048 mm (10 ft.) and 3,658 mm (12 ft.) (Table 1). The lumber originated from logs averaging 216 mm (8.5 in.) in small-end diameter. All specimens were kiln-dried at 95 °C for 96 hours to a target MC of about 16%. Visual grading was carried out by a certified inspector following the "Standard Grading Rules for Canadian Lumber" (NLGA, 2019). The material was then well wrapped and transported to the Wood Science and Technology Centre at the University of New Brunswick, where it was further conditioned for 3 months under ambient temperature and relative humidity (RH) and prepared for property assessment.

2.3. Physical properties

After conditioning, the dimensions of each specimen were measured. A vernier calliper was used to record cross-section dimension values at both ends and the mid-span, resulting in three readings for the narrow-face and three for the wide-face, respectively. The length was measured with a tape, however, due to negligible visual error and the limited precision of the tape, only the two lengths (3,048 mm and 3,658 mm) listed in Table 1 were adopted. MC was determined immediately using a pin-less moisture meter (Model: Wagner Orion 960), with six measurements taken on each specimen (three on each wide face) at evenly

Table 1
Quantity of aspen dimension lumber in different lengths and grades.

	Length	
	3,048 mm (10 ft.)	3,658 mm (12 ft.)
SS grade	49	79
No. 2 grade	-	193
Total	321	

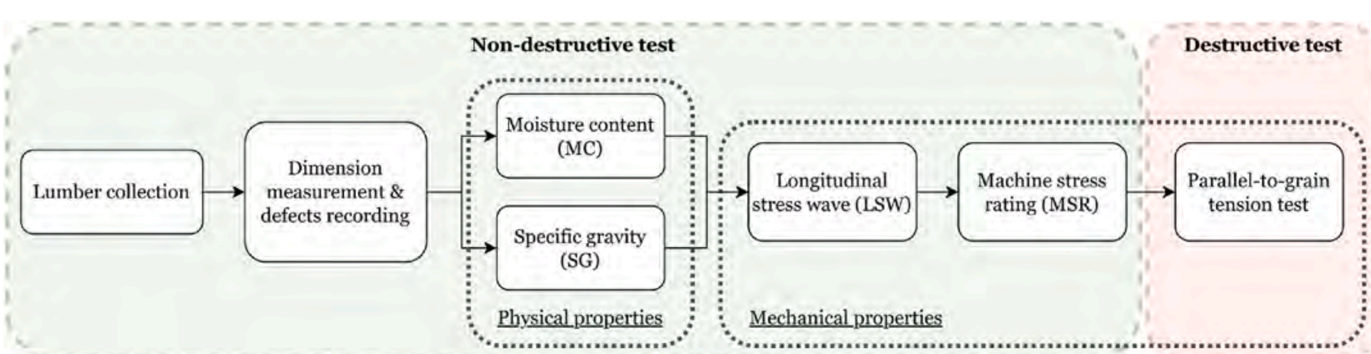

Fig. 1. Workflow for dataset constitution.

distributed locations. The mean value was calculated to minimize the influence of outliers. Since aspen is known for localized "wetwood", which has been previously reported in the literature (Lamb, 1967; Mackes and Lynch, 2001), this factor was considered. Finally, the mass of each specimen was obtained using an industrial balance and together with the measured volume, the SG at the current MC was calculated. For each specimen, the width-to-depth ratio, length, SG and its corresponding MC were documented as four input features for the ML model.

2.4. Visual defects

Visual grading rules are established by the National Lumber Grades Authority (NLGA) based on both qualitative and quantitative visual criteria used to evaluate each piece of lumber (Reid et al., 2009). The main objective is to correlate the visual characteristics of lumber with its tested mechanical performance, allowing the prediction of strength within an acceptable error range. As illustrated in Fig. 2, the UTS of aspen lumber in this study was effectively differentiated between SS and No. 2 grades. Previous research has demonstrated that lumber strength is closely related to the presence and severity of visible defects such as knots, slope of grain, unsound wood, pockets, wormholes, shakes, checks and compression or tension wood (Newlin and Johnson, 1924; Cramer and McDonald, 1989; Barrett and Lau, 1994; Fan et al., 2023). Among these, knots are consistently reported as the most critical factor in strength-reducing effects (Fan et al., 2023). Unlike softwoods which typically have long and well-aligned fibres with numerous small live knots along the grain, aspen is characterized by shorter and finer fibres but develops larger and more distinct dead knots (Lamb, 1967; Youngquist and Spelter, 1990; Balatinecz and Kretschmann, 2001). As a result, the strength of aspen lumber is more strongly influenced by knot presence and size (Wang et al., 2024). Therefore, this study placed particular emphasis on recording knot size as a key input feature for training the ML model.

The maximum strength-reducing defect (MSRD) method (Barrett and Hejja, 1984) was adopted to record the largest defects in each specimen. Following this approach, the three largest surface knots were measured using a ruler to record their maximum diameters (mm). Since the third knot (knot size #3) was negligible in nearly all cases in this study, only the two largest knots, denoted as "knot size #1" and "knot size #2", were retained for analysis considering as two input features. In addition, the combined size of these two knots (knot size sum) and the knot ratio of knot size #1 (can be calculated in Equation (1)) were included as extra two features for model development. The visually identified MSRD was positioned within the span and as close to the mid-length as possible.

$$Knot\ ratio = max\left(\frac{D_k}{b}, \frac{D_k}{d}\right) \tag{1}$$

where D_k is the knot maximum diameter in mm, and b and d are the specimen width and depth in mm, respectively. The larger value between D_k/b and D_k/d was taken to represent the most critical knot proportion regardless of the face orientation.

2.5. Mechanical properties

Two NDE methods, longitudinal stress wave (LSW) and MSR tests, were employed to determine the MOE of each piece of lumber. The MOE values obtained from these tests were considered as mechanical property features and incorporated into the input variables for developing the ML model. For the LSW test, a commercial handheld stiffness grader (Model: MTG-820), developed by Brookhuis (Enschede, the Netherlands) in collaboration with TNO (Delft, the Netherlands), was used to measure the mean MOE of each specimen (Fig. 3(a)). It should be noted that each specimen was measured individually after being lifted from the stack, and Fig. 3(a) illustrates the test arrangement only. The device is commonly employed for rapid on-site grading of lumber and is also utilized in industrial production lines as a substitute for conventional MSR machines proved in previous study (Wang et al., 2024). The device signal is analysed in terms of the natural frequency of the specimen under longitudinal vibration (Biechele et al., 2011). Based on these measurements, the MOE was calculated according to Equation (2) (Niederwestberg et al., 2025):

$$MOE = \frac{(2f_0 l)^2 \rho}{1 - 0.01(MC - 12)} \times 10^{-6} \tag{2}$$

where f_0 is the measured frequency in Hz, l is the measured length of the lumber in mm, ρ is the measured density of the lumber in kg/m^3, and MC is the measured moisture content in %.

The MSR test was performed using a "Cook-Bolinder" machine (Model: Tecmach Limited Stress Grading Machine System SG-AF), which

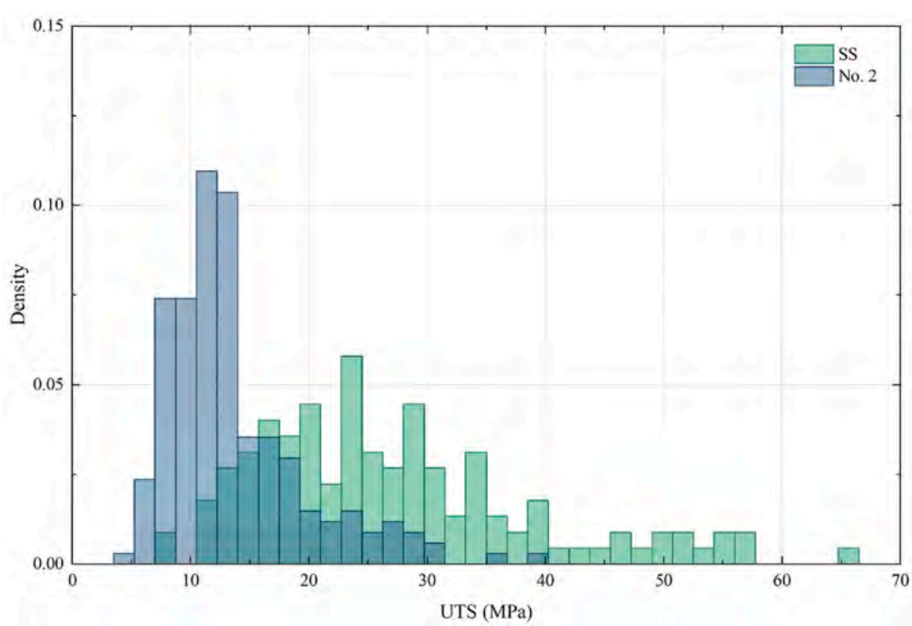

Fig. 2. Distribution of UTS for SS and No. 2 grades aspen lumber.

Fig. 3. (a) LSW test setup (Model: MTG-820); (b) MSR test setup (Model: Tecmach Limited Stress Grading Machine System SG-AF).

was modified for laboratory use. The flatwise centre-point bending method was applied to determine the MOE of each specimen, from which mean, maximum and minimum values were obtained (Fig. 3(b)) (Boström, 1994).

The parallel-to-grain tension test was carried out on each lumber specimen in accordance with ASTM D198 (American Society for Testing and Materials, 2022) using a Metriguard Testing Machine (Model: Metriguard 401). The testing span varied with specimen length, ranging from two distances, 2,438 mm (8 ft.) and 3,048 mm (10 ft.), while each end grip was fixed over a length of 610 mm (2 ft.). Only specimens that failed within the free span were retained in the dataset, and any specimen fracturing at or near the grips was excluded. A loading rate of 10 kN/min was applied, ensuring failure occurred in approximately 4 minutes. The UTS obtained from destructive testing was used as the target variable (output) for training of the ML model.

2.6. Statistical analysis of input and output variables for ML models

In total, 321 specimens were collected from the tests described in Section 2.2 to Section 2.5. The input variables were grade, length, width-to-depth ratio, MC, SG, the sizes of the two largest knots (namely, knot size #1 and knot size #2) and their sum (knot size sum), the knot ratio and the MOE values (namely, LSW-MOE, MSR-MOE min. and MSR-MOE mean) measured by both the LSW and MSR methods. The output variable was the UTS. In this study, the experimental dataset was cleaned and refined to ensure consistent and representative inputs. The grade variables were one-hot encoded. All physical quantities were converted to SI units and MOE were expressed in GPa to reduce scale bias. Width and depth were combined into a width-to-depth ratio. Knot-related inputs were derived from measured values, with the third-largest knot excluded to avoid sparsity. Two engineered features were added: knot size sum, which captures the cumulative effect of knots (Wright et al., 2019) and knot ratio, defined as the largest knot relative to the face dimension due to its strong link to the UTS (Roblot et al., 2010). Detailed information of input and output variables is presented in Table 2.

Fig. 4 illustrates the correlations between selected eight input variables (X3 to X5, X8 to X12) and the target variable (Y) along with their r values. Four input variables were neglected in Fig. 4 because grade (X1) and length (X2) were treated as a binary variable, the knot size sum was found to correlate more strongly with the target variable than the individual measures (knot size #1 and knot size #2). Among them, MSR-MOE min. showed the strongest positive correlation with the UTS ($r = 0.704$), showing a clear relationship between lumber stiffness and strength. In contrast, the width-to-depth ratio exhibited the weakest positive correlation ($r = 0.056$). The knot size sum showed the strongest negative correlation with target ($r = -0.523$), which is consistent with

Table 2
Statistical properties of input and output variables for the dataset.

