

Modulus of Rupture Prediction in *Pinus sylvestris* with Multivariate Models Constructed with Resonance, Ultrasound, and Wood Heterogeneity Variables

Álvaro Fernández-Serrano * and Antonio Villasante

Multivariate models with multiple linear regression (MLR), artificial neural network (ANN), and k-nearest neighbors (KNN) were developed to predict the modulus of rupture of *Pinus sylvestris* structural timber. The aim of this study was to develop and compare these models obtained from resonance and ultrasound tests, static modulus of elasticity tests, and different measured wood feature. Resonance tests were performed in the three vibration modes (edgewise, flatwise, and longitudinal) to obtain the fundamental resonant frequencies. To compare the goodness-of-fit of the different models, the 10-fold cross-validation method was used, which proved to be an adequate strategy to avoid overfitting. The variable with the best predictive capacity of the modulus of rupture was knottiness. The error was notably lower in the multivariate than the univariate models. The ANN and KNN algorithms showed no improvement over the MLR. The most suitable MLR for prediction of the modulus of rupture was the model with four variables: knottiness, edgewise dynamic modulus of elasticity, velocity of ultrasounds, and longitudinal resonant frequency.

DOI: 10.15376/biores.17.1.1106-1119

Keywords: Cross-validation; k-Nearest neighbors; Neural network; Non-destructive testing; Resonant frequency

Contact information: Department of Agricultural and Forest Engineering, University of Lleida, Av. Rovira Roure, 191, 25198 Lleida, Spain; *Corresponding author: alvaro.fernandezserrano@udl.cat

INTRODUCTION

Resonant frequencies have been employed to determine bending strength of *Pinus sylvestris* wood. Most of the works (Arriaga *et al.* 2012; Hassan *et al.* 2013; Aira *et al.* 2019) only used one vibration mode (edgewise, flatwise, or longitudinal) to predict the modulus of rupture (MOR). Only a few used more than one vibration mode (Hassan *et al.* 2013; Arriaga *et al.* 2014; Dahlen *et al.* 2018). Likewise, few studies have used a combination of vibration and ultrasound techniques (Halabe *et al.* 1997; Hassan *et al.* 2013). Some authors have analysed the effect on the MOR of features of sawn timber such as knots (Arriaga *et al.* 2014; França *et al.* 2019; Villasante *et al.* 2019), the slope of grain (Arriaga *et al.* 2014), or the rate of growth (Martins *et al.* 2017; França *et al.* 2019).

Some studies have taken into account the combined effect of different variables, while using multiple linear regression (MLR) for MOR prediction, with different pine species (Arriaga *et al.* 2012; Martins *et al.* 2017; Villasante *et al.* 2019). Usually, the coefficient of determination (R^2) is used to compare the model's capacity to predict MOR. However, some authors proposed use of the root-mean-square error (RMSE) as a more suitable measure of the fit of the model (Pommier *et al.* 2013; Yang *et al.* 2017; Villasante *et al.* 2019).

Most of the works consulted employed linear regressions, though some authors proposed nonlinear regressions based on artificial neural networks (ANNs). In some cases, prediction of the mechanical properties of wood through ANN has given good results (Mansfield *et al.* 2007; García Esteban *et al.* 2009; Fathi *et al.* 2020); however, in others no improvements over MLR-based models were found when using ANN-based models (Tanaka *et al.* 1996; Villasante *et al.* 2019)

Multivariate analysis can be affected by overfitting, producing misleading models valid only for the used dataset. To avoid overfitting (Lever *et al.* 2016), a few authors employed the cross-validation method in studies on pine (Villasante *et al.* 2019), chestnut (Vega *et al.* 2012), poplar (Fathi *et al.* 2020), and oak (Faydi *et al.* 2017).

The aim of this study was to develop and compare multivariate (linear, and non-linear) models for MOR prediction of a batch of *Pinus sylvestris* timber. These models were developed based on the values of different variables obtained in resonance and ultrasound tests and different measured features (blue stain, rate of growth, slope of grain, resin pockets, knottiness, and waness). Models were compared using RMSE values obtained with the cross-validation method. This study also aimed to find the variables that best explained the MOR. Grading of *Pinus sylvestris* timber from Montsec mountains (Spain) was not an objective of this work.

EXPERIMENTAL

Materials

Analyses were undertaken with samples of *Pinus sylvestris* from Montsec Range (NE of Spain) obtained from a local sawmill (Lleida, Spain). A total of 69 sawn wood samples were obtained by random sampling of timber stored at the sawmill. The nominal sample size was 70 mm × 100 mm × 2000 mm with different growth rings orientations. At the laboratory, 12 samples with bark pockets or rot were rejected, leaving a final total of 57 samples. The samples were stored in the test laboratory and subjected to periodic moisture content analyses using a wood moisture meter (Hydromette HB 30, GANN, Gerlingen, Germany) following EN 13183-2 (2002). The samples were removed from the drying process when they reached a constant weight ($\pm 0.1\%$ in 6 h), in accordance with EN 408:2011+A1 (2012). Final moisture was calculated using the oven dry method following EN 13183-1 (2002), with 20-mm-thick slices obtained from similar specimens that were not included in the test batch.

Methods

Resonance tests

The resonance tests were conducted following ASTM E1876-15 (2015). The specimens were suspended using elastic cords situated in the nodes corresponding to the fundamental flexural resonance (Fig. 1) and were tapped in the centre of the edge, face, and end to obtain the three vibration modes: edgewise, flatwise, and longitudinal, respectively. A 22.7 g impulser was used, consisting of a 230 mm long × 4 mm diameter wooden handle with a 26-mm diameter glass marble at the tip. The vibration was registered using a microphone with cardioid polar pattern and frequency range of 20 Hz to 20 kHz (Rode NT-USB, Rode Microphones, Silverwater, Australia). The signal obtained was recorded and analysed with Audacity® (2015) (Audacity Team, Pittsburgh, PA, USA). A project sampling rate of 384 kHz and resolution of 24 bits were used for the recording. The

frequency spectrum was analysed with the Hann function to obtain the fundamental resonant frequencies for the longitudinal (f_{LV}), flatwise (f_{FV}), and edgewise (f_{EV}) vibration modes.

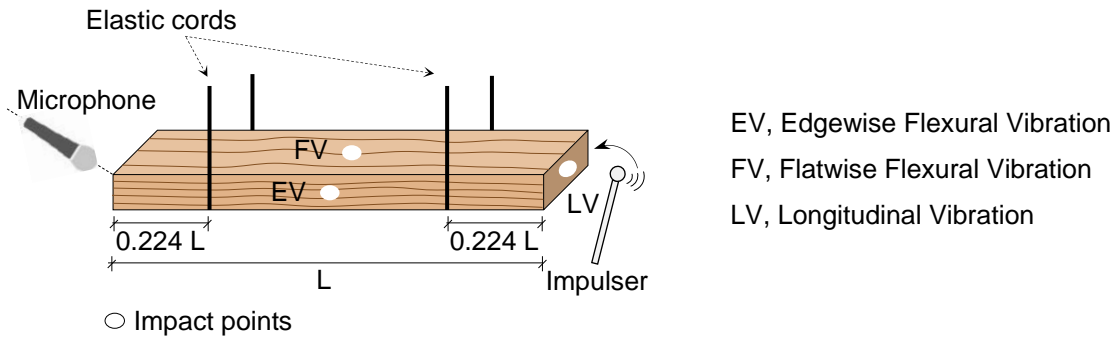


Fig. 1. Scheme of the resonant frequencies tests

On the basis of the frequencies, the corresponding MOE_{dyn} were obtained through the following Eqs. 1 through 3 (Weaver *et al.* 1990),

$$MOE_{LV} = 4 \times (f_{LV})^2 \times L^2 \times \rho \quad (1)$$

$$MOE_{FV} = [48 \times \pi^2 \times \rho \times (f_{FV})^2 \times L^4] / [4.73^4 \times b^2] \quad (2)$$

$$MOE_{EV} = [48 \times \pi^2 \times \rho \times (f_{EV})^2 \times L^4] / [4.73^4 \times h^2] \quad (3)$$

where MOE_{LV} , MOE_{FV} , and MOE_{EV} are the dynamic modulus of elasticity (all expressed in MPa) for each type of vibration (longitudinal, flatwise, and edgewise, respectively), f_{LV} , f_{FV} , and f_{EV} are the fundamental resonant frequencies (kHz) for each type of vibration, ρ is the density ($g \cdot mm^{-3}$) of the sample, and L , h , and b are the actual length, width, and thickness (all expressed in mm) of each sample, respectively. The effect of shear in the MOE_{dyn} calculation was not taken into account because it had a reduced influence, as the ratios L/b and L/h were equal to or greater than 20 (Arriaga *et al.* 2014).