Variables			Min.	Max.	Mean	Median	St. Dev.
Input	X1	Grade (SS/No. 2)	0	1	-	-	-
	X2	Length (m)	3.048	3.658	-	-	-
	X3	Width-to-depth ratio	2.213	2.403	2.313	2.314	0.026
	X4	MC (%)	5.5	9.9	7.7	7.7	0.796
	X5	SG	0.31	0.54	0.42	0.41	0.035
	X6	Knot size #1 (mm)	0	78	29	27	14.803
	X7	Knot size #2 (mm)	0	73	10	0	14.625
	X8	Knot size sum (mm)	0	107	39	37	22.768
	X9	Knot ratio	0	1	0.472	0.414	0.287
	X10	LSW-MOE (GPa)	7.656	13.130	10.308	10.144	1.010
	X11	MSR-MOE min. (GPa)	4.385	14.648	8.514	8.303	1.794
	X12	MSR-MOE mean (GPa)	6.907	15.916	10.675	10.511	1.581
Output	Y	UTS (MPa)	4.080	65.671	18.991	15.523	10.873

previous findings that larger knots reduce lumber strength (Fan et al., 2023; Wang et al., 2024).

3. Development of ML models

In this study, ML models were developed to predict the UTS of aspen lumber. Fig. 5 illustrates the overall workflow for developing the ML models. To minimize the risk of overfitting, the dataset was partitioned into training and testing subsets using the holdout method (Raschka, 2018) with 80% of the data used for training and 20% reserved for testing. The data were randomly partitioned to mitigate locality effects. The training set was used for model development including iterative learning and hyperparameter optimization. Given the limited dataset size, a grid search combined with five-fold cross-validation was employed for hyperparameter tuning (Anguita et al., 2012). In this

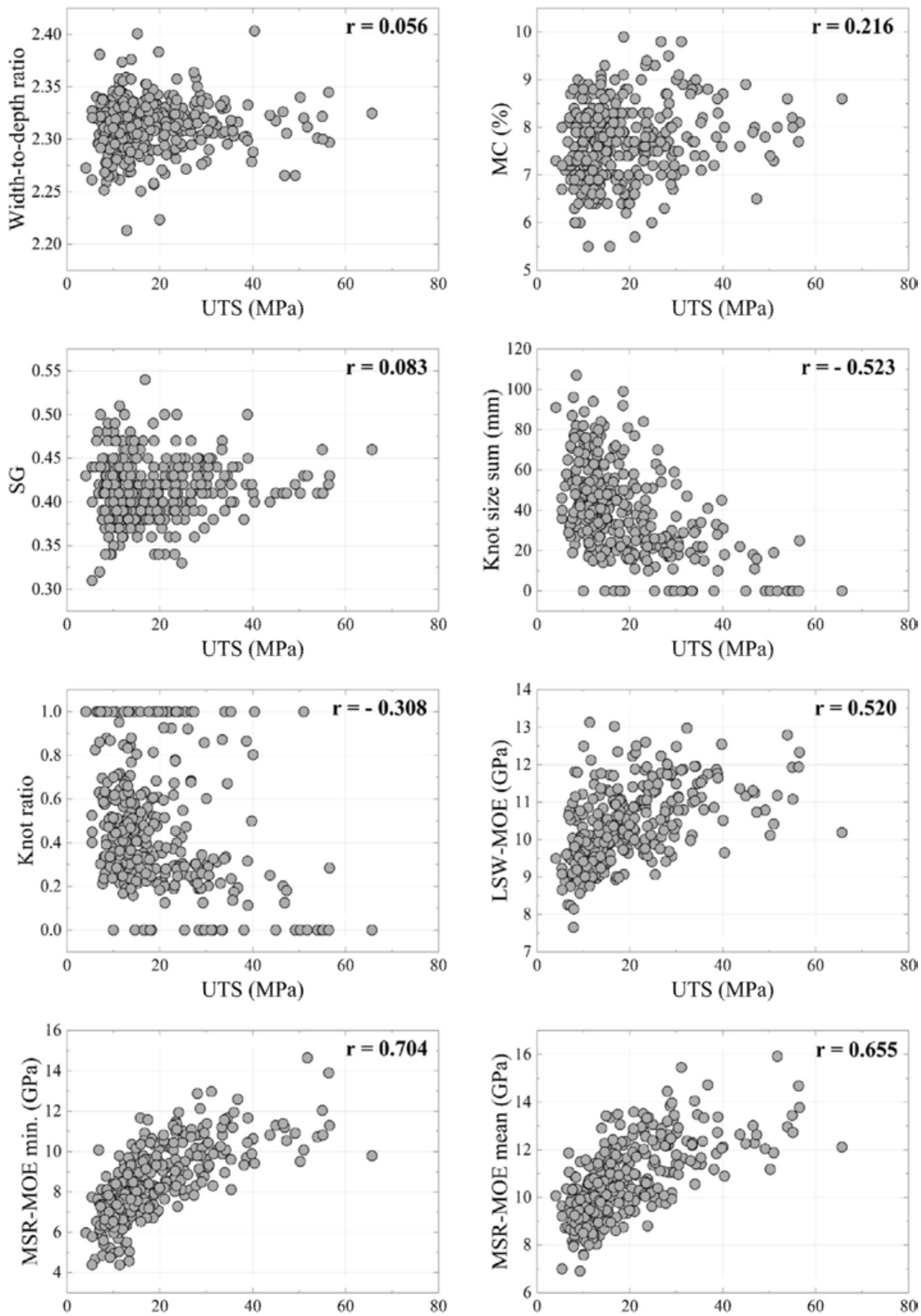

Fig. 4. Correlation between selected 8 input variables and UTS of the dataset.

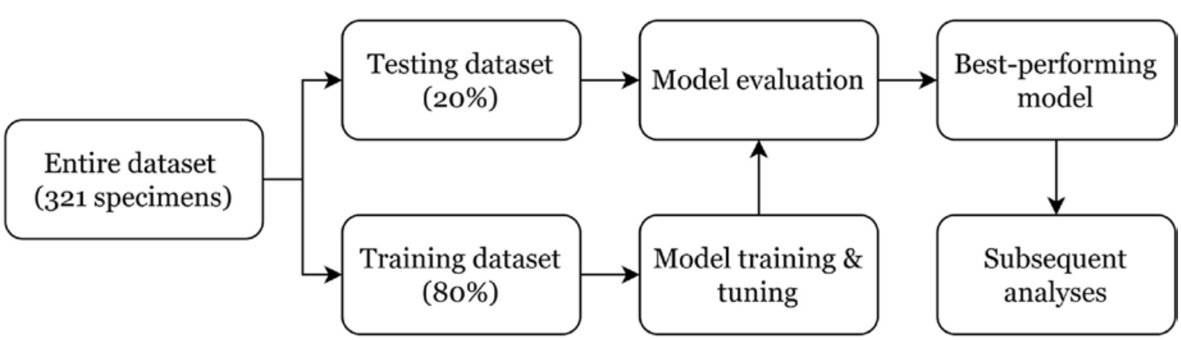

Fig. 5. Flowchart of the ML model development process.

procedure, the training data are randomly divided into five equal subsets, each subset serves once as a testing set while the remaining four are used for training. The MSE is used as the primary evaluation metric for model selection, and the configuration yielding the lowest MSE is considered optimal (Bardenet et al., 2013).

This study considered ensemble models due to the complexity inherent in predicting lumber strength. As an anisotropic material, lumber exhibits significant directional variability in its mechanical and physical properties. Ensemble methods are well-suited to model such complexity, as they are recognized for providing high accuracy, stability and generalization (Nguyen et al., 2023). Furthermore, their application to dimension lumber strength prediction also remains underexplored, presenting a clear opportunity for this research.

Four ensemble models, including RF, XGBoost, GBRT and AdaBoost were employed in this study. These models are based on decision tree ensembles that aggregate multiple weak learners to form a strong predictive model (Nguyen et al., 2021a; Nguyen et al., 2021b; Nguyen et al., 2022). RF (Breiman, 2001) represents the bagging approach, building multiple decision trees in parallel from bootstrap specimens and randomly selected feature subsets, and averaging their predictions to reduce variance and limit overfitting (Speiser et al., 2019). XGBoost (Chen and Guestrin, 2016), GBRT (Friedman, 2001) and AdaBoost (Freund and Schapire, 1997) are boosting methods that train trees sequentially, with each learner addressing the errors left by the previous one. XGBoost enhances this framework with regularization, subsampling and second-order gradient information to improve robustness and computational efficiency (Lim and Chi, 2019). GBRT reduces prediction error by fitting trees to the negative gradients of the loss function (Prettenhofer and Louppe, 2014), while AdaBoost increases the influence of difficult specimens to strengthen generalization. These algorithms have shown strong performance and reliability in modelling wood mechanical properties (Esteghamati et al., 2023). Detailed information of four selected ML models can be found in the literature (Freund and Schapire, 1997; Breiman, 2001; Friedman, 2001; Chen and Guestrin, 2016).

In this study, for each ML algorithm, key hyperparameters such as the number of estimators, tree depth, learning rate and regularization terms were tuned using a grid search with five-fold cross-validation to obtain the highest accuracy model. To make the tuning process consistent across models, the search covered a representative and sufficiently wide set of hyperparameters commonly used in tree-based ensemble methods. For the RF model, the grid search examined the number of trees, maximum tree depth, the minimum specimens required for node splitting and leaf formation, feature selection strategies and whether bootstrap sampling was used. For XGBoost, the search space included the number of boosting rounds, tree depth, learning rate, row and column subsampling ratios and the L1 and L2 regularization terms. The GBRT model was tuned by varying the learning rate, number of estimators, tree depth, splitting rules, leaf size and subsampling proportion. For AdaBoost, the search focused on the number of estimators, learning rate and the depth of the base decision tree. These hyperparameter ranges were chosen to offer enough flexibility to capture nonlinear behaviour without enlarging the search space unnecessarily. A complete list of the hyperparameters and their corresponding ranges are provided in Table 3. After tuning, the models were retrained on the full training dataset using the optimal settings and evaluated on a hold-out testing set. In this study, model performance was assessed using R^2, RMSE, MAPE and r. R^2 describes the proportion of variance explained by the model, RMSE and MAPE quantify absolute and relative prediction errors, providing a balanced view of model accuracy and robustness, r reflects the strength of the linear association between predicted and observed values (Cameron and Windmeijer, 1997).

After training of the ML models, interpreting predictive models are important for understanding how individual input features influence the predictions. In this study, SHAP approach was used. The SHAP framework proposed by Lundberg and Lee (2017) employs a game-theoretic

Table 3
Hyperparameter ranges used in the grid search for four ML models.

Model	Hyperparameter range
RF	n_estimators: {100, 300, 500}
	max_depth: {none, 10, 20, 30, 50, 100}
	min_samples_split: {2, 5, 10, 20}
	min_samples_leaf: {1, 2, 4, 8}
	max_features: {“sqrt”, “log2”, “none”}
	bootstrap: {true, false}
XGBoost	n_estimators: {100, 300, 500}
	max_depth: {3, 6, 10}
	learning_rate: {0.01, 0.1, 0.3}
	subsample: {0.6, 0.8, 1.0}
	colsample_bytree: {0.6, 0.8, 1.0}
	reg_alpha (L1): {0, 0.1, 1}
	reg_lambda (L2): {1, 1.5, 2}
GBRT	n_estimators: {100, 300, 500}
	learning_rate: {0.01, 0.05, 0.1}
	max_depth: {3, 10, 20, 30}
	min_samples_split: {2, 5, 10}
	min_samples_leaf: {1, 3, 5}
	subsample: {0.6, 0.8, 1.0}
	max_features: {“sqrt”, “log2”, “none”}
AdaBoost	n_estimators: {100, 300, 500}
	learning_rate: {0.01, 0.05, 0.1, 0.2}
	base estimator: decision trees with max_depth {3, 5, 10, 15}

approach. Specifically, it quantifies the marginal contribution of each feature to a prediction, including both the direction and magnitude of its effect (Li, 2022). Several implementations exist depending on the model type, such as “kernel SHAP”, “deep SHAP” and “tree SHAP” (Rodríguez-Pérez and Bajorath, 2020). Because the models used in this study are based on decision-tree ensembles, the tree SHAP method was applied (Lundberg et al., 2018; Nguyen et al., 2022a;Nguyen et al., 2022b) to compute and visualize feature contributions. Tree SHAP satisfies the principles of local accuracy, missingness and consistency. This means that the feature attributions sum to the model output, features not used by the model have no effect, and the contribution of an important feature does not decrease when irrelevant features are removed (Ponce-Bobadilla et al., 2024).

4. Performance of developed ML models

In this section, the performance of the developed ML models is presented and discussed. Although there is no universal standard, many works adopt a training-testing split in which 60%-80% of the data are used for training, and the remaining 40%-20% are used for testing, respectively (Nguyen et al., 2023; Huang et al., 2023). Therefore, the choice of split ratio was examined through a sensitivity analysis in this study. Fig. 6 compares the performance of the four ML models under five training-testing split ratios: 6:4, 6.5:3.5, 7:3, 7.5:2.5, and 8:2, using the confirmed set of input features. Model performance is presented in terms of R^2 (Fig. 6(a)), RMSE (Fig. 6(b)), and MAPE (Fig. 6(c)). None of the three metrics changes monotonically with the training proportion. R^2 remains within 0.614-0.671 up to the 7.5:2.5 split and then increases to 0.678-0.718 at the 8:2 split, while MAPE shows a similar pattern. RMSE reaches its lowest value at the 6.5:3.5 split but remains relatively stable across the other split ratios. Overall, the 8:2 split provides the strongest performance, with the highest R^2 and the lowest MAPE for most models. It was therefore adopted as the standard configuration for all subsequent analyses in this study. After the full training and hyperparameter-tuning process, the optimal hyperparameters for each model were determined, and the final configurations are summarized in Table 4. Subsequent comparisons were conducted using the models developed with these optimal hyperparameters.

As shown in Table 5, the accuracy of the four ML models in predicting the UTS was evaluated using four indexes: R^2, RMSE, MAPE and r. The overall predictive performance of the four models on the testing set was comparable. Among them, XGBoost and GBRT achieved the best

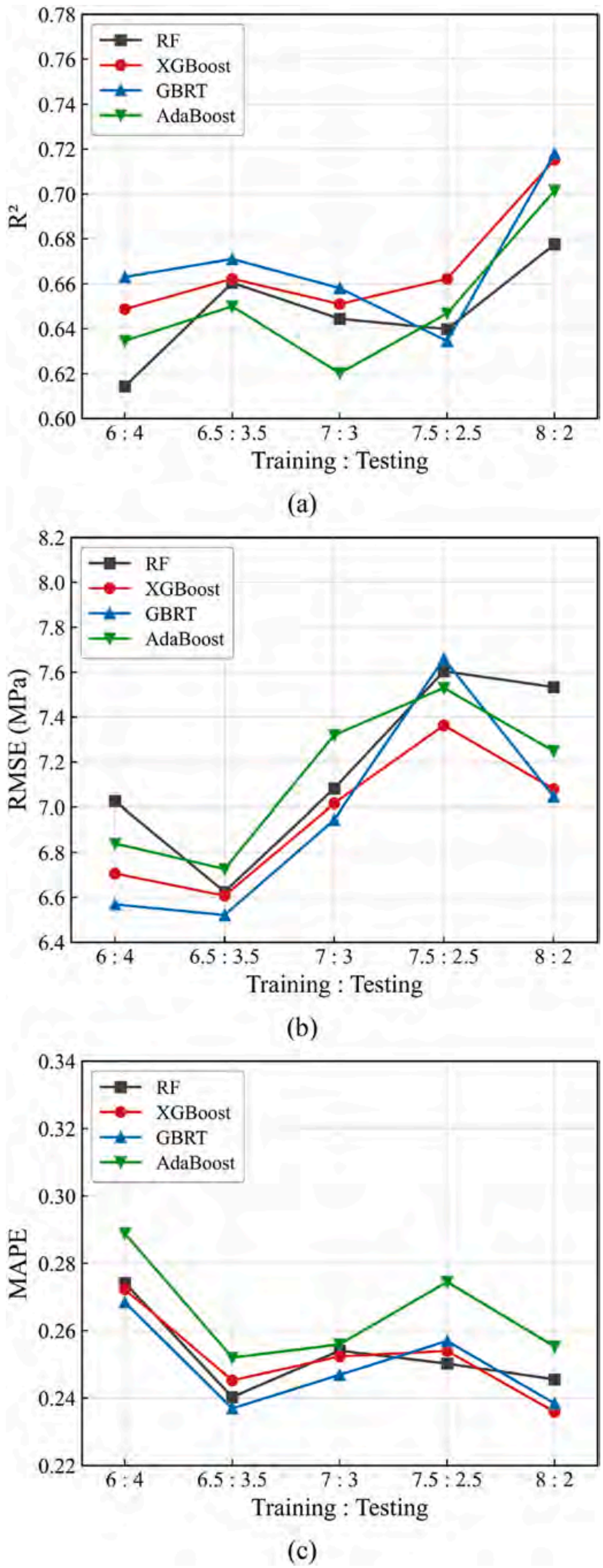

Fig. 6. Performance comparison of four ensemble models under different training-testing data split ratios: (a) R^2; (b) RMSE; (c) MAPE.

performance, with R^2 values of 0.715 and 0.718, RMSEs of 7.080 MPa and 7.047 MPa, respectively. In contrast, RF showed the weakest overall testing set performance, with the lowest R^2 (0.678) and the highest RMSE (7.533 MPa), although its MAPE (0.246) was slightly lower than that of AdaBoost (0.255). These results indicate that RF produced less accurate UTS predictions than XGBoost and GBRT for the present dataset.

Fig. 7 presents the prediction results of the four ML models on both the training and testing sets, showing their fitting performance and generalization capability. In the scatter plots, red and grey markers represent the relationships between the predicted and actual (true) values for the training and testing sets, respectively, with the fitting curve applied. The black dashed line ($y = x$) denotes the ideal prediction line where predicted and actual values coincide. As shown in Fig. 7, the prediction lines for XGBoost (Fig. 7(b)) and GBRT (Fig. 7(c)) were closely aligned with the perfect line ($y = x$). It can be seen that these models accurately captured the nonlinear relationship between the input features and the actual UTS. RF (Fig. 7(a)) displays a slightly wider scatter on the testing set, but the overall trend remains consistent, suggesting reasonable generalization. AdaBoost (Fig. 7(d)) also shows reasonable predictive performance, although its regression line has a slightly greater deviation from the ideal line and its testing set R^2 (0.701) is lower than those of XGBoost and GBRT. In comparison, RF (Fig. 7(a)) shows a wider scatter on the test set and the lowest testing set R^2 (0.678), indicating the weakest overall predictive performance among the four models. Overall, XGBoost and GBRT provide the most accurate and stable predictions, followed by AdaBoost, while RF shows relatively larger prediction errors.

Fig. 8(a) presents the comparative performance plot of each ML model, where the optimal ML model should combine a high R^2 with a low RMSE. Therefore, R^2 and RMSE were used as the two key indicators for comparison. The results show that GBRT ($R^2 = 0.718$, RMSE = 7.047 MPa) achieved the strongest overall predictive performance, appearing in the lower-right region of the plot, which indicates high prediction accuracy with small errors. In contrast, XGBoost ($R^2 = 0.715$, RMSE = 7.080 MPa), AdaBoost ($R^2 = 0.701$, RMSE = 7.249 MPa) and RF ($R^2 = 0.678$, RMSE = 7.533 MPa) are followed and exhibited lower predictive accuracy. Fig. 8(b) shows the R^2 difference between the training and testing datasets. A smaller difference indicates better generalization, while a larger difference may suggest potential overfitting or underfitting. The results point out that RF had the largest difference (0.0784), which can be noted as slightly overfitting to the training data. GBRT (0.0544) and XGBoost (0.0461) exhibited smaller differences, reflected more stable generalization. Although AdaBoost had the lowest overfitting risk (0.0205), its predictive accuracy was limited as stated above. Based on results shown above, GBRT provided the most reliable results of the four ML models, with comparatively higher accuracy and stable generalization.

To obtain more comprehensive comparison, the ratio of predicted-to-experimental values (predicted value divided by experimental value) was calculated for the full dataset. The mean, standard deviation (St. Dev.) and coefficient of variation (CoV) were examined, as shown in Table 6. A ratio close to 1 represents a good match between predicted and measured strength. A ratio above 1 indicates that the model overestimates the strength, whereas a ratio below 1 indicates underestimation. RF produced the mean ratio closest to the reference value of 1 (1.079), followed by GBRT (1.092), XGBoost (1.094) and AdaBoost (1.123). The corresponding St. Dev. values were 0.279, 0.311, 0.318 and 0.345, and the CoV values were 0.258, 0.285, 0.291 and 0.307, reflecting relatively consistent predictions with limited dispersion.

To examine the influence of individual input variables on model predictions, Fig. 9 presents the SHAP summary plots for GBRT, the ML model that achieved the best predictive performance compared to the remaining models comprehensively. Feature importance and interpretability for the model was assessed using the tree SHAP algorithm.

In Fig. 9, each row corresponds to one input feature, and each point

Table 4
Primary parameters for the four ML models.

RF	XGBoost	GBRT	AdaBoost
Number of trees = 100	Number of trees = 300	Number of trees = 300	Number of trees = 100
Max_depth = 10	Max_depth = 3	Max_depth = 3	Max_depth = 3
Max_features = sqrt	Subsample = 0.6	Max_features = none	Learning rate = 0.1
Min_samples_leaf = 2	Colsample_bytree = 0.8	Min_samples_leaf = 3	-
Min_samples_split = 20	Learning rate = 0.01	Min_samples_split = 10	
Bootstrap = false	reg_alpha (L1) = 1	Learning rate = 0.01	
-	reg_lambda (L2) = 1	Subsample = 0.6	

Table 5
Prediction accuracy of the four ML models for UTS.

Parameter	RF		XGBoost		GBRT		AdaBoost	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
R^2	0.756	0.678	0.761	0.715	0.772	0.718	0.722	0.701
RMSE (MPa)	4.946	7.533	4.893	7.080	4.779	7.047	5.280	7.249
MAPE	0.206	0.246	0.238	0.236	0.232	0.238	0.273	0.255
r	0.883	0.849	0.887	0.872	0.890	0.866	0.857	0.852

represents the SHAP value of that feature for an individual observation. The horizontal axis indicates the positive or negative contribution of a feature to the predicted UTS, while the vertical ordering reflects overall feature importance based on mean absolute SHAP values. Positive SHAP values indicate an increase in the prediction, whereas negative values indicate a decrease. Point colour represents the relative magnitude of the feature value, with red indicating higher values and blue indicating lower values. As shown in Fig. 9, the GBRT model identified MSR-MOE min., length and MSR-MOE mean as the most influential variables governing UTS prediction. MSR-MOE min. exhibited the strongest contribution among all features, consistent with the high correlation observed in Fig. 4. MSR-MOE min. captures the weakest stiffness zone along the lumber, which typically corresponds to the largest knot region (Cramer and Goodman, 1983). Its dominant influence on the SHAP plot reinforces that tensile behaviour is governed by knot-controlled weak sections, consistent with previous studies (Samson and Blanchet, 1992; As et al., 2006; Fan et al., 2023). Length also displayed a substantial effect, as longer specimens have a greater likelihood of containing strength-reducing defects. In contrast, MSR-MOE mean represents the average stiffness measured along the entire specimen and reflects its global stiffness profile rather than the localized weakest zone captured by MSR-MOE min. Collectively, these factors provide a clearer understanding of how defects form and spread, and how their presence ultimately influences the strength response.