Ultrasound tests

The ultrasound tests were performed with a Sylvatest Duo ultrasound device with a frequency of 22 kHz (Concept Bois Structure, Montandon, France). The transducers were placed at the ends of the samples. Information about species, length, and MC of each sample had to be introduced in the ultrasounds device. After the test, the device's screen showed the velocity of ultrasound waves (V_U), the attenuation of the ultrasonic wave ($peak_U$), the modulus of elasticity by ultrasound (MOE_U), and the modulus of rupture by ultrasound (MOR_U).

Features and biological degradations

The samples were measured and weighed to obtain their dimensions and density following EN 408:2011+A1 (2012). The blue stain (BS), rate of growth (RG), slope of grain (SLG), and resin pockets were also measured in accordance with the procedure outlined in EN 1309-3 (2018).

Two different procedures were used for knot measurement. The first procedure (Fig. 2a) was based on the knot area ratio (KAR). The second one (Fig. 2b) was based on

Annex A of EN 1309-3 (2018). The KAR indicates the proportion of the cross-section occupied by knots (Walker 1993), while MKAR (Fig. 2a) indicates the proportion of the margin cross-section (outer quarters) occupied by knots (Lam *et al.* 2005). The MKAR in compression (MKAR_C) refers to the outer quarter subjected to compression, and the MKAR in tension (MKAR_T) denotes to the outer quarter subjected to tension. The KAR-based measurements were made in the centre third of the length of the sample, in the section where the KAR values were higher.

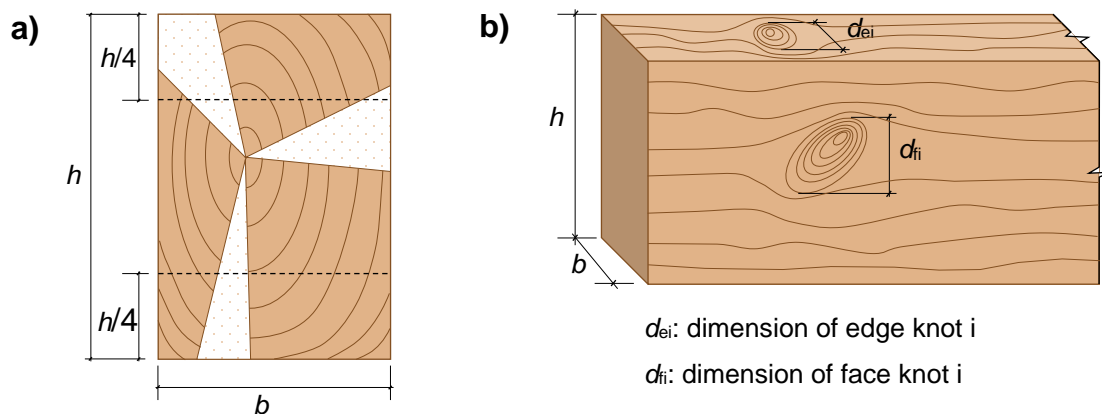


Fig. 2. Scheme for the measurement of knottiness: a) Cross-section, KAR, MKAR, MKAR_C, and MKAR_T. The dotted areas represent knots; b) Measure of knots following Annex A of EN 1309 3 (2018)

The knottiness was also measured following the procedure set out in Annex A of EN 1309-3 (2018), measuring each knot in the direction perpendicular to the length of the sample (Fig. 2b). For the variable $knot_T$, the sum of the measurements of knots in the four faces was calculated. For the variable $knot_{1/3}$, only the knots situated in the centre third of length of the sample were included.

Two procedures were also followed to measure the waness. The procedure described in EN 1309-3 (2018) was performed first (see Fig. 3). The wane was evaluated as the highest ratio of the dimension of the wane on the face with respect to the complete dimension of the face (WN_F). The same procedure was performed with respect to the edge (WN_E). The mean of these two values was also calculated (WN_M). These same measurements were also taken exclusively in the centre third of the length of the sample ($WN_{F1/3}$, $WN_{E1/3}$, and $WN_{M1/3}$). The second procedure was to measure the waness with the loss of volume with respect to the theoretical volume without waness (WN_{VOL}).