To conclude, all the results and discussion above show that GBRT provided the best balance between accuracy and consistency, with the values closest to the actual results with relatively small variation. XGBoost showed comparable performance but was slightly less effective. AdaBoost performed lower yet remained stable, whereas RF exhibited the greatest bias and variability, indicating weaker generalization. Therefore, the best performance GBRT model was proposed and used for the subsequent analyses.

5. Performance comparison of the proposed GBRT model with empirical models, multivariable linear regression, and previous studies

5.1. Comparison with MOE-based empirical models and multivariable linear regression

In 1983, the Canadian Wood Council and the Canadian lumber industry launched the “Lumber Properties Project”, which led to the Canadian in-grade report (Barrett and Lau, 1994). This report has long served as an important reference for estimating structural lumber properties in Canada based on extensive standard testing. Among the model forms discussed in the report, linear and three-parameter power-law models have commonly been used to describe the relationship between lumber strength and MOE.

In this study, these model forms were not used with their original softwood-based parameters. Instead, they were refitted using the present aspen dataset to allow comparison under the same material source as the proposed GBRT model. The MOE-based empirical models were developed using one stiffness variable at a time, including MSR-MOE min., MSR-MOE mean, and LSW-MOE, respectively. In addition, an MLR model was developed using the same 12 input variables as the GBRT model. Therefore, this section includes two baseline levels: traditional MOE-based models and an MLR model with the same input information as the ML models.

The ensemble ML models developed in the previous section were trained and tuned under a fixed 8:2 training-testing split, with hyperparameters selected using five-fold cross-validation within the training set. For the empirical models and the MLR model in this section, performance was evaluated using 100 random 8:2 repetitions to reduce the influence of a single holdout split. The mean testing set R^2 across the 100 repetitions was used as the main performance metric. For the measured-versus-predicted scatter plots, the repetition with a testing set R^2 closest to the 100-repetition mean was selected as the representative split, so that the plotted results reflect the average model behaviour rather than an unusually favourable or unfavourable split.

This comparison is presented in three steps. First, single-variable linear regression models were used to examine the predictive value of each MOE variable. Second, three-parameter power-law models were used to evaluate whether a nonlinear empirical relationship between MOE and UTS could improve prediction. Third, the MLR model was compared with the proposed GBRT model to examine whether nonlinear ML model can provide additional predictive value when the same 12 input variables were used. This comparison focuses on individual-specimen UTS prediction and is not intended to establish a new grading rule or design value.

5.1.1. Single-variable linear regression models

Linear regression has been widely used to describe the relationship between lumber strength and MOE (Green and Kretschmann, 1991; Barrett and Lau, 1994). In this study, three MOE-related variables, namely MSR-MOE min., MSR-MOE mean, and LSW-MOE, were separately used as single input variables to develop linear regression models. These models provide a simple MOE-based reference for evaluating how much UTS variation can be explained by stiffness alone.

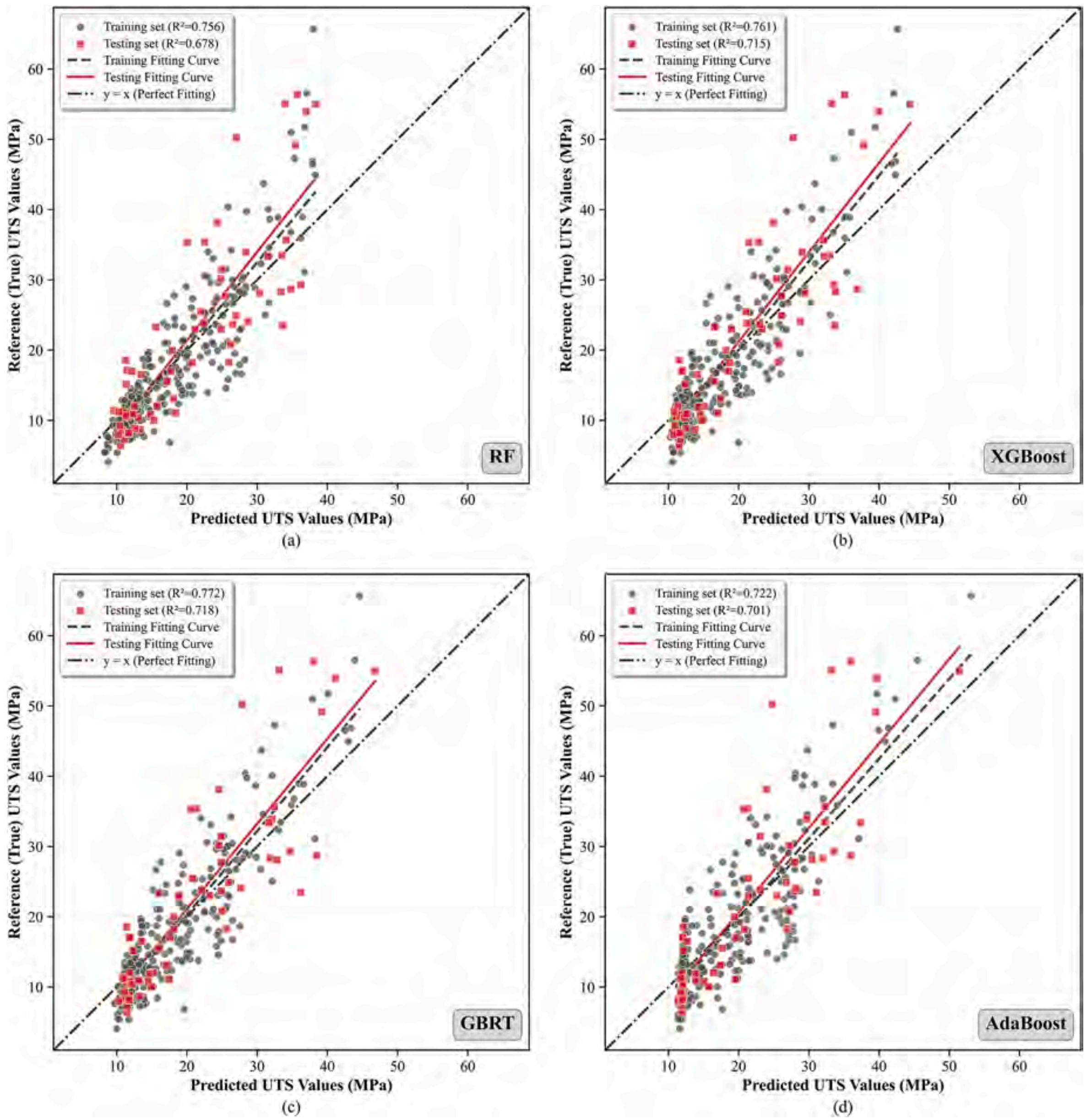

Fig. 7. Comparison of predicted and actual UTS values for four ensemble ML models (a) RF; (b) XGBoost; (c) GBRT; (d) AdaBoost.

To reduce the influence of a favourable or unfavourable training-testing split, the three single-variable linear models were evaluated using 100 random 8:2 repetitions. For each MOE variable, the repetition with a testing set R^2 closest to the 100-repetition mean was selected as the representative split and is shown in Fig. 10. In this representative comparison, MSR-MOE min. provided the best single-variable linear prediction of UTS (Fig. 10(a)), with a testing set R^2 of 0.476. MSR-MOE mean also showed moderate predictive ability, with a testing set R^2 of 0.402 (Fig. 10(b)). In comparison, LSW-MOE gave the weakest performance, with a testing set R^2 of 0.255 (Fig. 10(c)). These results show that MOE is useful for predicting UTS, but a single stiffness variable cannot fully explain the strength variation of aspen lumber.

The better performance of MSR-MOE min. is reasonable because it reflects the lowest stiffness region along the specimen. This local stiffness information is more likely to be associated with weak zones that control tensile failure. In contrast, MSR-MOE mean represents the average stiffness of the full specimen and may reduce the effect of localized defects. LSW-MOE also reflects a global dynamic stiffness response and may be less sensitive to local weak zones. Therefore, for this dataset, local stiffness information was more useful than global stiffness indicators for predicting UTS.

However, these single-variable models should not be treated as direct alternatives to the 12-input GBRT model. They are included as traditional MOE-based reference models. A more balanced comparison is provided by the MLR model in Section 5.1.3, where the same 12 input variables were used.

5.1.2. *Power-law models based on the three-parameter form*

In addition to the linear model, the Canadian in-grade report also used a three-parameter power-law model to describe the relationship between lumber strength and stiffness. This model form was originally developed from large in-grade datasets for commercial softwood species groups. Since the original fitted parameters were not developed for aspen, only the model form was adopted here and refitted using the present aspen dataset. The original model form is given in Equation (3):

$$UTS = a(MOE)^b(c)^{MOE} \tag{3}$$

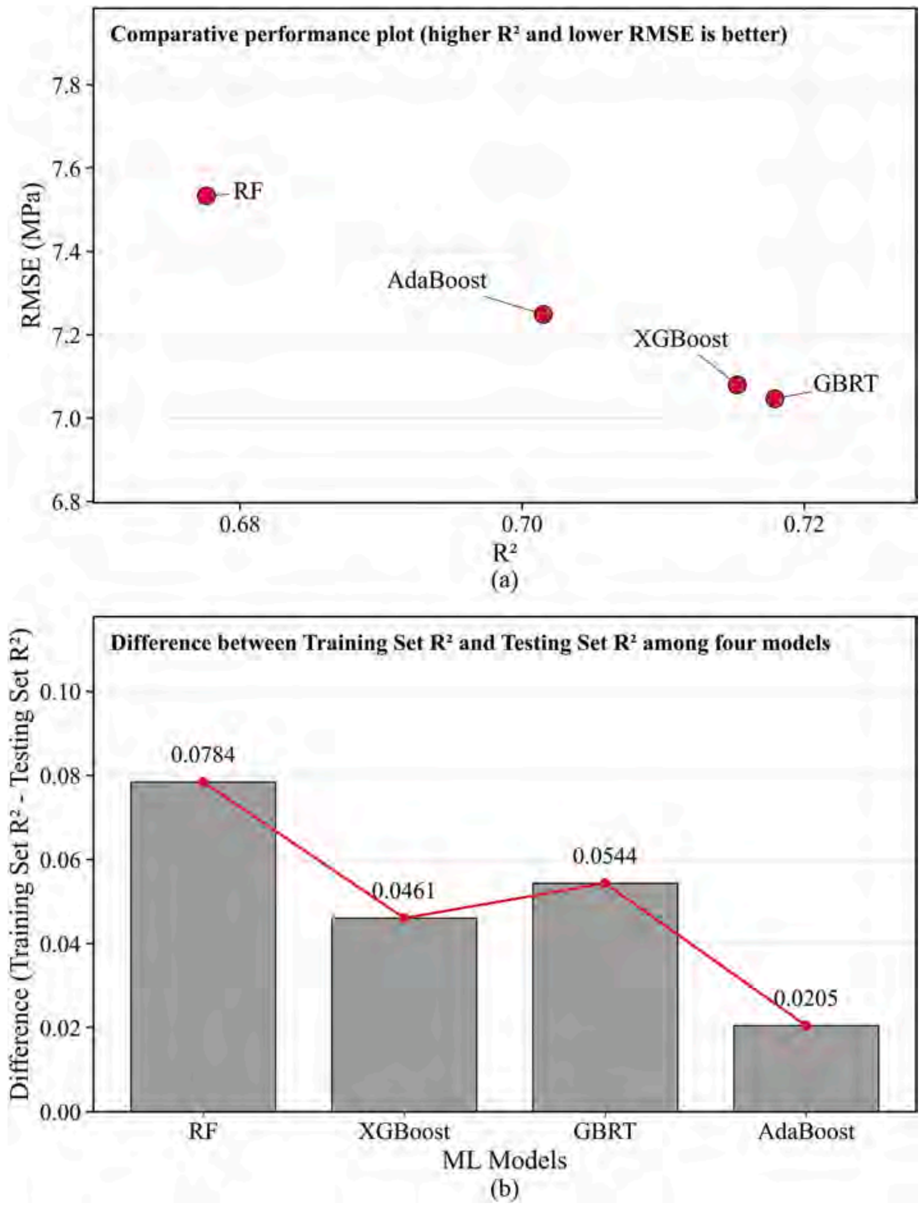

Fig. 8. (a) Comparative performance plot between R^2 and RMSE for four ML models; (b) Comparison based on the difference between training and testing R^2 values.

Table 6
Comparison of predicted-to-experimental value ratios among four ML models.

	RF/ Experiment	XGBoost/ Experiment	GBRT/ Experiment	AdaBoost/ Experiment
Mean	1.079	1.094	1.092	1.123
St. Dev.	0.279	0.318	0.311	0.345
CoV	0.258	0.291	0.285	0.307

where *a*, *b*, and *c* are model parameters, and MOE represents one of the three MOE-related variables. In the present refitting, *c* was fixed at 1 to avoid unstable fitting caused by the limited specimen size, while *a* and *b* were fitted from the training data. Therefore, the fitted model kept the same power-law structure but avoided additional sensitivity from the exponential term. Similar to the linear models, three separate power-law models were developed using MSR-MOE min. (Fig. 11(a)), MSR-MOE mean (Fig. 11(b)), and LSW-MOE Fig. 11(c)) as the single predictor, respectively.

Following the same procedure as in Section 5.1.1, the representative split whose testing set R^2 was closest to the 100-repetition mean was selected for each power-law model. The power-law models showed a performance pattern similar to that of the linear models, with MSR-MOE min. again providing the best single-variable prediction of UTS ($R^2 = 0.487$), followed by MSR-MOE mean ($R^2 = 0.405$) and LSW-MOE ($R^2 = 0.239$) in Fig. 11. Compared with the corresponding linear models, the differences in testing set R^2 were small for all three MOE variables, indicating that changing the model form from linear regression to a power-law model had little practical effect when only one MOE variable was used. These results suggest that the main limitation of the MOE-based empirical models was not only the mathematical form, but also the limited input information, because UTS is affected by MOE, SG, grade, and defect-related characteristics. Therefore, a single MOE variable, whether used in a linear or power-law form, cannot fully describe the strength variation of aspen lumber.

5.1.3. Multivariable linear regression model

The single-variable linear and power-law models discussed in Sections 5.1.1 and 5.1.2 provide useful MOE-based empirical references, but they are not a balanced comparison for the proposed GBRT model

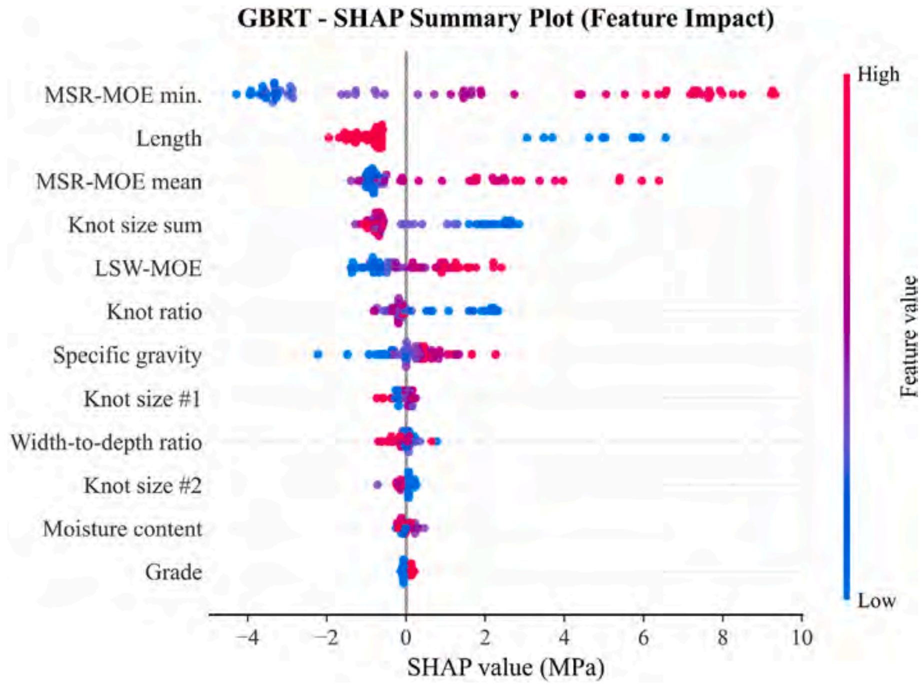

Fig. 9. SHAP summary plots showing the feature impacts on UTS prediction from GBRT model.

because they use only one predictor at a time. To provide a more reasonable linear baseline, an MLR model was developed using the same 12 input variables as the GBRT model listed in Table 2. The visual grade variable was encoded in the same way as in the ML models, with SS assigned as "1" and No. 2 assigned as "0".

The MLR model was evaluated using the same 100-repetition 8:2 data splitting procedure used for the empirical models in Sections 5.1.1 and 5.1.2. The representative split whose testing set R^2 was closest to the 100-repetition mean was selected and is shown in Fig. 12. In this representative comparison, the MLR model achieved a testing set R^2 of 0.585, which was higher than those of the single-variable MOE-based models. For example, compared with the best single-variable power-law model based on MSR-MOE min. ($R^2 = 0.487$), the MLR model increased the R^2 by 0.098. This improvement suggests that knot-related variables, physical properties, and grade provided useful information beyond MOE alone.

However, the MLR model still performed lower than the GBRT model, which achieved a testing set R^2 of 0.718. The absolute increase in R^2 from MLR to GBRT was 0.133. This result suggests that nonlinear modelling provided additional predictive value when the same set of measurable input variables was used. Therefore, the value of the GBRT model should not be interpreted only as the result of using more input variables. Rather, it shows that a nonlinear model can make better use of the same input information for individual-specimen UTS prediction.

The coefficients of the MLR model were not further discussed because the input variables had different units and scales. In this section, the MLR model was used only as a linear reference model with the same input information as the GBRT model. This comparison helped separate the effect of adding more input variables from the effect of using a nonlinear ML model.

5.1.4. Summary of same-dataset model comparison

Table 7 summarizes the performance of the MOE-based empirical models, the MLR model, and the proposed GBRT model based on the same aspen dataset in this study. Among the single-variable empirical models, MSR-MOE min. provided the best performance in both the linear and power-law forms, with testing set R^2 values of 0.476 and 0.487, respectively. This confirms that MSR-MOE min. was the most useful single MOE predictor for UTS in this dataset.

When the same 12 input variables were included in the MLR model, the testing set R^2 increased to 0.585. Compared with the best single-variable power-law model based on MSR-MOE min. ($R^2 = 0.487$), the MLR model increased the R^2 by 0.098, corresponding to a 20.1% relative increase. This result shows that measurable variables beyond MOE, including SG, grade, and knot-related parameters, provided additional information for UTS prediction.

The proposed GBRT model further increased the testing set R^2 to 0.718. Compared with the MLR model, the R^2 increased by 0.133, corresponding to a 22.7% relative increase. Compared with the best single-variable empirical model, the GBRT model increased the R^2 by 0.231, corresponding to a 47.4% relative increase. This result indicates that nonlinear modelling provided additional predictive value when the same set of 12 input variables was used. Therefore, the improvement of GBRT was not only due to the inclusion of more variables, but also due to its ability to capture nonlinear relationships and possible feature interactions that were not represented by the linear additive form of MLR.

Overall, the same-dataset comparison shows that UTS prediction benefited from two factors: the inclusion of multiple measurable variables and the use of a nonlinear ensemble model. However, because the improvement of GBRT over MLR was moderate, the result should be interpreted with caution. The proposed model should be viewed as an aspen-specific individual-specimen UTS prediction model developed from the present dataset, rather than as a general grading model or a design-value model.

5.2. Comparison with previous studies

The same-dataset comparison in Section 5.1 provides a controlled basis for evaluating the proposed GBRT model because the baseline models and the GBRT model were developed from the same aspen dataset and assessed under clearly defined evaluation procedures. In contrast, comparison with previous studies should be interpreted with caution because different studies used different species, specimen sizes, testing methods, specimen sizes, input variables, and model forms. Therefore, this section does not aim to directly rank the present model against previous models. Instead, it places the present results within the