Static bending test

The samples were subjected to a four point static bending test following EN 408:2011+A1 (2012). This test was performed to obtain the global static modulus of elasticity (MOEs) and the MOR. The tests were completed with a 50-kN universal testing machine (Cohiner, Lleida, Spain) and data acquisition with LabVIEW 7.1 (National Instruments, Austin, TX, USA). The distance between supports was 1.8 m and the distance between load points was 0.6 m. A displacement transducer was placed at the midpoint between the two supports. After the bending test, a 20-mm-thick slice was obtained close to the area of rupture to determine moisture content using the oven dry method following EN 13183-1 (2002).

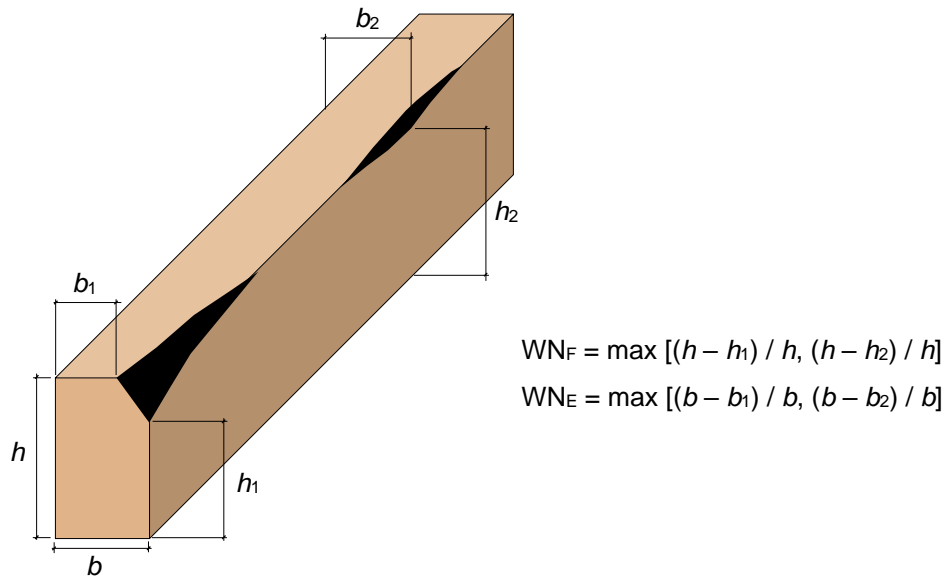


Fig. 3. Measure of wanes following EN 1309-3 (2018)

Statistical analyses

Simple linear regressions (SLR) were developed to predict the MOR from each of the studied variables. The significance level for each variable was calculated with F-test. In the present work, the RMSE (Eq. 4) was used to compare the fit of the MOR prediction models.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\widehat{MOR}_i - MOR_i)^2}{n}} \quad (4)$$

where \widehat{MOR}_i is the predicted value from non-destructive testing (NDT), MOR_i is the observed value with EN 408:2011+A1 (2012), and n the number of observations.

To evaluate the usefulness of a prediction model, it is more important to know the precision of the values that it generates (RMSE) than to quantify the variability (R^2) of the predicted values (Mansfield *et al.* 2007; Alexander *et al.* 2015). The R^2 value obtained with the whole dataset was used exclusively to compare the results with those of other authors. The MOR prediction was also carried out using multivariate models with different algorithms: multiple linear regression (MLR), artificial neural network (ANN), and k -nearest neighbors (KNN). For the multivariate algorithms, the variables that did not present a statistically significant SLR were discarded. Similarly, only one knottiness and one wane variable were selected.

For the MLR, a greedy selection method was applied using the Akaike information metric to reduce the number of variables (Frank *et al.* 2016). A stepwise regression method was then applied, removing one-by-one the variables that caused the lowest increase in the RMSE until a univariate model was obtained. The predictive capacity of the three multivariate algorithms was compared for each group of variables.