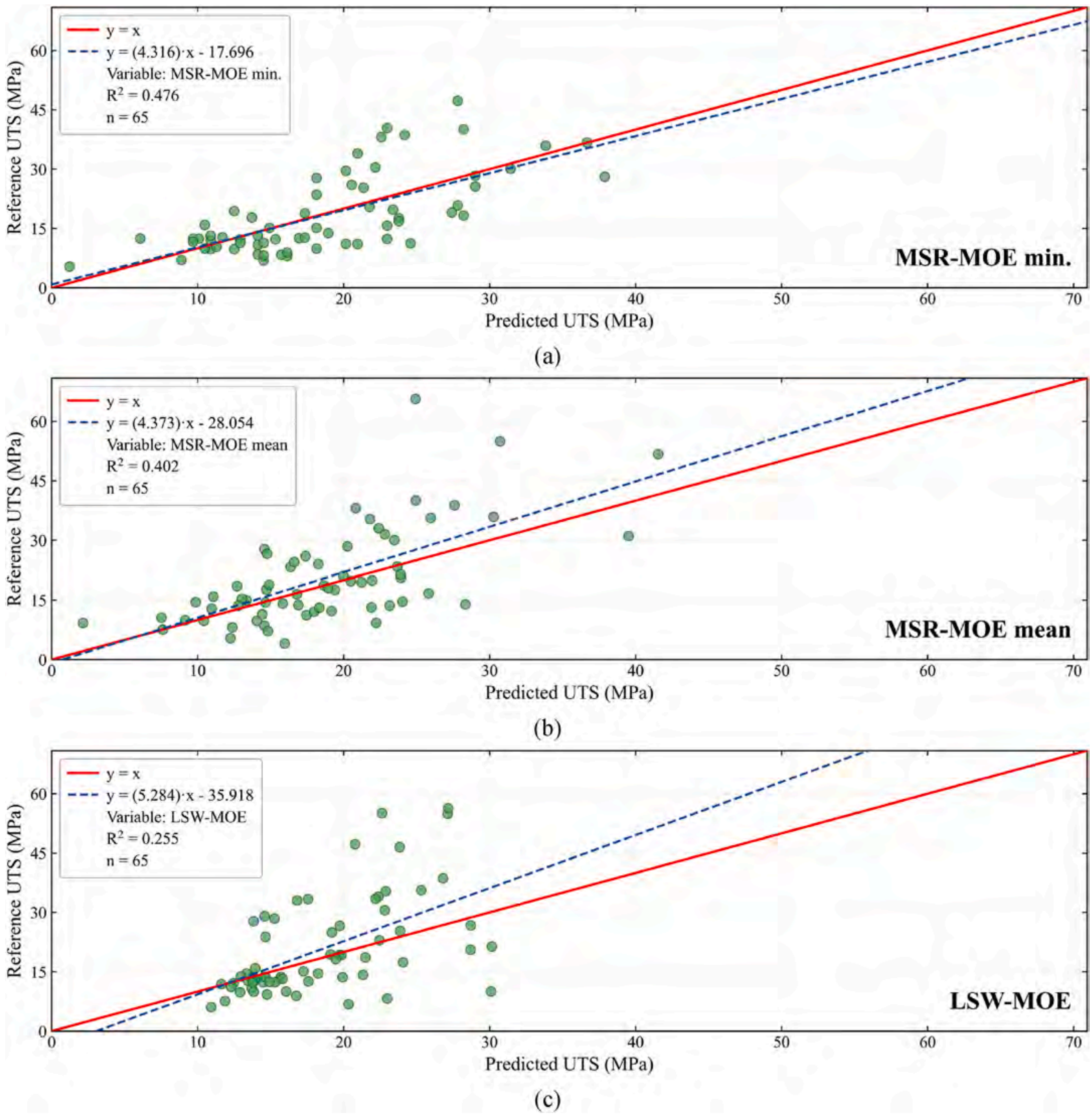

Fig. 10. Testing set measured-versus-predicted parity plots for single-variable linear regression models based on: (a) MSR-MOE min.; (b) MSR-MOE mean; and (c) LSW-MOE.

range of reported studies on UTS prediction of timber.

Saboksayr et al. (2003) studied 1,000 pieces of SPF lumber and used X-ray density imaging, MSR-MOE, slope of grain, and more than 40 image-derived features to develop multiple regression and ANN models. Senalik et al. (2020) studied 103 southern yellow pine specimens with dimensions of 38 mm × 38 mm, which were cut from No. 2 grade 38 mm × 184 mm specimens. Their models used SG, longitudinal dynamic MOE, stress-wave frequency, and acoustic energy parameters to predict UTS. Uzcategui et al. (2023) analysed 702 No. 2 southern yellow pine specimens with dimensions of 38 mm × 140 mm and developed MLR models using SG, longitudinal dynamic MOE, and frequency-domain energy parameters. These studies provide useful references for timber strength prediction, but their models were developed using different species, specimen sizes, measurement systems, and input variables. Some of these approaches used advanced imaging or frequency-domain features, which can provide detailed information on material variation or internal defects but may require more specialized equipment. In contrast, the present study used variables that are relatively accessible in industrial production-line or industry-oriented testing conditions, including grade, SG, knot measurements, MSR-MOE, and LSW-MOE.

As shown in Fig. 13, the testing set R^2 of the proposed GBRT model

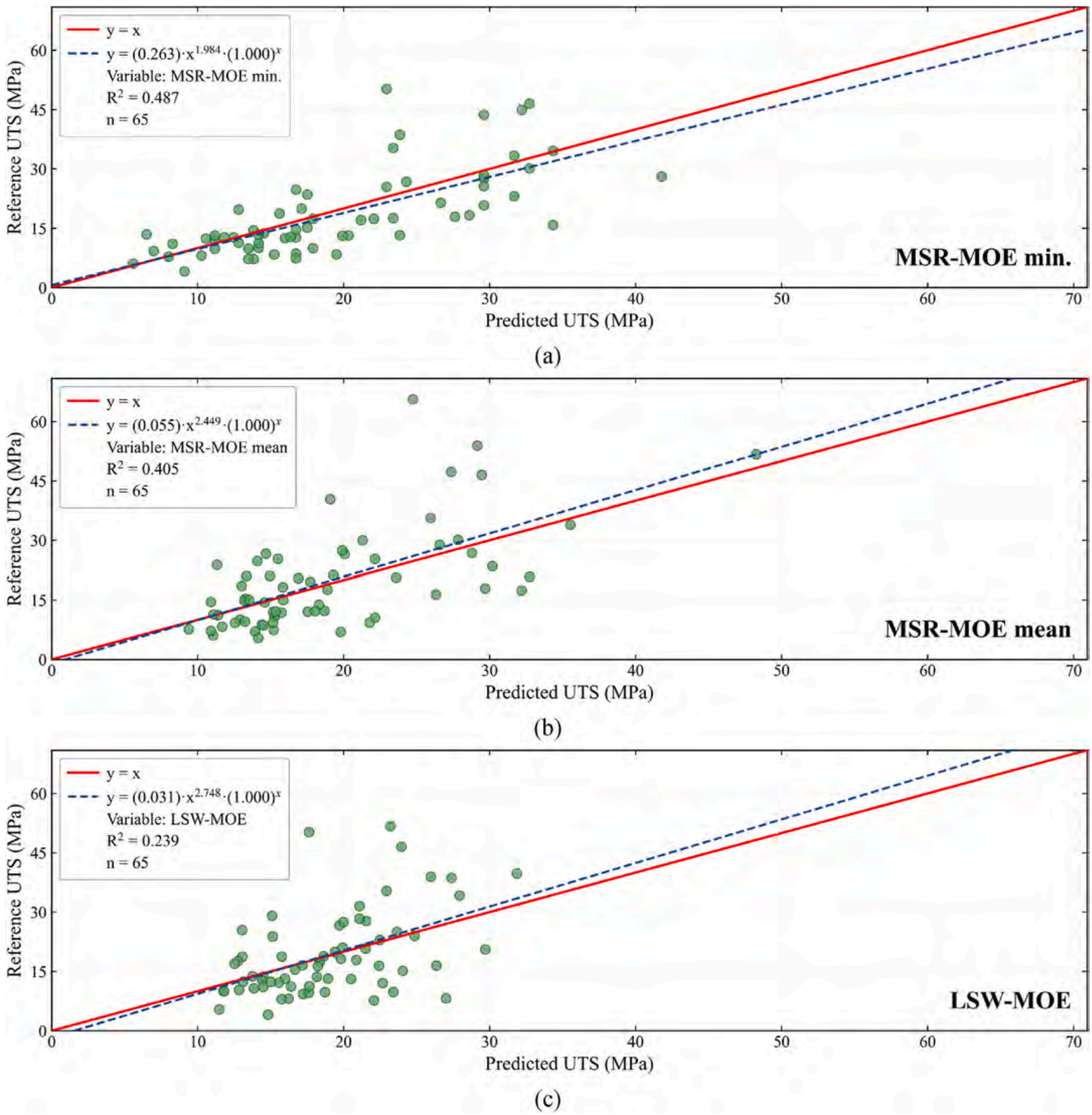

Fig. 11. Testing set measured-versus-predicted parity plots for power-law models based on: (a) MSR-MOE min.; (b) MSR-MOE mean; and (c) LSW-MOE.

was within the higher range of the reported results. However, this figure should be understood as a contextual comparison rather than evidence that the present model is directly superior to previous models. The main point is that the proposed model achieved a comparable level of prediction accuracy while relying on a practical set of input variables. This supports the potential of using measurable grading-related and NDE-related variables for individual-specimen UTS prediction of aspen lumber. However, external validation using independent aspen lumber batches is still needed before the model can be used for routine grading, design-value estimation, or broader industrial application.

6. Development of a graphical user interface for proposed model

To make the developed GBRT model for predicting the UTS of aspen lumber conveniently available to other users, a GUI was created in Python with the built-in tkinter library. Fig. 14 presents the layout of the interface. The program accepts 12 input variables corresponding to those used in the GBRT model. For accurate predictions, the input values should be kept remained within the ranges summarized in Table 2.

As demonstrated in Fig. 14, the UTS of aspen lumber can be obtained quickly through the GUI. This shows the applicability of the proposed model in practical use. The executable file compiled for the GUI can be downloaded at the link (https://github.com/Dawei-Wood/Trembling-

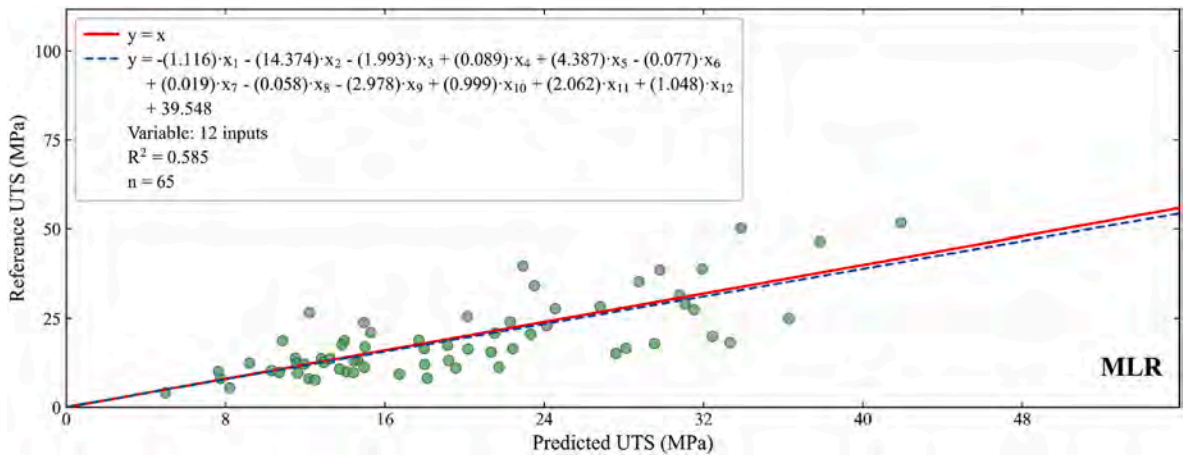

Fig. 12. Testing set measured-versus-predicted parity plot for the MLR model using 12 input variables.

Table 7
Performance comparison of different models based on the same aspen dataset.

Model	Input Variable(s)	Training R^2	Testing R^2
Linear	MSR-MOE min.	0.499	0.476
	MSR-MOE mean	0.432	0.402
	LSW-MOE	0.267	0.255
Three-parameter power-law	MSR-MOE min.	0.514	0.487
	MSR-MOE mean	0.442	0.405
	LSW-MOE	0.264	0.239
MLR	12 inputs	0.634	0.585
GBRT (proposed)	12 inputs	0.772	0.718

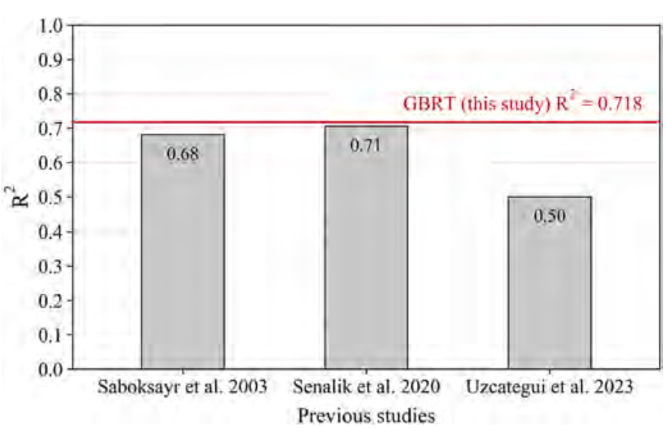

Fig. 13. Contextual comparison of reported R^2 values for UTS prediction models.

Aspen-UTS-Prediction-GUI).

7. Conclusions

This study developed and evaluated ensemble ML models for predicting the UTS of aspen dimension lumber. A total of 321 specimens were tested, and 12 input variables related to grade, geometry, physical properties, knot features, and stiffness measurements were collected. Four ensemble models, including RF, XGBoost, GBRT, and AdaBoost, were trained and tuned using grid search and five-fold cross-validation. The proposed GBRT model was further benchmarked against MOE-based empirical models and an MLR model using the same aspen dataset. Based on the results, the main findings are summarized as follows:

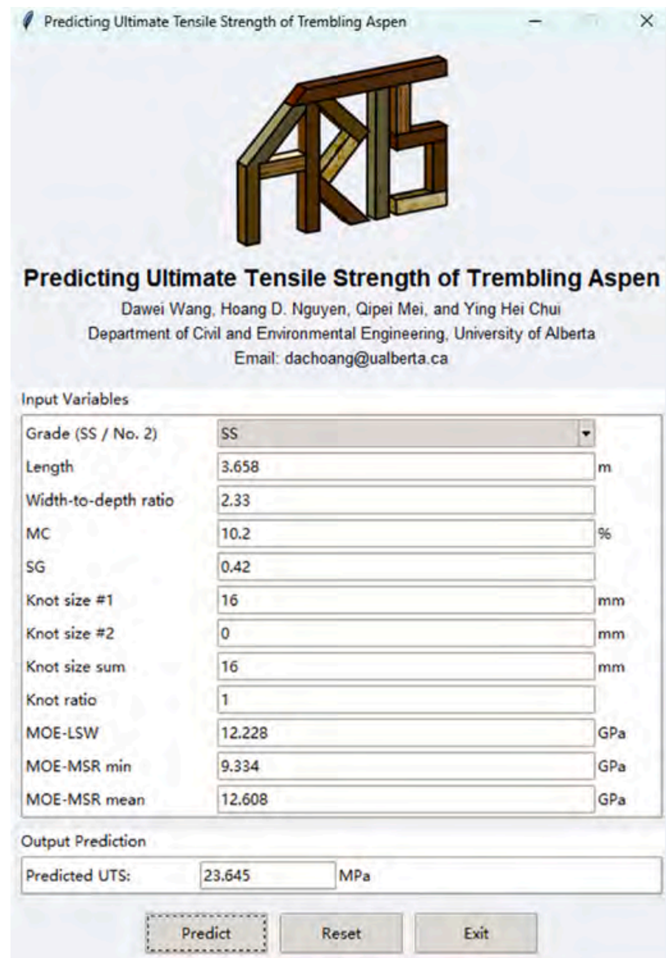

Fig. 14. GUI for the proposed GBRT model.

1. The ensemble ML models showed useful predictive performance for individual-specimen UTS prediction of aspen lumber. Among the four models, GBRT achieved the best overall performance, with a testing set R^2 of 0.718, RMSE of 7.047 MPa, MAPE of 0.238, and r of 0.866. XGBoost showed comparable performance with a testing set R^2 of 0.715, while AdaBoost and RF showed lower predictive accuracy.
2. The same-dataset model comparison showed that adding measurable variables beyond MOE improved UTS prediction. The best single-

variable empirical model was the power-law model based on MSR-MOE min., with a testing set R^2 of 0.487. The MLR model using the same 12 input variables increased the R^2 to 0.585, corresponding to an absolute increase of 0.098 and a relative increase of 20.1%. This result indicates that physical properties, grade, and knot-related variables provided useful information beyond stiffness alone.

3. The GBRT model further improved the prediction compared with the MLR model using the same 12 input variables. The testing set R^2 increased from 0.585 to 0.718, corresponding to an absolute increase of 0.133 and a relative increase of 22.7%. Compared with the best single-variable empirical model, GBRT increased the R^2 by 0.231, corresponding to a relative increase of 47.4%. These results suggest that nonlinear modelling provided additional predictive value for individual-specimen UTS prediction.
4. The SHAP analysis identified MSR-MOE min. as the most influential predictor of UTS, showing the importance of the weakest stiffness region along the specimen. MSR-MOE mean and lumber length also contributed to the model prediction. However, the effect of length should be interpreted within the present dataset because only two lengths were included.
5. The results indicate that the ML model developed in this study can support aspen-specific individual-specimen UTS prediction using stiffness, physical-property, and knot-related variables obtained through industry-oriented testing. A GUI was also developed based on the GBRT model to support practical use in research and industry. However, the model should be viewed as a prediction tool developed from the present dataset, rather than as a general grading model or design-value model.

Future work should focus on expanding the dataset with independent aspen lumber batches and broader grade representation to improve model generalization. External validation is also needed before the model can be used for routine industrial decision-making. Additional input variables, such as growth-ring characteristics or improved defect measurements, may also be considered if they can be collected efficiently under industry-oriented testing conditions.

CRediT authorship contribution statement

Dawei Wang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Hoang D. Nguyen:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Qipei Mei:** Conceptualization, Supervision, Writing – review & editing. **Ying Hei Chui:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors gratefully acknowledge the financial support provided by the National Research Council of Canada's (NRC) Collaborative Science, Technology and Innovation Program (CSTIP), Alberta Innovates' Bioindustrial Materials, Natural Sciences and Engineering Research Council (NSERC) of Canada through its Alliance Grant program (Grant number ALLRP 581074-22 and Grant number ALLRP 571090-2). The authors further wish to thank Wood Science and Technology Centre at the University of New Brunswick, Canada, for their valuable technical assistance.

Data availability

Data will be made available on request.

References

Alberta Agriculture and Forestry, 2021. Alberta's Forest Economy 2021: a Handbook of Public Economic and Socioeconomic Accounts. Government, of Alberta.
American Society for Testing and Materials (ASTM), 2022. Standard Test Methods of Static Tests of Lumber in Structural Sizes. ASTM International, Philadelphia, USA. https://doi.org/10.1520/D0198-22. ASTM D198-22).
American Society for Testing and Materials (ASTM), 2025. *Standard Methods for Establishing Structural Grades and Related Allowable Properties for Visually Graded Lumber* (ASTM D245-22). ASTM International, Philadelphia, USA. https://doi.org/10.1520/D0245-22.
Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., Ridella, S., 2012. The "K" in K-fold cross validation. ESANN 102, 441–446.
Arnold, D.H., Mauseth, J.D., 1999. Effects of environmental factors on development of wood. Am. J. Bot. 86 (3), 367–371. https://doi.org/10.2307/2656758.
As, N., Goker, Y., Dundar, T., 2006. Effect of knots on the physical and mechanical properties of Scots pine. Wood Res. 51 (3), 51–58.
Balatinecz, J.J., Kretschmann, D.E., 2001. Properties and utilization of poplar wood. In: Poplar Culture in North America. Press, NRC Research, pp. 277–291.
Bardenet, R., Brendel, M., Kégl, B., Sebag, M., 2013. Collaborative hyperparameter tuning. In: Proceedings of the 30th International Conference on Machine Learning. JMLR, pp. 199–207. In: https://proceedings.mlr.press/v28/bardenet13.html.
Barrett, J.D., Hejja, A., 1984. Compression strength of Canadian softwood lumber. Forintek Canada Corp., Western Laboratory.
Barrett, J.D., Lau, W., 1994. Canadian Lumber Properties. Canadian Wood Council.
Biechele, T., Chui, Y.H., Gong, M., 2011. Comparison of NDE techniques for assessing mechanical properties of unjointed and finger-jointed lumber. Holzforschung 65 (3), 449–456. https://doi.org/10.1515/hf.2011.038.
Boström, L., 1994. Machine strength grading. In: *Comparison of Four Different Systems* (SP Report 49). Swedish National Testing and Research Institute.
Boyle, D., 1998. Nutritional factors limiting the growth of Lentinula edodes and other white-rot fungi in wood. Soil Biol. Biochem. 30 (6), 817–823. https://doi.org/10.1016/S0038-0717(97)00159-4.
Brashaw, B., Bergman, R., 2021. Wood as a renewable and sustainable resource. In: FPL-GTR-282. USDA Forest Service, 1-1.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.
Briggert, A., Olsson, A., Oscarsson, J., 2020. Prediction of tensile strength of sawn timber: definitions and performance of indicating properties based on surface laser scanning and dynamic excitation. Mater. Struct. 53 (3), 54. https://doi.org/10.1617/s11527-020-01460-5.
Burdon, R.D., Thulin, I.J., 1966. In: The Improvement of Pinus Radiata: a Record of a Symposium Held at the Forest Research Institute, Rotorua, 24–25 November 1965 (F.R.I. Symposium No. 6). New Zealand Forest Service, Forest Research Institute.
Cameron, A.C., Windmeijer, F.A., 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. J. Econom. 77 (2), 329–342. https://doi.org/10.1016/S0304-4076(96)01818-0.
Canadian Wood Council (CWC), 2019. http://www.cwc.ca.
Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. https://doi.org/10.1145/2939672.2939785.
Cramer, S.M., Goodman, J.R., 1983. Model for stress analysis and strength prediction of lumber. Wood Fiber Sci. 15 (3), 338–349.
Cramer, S.M., Goodman, J.R., 1986. Failure modelling: a basis for strength prediction of lumber. Wood Fiber Sci. 18 (4), 446–459.
Cramer, S.M., McDonald, K.A., 1989. Predicting lumber tensile stiffness and strength with local grain angle measurements and failure analysis. Wood Fiber Sci. 21 (3), 393–410.
Dickmann, D.I., Isebrands, J.G., Eckenwalder, J.E., Richardson, J. (Eds.), 2001. Poplar Culture in North America. Press, NRC Research.
Divos, F., Tanaka, T., 1997. Lumber strength estimation by multiple regression. Holzforschung 51 (5), 467–471. https://doi.org/10.1515/hfsg.1997.51.5.467.
Esteghamati, M.Z., Gernay, T., Banerji, S., 2023. Evaluating fire resistance of timber columns using explainable machine learning models. Eng. Struct. 296, 116910. https://doi.org/10.1016/j.engstruct.2023.116910.
Fan, S., Wong, S.W., Zidek, J.V., 2023. Knots and their effect on the tensile strength of lumber: a case study. J. Qual. Technol. 55 (4), 510–522. https://doi.org/10.1080/00224065.2023.2180457.
Fathi, H., Nasir, V., Kazemirad, S., 2020. Prediction of the mechanical properties of wood using guided wave propagation and machine learning. Constr. Build. Mater. 262, 120848. https://doi.org/10.1016/j.conbuildmat.2020.120848.
Feng, Y., Mekhilef, S., Hui, D., Chow, C.L., Lau, D., 2024. Machine learning-assisted wood materials: applications and future prospects. Extreme Mech. Lett. 71, 102209. https://doi.org/10.1016/j.eml.2024.102209.
Fink, G., Kohler, J., 2014. Model for the prediction of the tensile strength and tensile stiffness of knot clusters within structural timber. Eur. J. Wood Wood Prod. 72 (3), 331–341. https://doi.org/10.1007/s00107-014-0781-0.
Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. J. Comput. Syst. Sci. 55 (1), 119–139. https://doi.org/10.1006/jcss.1997.1504.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232. https://doi.org/10.1214/aos/1013203451.
Gong, M., 2019. Lumber-based mass timber products in construction. In: Timber Buildings and Sustainability. IntechOpen. https://doi.org/10.5772/intechopen.85808.
Gong, M., 2021. Wood and engineered wood products: stress and deformation. In: Engineered Wood Products for Construction. IntechOpen. https://doi.org/10.5772/intechopen.101199.
Green, D.W., Kretschmann, D.E., 1991. Lumber property relationships for engineering design standards. Wood Fiber Sci. 23 (3), 436–456.
Guntekin, E., Emiroglu, Z.G., Yilmaz, T., 2013. Prediction of bending properties for Turkish red pine (Pinus brutia Ten.) lumber using stress wave method. Bioresources 8 (1), 231–237. https://doi.org/10.15376/biores.8.1.231-237.
Haftkhani, A.R., Abdoli, F., Sepehr, A., Mohebby, B., 2021. Regression and ANN models for predicting MOR and MOE of heat-treated fir wood. J. Build. Eng. 42, 102788. https://doi.org/10.1016/j.jobe.2021.102788.
Huang, F., Teng, Z., Guo, Z., Catani, F., Huang, J., 2023. Uncertainties of landslide susceptibility prediction: influences of different spatial resolutions, machine learning models and proportions of training and testing dataset. Rock Mech. Bull. 2 (1), 100028. https://doi.org/10.1016/j.rockmb.2023.100028.
Kaiserlik, J.H., Pellerin, R.F., 1977. Stress wave attenuation as an indicator of lumber strength. USDA Forest Service Research Note FPL-0252.
Lam, F., Varoğlu, E., 1991. Variation of tensile strength along the length of lumber: part 1: experimental. Wood Sci. Technol. 25 (5), 351–359. https://doi.org/10.1007/BF00226174.
Lamb, F.M., 1967. *Aspen Wood Characteristics, Properties and Uses: a Review of Recent Literature* (Research Paper NC-13). US Department of Agriculture, Forest Service, North Central Forest Experiment Station.
Lei, Y.C., Zhang, S.Y., Jiang, Z., 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. https://doi.org/10.1007/s00226-004-0269-x.
Li, Z., 2022. Extracting spatial effects from machine learning model using local interpretation method: an example of SHAP and XGBoost. Comput. Environ. Urban Syst. 96, 101845. https://doi.org/10.1016/j.compenvurbsys.2022.101845.
Lim, S., Chi, S., 2019. XGBoost application on bridge management systems for proactive damage estimation. Adv. Eng. Inform. 41, 100922. https://doi.org/10.1016/j.aei.2019.100922.
Lundberg, S.M., Erion, G.G., Lee, S.I., 2018. Consistent individualized feature attribution for tree ensembles. arXiv preprint arXiv:1802.03888. https://doi.org/10.48550/arXiv.1802.03888.
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30, 4765–4774. In: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.
Mackes, K.H., Lynch, D.L., 2001. The effect of aspen wood characteristics and properties on utilization. In: Sustaining Aspen in Western Landscapes: Symposium Proceedings. USDA Forest Service, Rocky Mountain Research Station, pp. 429–440.
Nasir, V., Fathi, H., Fallah, A., Kazemirad, S., Sassani, F., Antov, P., 2021. Prediction of mechanical properties of artificially weathered wood by color change and machine learning. Materials 14 (21), 6314. https://doi.org/10.3390/ma14216314.
National Lumber Grades Authority (NLGA), 2019. Grading Rules for Canadian Lumber. National Lumber Grades Authority.
Natural Resources Canada, 2023. The state of Canada's forests: annual report 2022. https://natural-resources.canada.ca/our-natural-resources/forests/state-canadas-forests-report/16496.
Natural Resources Canada, 2025. Canada's softwood lumber industry. https://natural-resources.canada.ca/forest-forestry/forest-industry-trade/canada-s-softwood-lumber-industry.
Nazerian, M., Razavi, S.A., Partovinia, A., Vatankhah, E., Razmpour, Z., 2020. Prediction of the bending strength of a laminated veneer lumber (LVL) using an artificial neural network. Mech. Compos. Mater. 56 (5), 649–664. https://doi.org/10.1007/s11029-020-09911-4.
Newlin, J.A., Johnson, R.P.A., 1924. Structural timbers, defects and their influence on strength. USDA Bulletin No. 1086.
Nguyen, H.D., Dao, N.D., Shin, M., 2021a. Prediction of seismic drift responses of planar steel moment frames using artificial neural network and extreme gradient boosting. Eng. Struct. 242, 112518. https://doi.org/10.1016/j.engstruct.2021.112518.
Nguyen, H.D., Dao, N.D., Shin, M., 2022a. Machine learning-based prediction for maximum displacement of seismic isolation systems. J. Build. Eng. 51, 104251. https://doi.org/10.1016/j.jobe.2022.104251.
Nguyen, H.D., LaFave, J.M., Lee, Y.J., Shin, M., 2022b. Rapid seismic damage-state assessment of steel moment frames using machine learning. Eng. Struct. 252, 113737. https://doi.org/10.1016/j.engstruct.2021.113737.
Nguyen, H.D., Lee, Y.J., LaFave, J.M., Shin, M., 2023. Seismic fragility analysis of steel moment frames using machine learning models. Eng. Appl. Artif. Intell. 126, 106976. https://doi.org/10.1016/j.engappai.2023.106976.
Nguyen, H.D., Truong, G.T., Shin, M., 2021b. Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. Eng. Struct. 235, 112067. https://doi.org/10.1016/j.engstruct.2021.112067.
Niederwestberg, J., Weijs, H.J.J., Vonk, N.H., de Jong, J., 2025. Structural potential of reclaimed and local timber as new resources in the Netherlands. In: Rinke, R., Hvejsel, F. (Eds.), Structures and Architecture: Restructure Rematerialize Rethink Reuse. CRC Press, pp. 137–144. https://doi.org/10.1201/9781003658641-17.
Panshin, A.J., de Zeeuw, C., 1981. Textbook of Wood Technology, fourth ed. McGraw-Hill.
Ponce-Bobadilla, A.V., Schmitt, V., Maier, C.S., Mensing, S., Stodtmann, S., 2024. Practical guide to SHAP analysis: explaining supervised machine learning model predictions in drug development. Clin. Transl. Sci. 17 (11), e70056. https://doi.org/10.1111/cts.70056.
Prettenhofer, P., Louppe, G., 2014. Gradient boosted regression trees in scikit-learn. PyData 2014 Conference.
Qiu, L.P., Zhu, E.C., Van de Kuilen, J.W.G., 2014. Modeling crack propagation in wood by extended finite element method. Eur. J. Wood Wood Prod. 72 (2), 273–283. https://doi.org/10.1007/s00107-013-0773-5.
Rahimi, S., Avramidis, S., 2022. Predicting moisture content in kiln dried timbers using machine learning. Eur. J. Wood Wood Prod. 80 (3), 681–692. https://doi.org/10.1007/s00107-022-01794-7.
Raschka, S., 2018. Model evaluation, model selection, and algorithm selection in machine learning. arXiv preprint arXiv:1811.12808. https://doi.org/10.48550/arXiv.1811.12808.
Reid, D.E., Young, S., Tong, Q., Zhang, S.Y., Morris, D.M., 2009. Lumber grade yield, and value of plantation-grown black spruce from 3 stands in northwestern Ontario. For. Chron. 85 (4), 609–617. https://doi.org/10.5558/tfc85609-4.
Roblot, G., Bléron, L., Mériaudeau, F., Marchal, R., 2010. Automatic computation of the knot area ratio for machine strength grading of Douglas-fir and Spruce timber. Eur. J. Environ. Civil Eng. 14 (10), 1317–1332. https://doi.org/10.1080/19648189.2010.9693296.
Rodríguez-Pérez, R., Bajorath, J., 2020. Interpretation of machine learning models using Shapley values: application to compound potency and multi-target activity predictions. J. Comput. Aided Mol. Des. 34 (10), 1013–1026. https://doi.org/10.1007/s10822-020-00314-0.
Saboksayr, H., Saravi, A., Lawrence, P.D., Lam, F., 2003. Implementation of a neural network based system for estimating the strength of a board using mixed signals. In: 2003 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM 2003), 1. IEEE, pp. 342–347. https://doi.org/10.1109/PACRIM.2003.1235787.
Samson, M., Blanchet, L., 1992. Effect of knots on the flatwise bending stiffness of lumber members. Holz als Roh- Werkst. 50 (4), 148–152. https://doi.org/10.1007/BF02663256.
Schajer, G.S., 2001. Lumber strength grading using X-ray scanning. For. Prod. J. 51 (1), 43–50.
Senalik, C.A., França, F.J.N., Seale, R.D., Ross, R.J., Shmulsky, R., 2020. Estimating lumber properties with acoustic-based technologies---Part 2: ultimate tension stress estimation from time- and frequency-domain parameters. Wood Fiber Sci. 52 (4), 390–399. https://doi.org/10.22382/wfs-2020-037.
Smith, I., Landis, E., Gong, M., 2003. Fracture and Fatigue in Wood. John Wiley & Sons.
Speiser, J.L., Miller, M.E., Tooze, J., Ip, E., 2019. A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst. Appl. 134, 93–101. https://doi.org/10.1016/j.eswa.2019.05.028.
Steele, P.H., Cooper, J.E., 2003. Estimating lumber strength with radio frequency scanning. Comput. Electron. Agric. 41 (1–3), 77–83. https://doi.org/10.1016/S0168-1699(03)00043-7.
Uzcategui, M.G., França, F.J.N., Seale, R.D., Ross, R.J., Senalik, C.A., 2023. Assessing 2 by 6 Southern pine lumber flexural and tensile properties with nondestructive techniques. For. Prod. J. 73 (1), 75–81. https://doi.org/10.13073/FPJ-D-22-00055.
Wang, D., Zhang, M., Gong, M., Chui, Y.H., 2024. Evaluation of major physical and mechanical properties of trembling aspen lumber. Materials 17 (12), 2952. https://doi.org/10.3390/ma17122952.
Wang, Q., Shi, G., Chan-Halbrendt, C., 2004. Market potential for fine furniture manufactured from low-grade hardwood. For. Prod. J. 54 (5), 49–55.
Wang, X.Q., Chen, P., Chow, C.L., Lau, D., 2023. Artificial-intelligence-led revolution of construction materials: from molecules to Industry 4.0. Matter 6 (6), 1831–1859. https://doi.org/10.1016/j.matt.2023.04.016.
Wang, Y.T., Lam, F., Barrett, J.D., 1995. Simulation of correlated modulus of elasticity and compressive strength of lumber with gain factor. Probab. Eng. Mech. 10 (2), 63–71. https://doi.org/10.1016/0266-8920(94)00009-A.
Wright, S., Dahlen, J., Montes, C., Eberhardt, T.L., 2019. Quantifying knots by image analysis and modeling their effects on the mechanical properties of loblolly pine lumber. Eur. J. Wood Wood Prod. 77 (5), 903–917. https://doi.org/10.1007/s00107-019-01441-8.
Xin, Z., Ke, D., Zhang, H., Yu, Y., Liu, F., 2022. Non-destructive evaluating the density and mechanical properties of ancient timber members based on machine learning approach. Constr. Build. Mater. 341, 127855. https://doi.org/10.1016/j.conbuildmat.2022.127855.
Yamamoto, H., Sakagami, H., Kijidani, Y., Matsumura, J., 2013. Dependence of microcrack behavior in wood on moisture content during drying. Adv. Mater. Sci. Eng. 2013, 802639. https://doi.org/10.1155/2013/802639.
Youngquist, J.A., Spelter, H., 1990. Aspen wood products utilization: impact of the Lake states composites industry. In: Proceedings of the Aspen Symposium 89. USDA Forest Service, North Central Forest Experiment Station, pp. 91–102.
Zobel, B.J., Van Buijtenen, J.P., 2012. Wood Variation: Its Causes and Control. Springer Science & Business Media. https://doi.org/10.1007/978-3-642-74069-5.