The ANN prediction model was generated with the WEKA software (Waikato University, version 3.6, Hamilton, New Zealand). A multilayer perceptron was used with sigmoid nodes and learning by backpropagation (Frank *et al.* 2016). The ANN parameters

were adjusted in a prior test (learning rate, training time, momentum applied to the weights, and number of neurons).

Finally, a prediction model was also generated using a KNN model with the IBk algorithm in WEKA software. The KNN model is a non-parametric classification method that constructs a regression from the weighted average of the values of KNN (Frank *et al.* 2016). The KNN parameters were adjusted in a prior test (number of neighbor values and distance weighting method).

The prediction models generated for a dataset can be affected by overfitting. This effect appears when the model is fitted even to the noise of the sample, confusing the noise with the internal structure of the model (Lever *et al.* 2016). For small datasets, the K-fold cross-validation can help avoid overfitting (Lever *et al.* 2016). In this respect, according to Tetko *et al.* (1995), the cross-validation method avoids overfitting in complex predictive models.

The 10-fold cross-validation method (Hashim *et al.* 2016; Faydi *et al.* 2017; Villasante *et al.* 2019) was used to calculate the RMSE of the models. The samples were randomly distributed into ten groups or folds, using each group to validate the model generated with the remaining nine. Using the WEKA software, this procedure was repeated 5 times to obtain 50 RMSE values of each model (MLR, ANN, and KNN). The nonparametric Kruskal-Wallis test was used to detect significant differences between the RMSEs of the models. In those cases in which significant differences were detected, a *post hoc* analysis was performed with Dunn's test and Bonferroni adjustment (Brunetti *et al.* 2016). A significance level of 0.05 was applied. The analysis of significant differences was performed using R-software (The R Foundation, version 3.6.1, Vienna, Austria).

RESULTS AND DISCUSSION

The values obtained for the measured variables are shown in Table 1. The mean MOR value (40.0 MPa) was similar to those obtained in previous studies (between 35 and 45 MPa) for *Pinus sylvestris* (Arriaga *et al.* 2012; Villasante *et al.* 2019) and other pine species, including *Pinus radiata* (Arriaga *et al.* 2014), *Pinus nigra* (Íñiguez González *et al.* 2007; Arriaga *et al.* 2012), and southern pine (Yang *et al.* 2017).

Simple Linear Regression

For most of the variables in this study, a statistically significant correlation, though weak, was found for the MOR prediction while using a SLR (Table 1). In this work, the R^2 value was used only to facilitate a comparison with results obtained by other authors in previous studies. The R^2 values for the variables of knottiness were similar or slightly higher than those observed in previous studies with different pine species (Conde García *et al.* 2007; Arriaga *et al.* 2014; França *et al.* 2019). In addition, the R^2 values for the different MOE_{dyn} were similar to those found by other authors in different pine species (Pommier *et al.* 2013; Arriaga *et al.* 2014; França *et al.* 2019).

Table 2 contains the RMSE results obtained through the 10-fold cross-validation method with 5 repetitions. The variables that did not present a statistically significant SLR were excluded. In this sense, the 29 initial variables were reduced to 15.

For knot-based MOR prediction, no significant differences were observed between the RMSE of the different measurement methods employed (KAR based or EN 1309-3 (2018)-based), except for MKAR_c that gave a worse predictive capacity than the rest

(Table 1). Conde Garcia *et al.* (2007) in *Pinus sylvestris* and *Pinus nigra* and França *et al.* (2019) in southern pine also found small differences in the R^2 values between different knottiness variables and MOR. In the present study, the variable $knot_T$ was chosen because it is easier to measure than any of the KAR-based variables. With respect to the different methods used to measure the influence of waness, no statistically significant differences were observed between the RMSE values in MOR prediction (Table 1). The variable WN_M was chosen because it was easier to measure than the loss of volume.

Table 1. Summary of the Study Variables (Regression with the Whole Dataset)

Variable	Units	Mean Value	CV (%)	Linear Regression Model (MPa)	R^2	SSD ⁽¹⁾
Features						
SLG	%	5.2	66.6	MOR = -1.702·SLG + 48.90	0.14**	
RG	mm	3.2	23.8	MOR = -11.29·RG + 76.70	0.31***	
BS	%	10.4	147.0	MOR = -0.085·BS + 40.920	0.01 ^{ns}	
WN_E	mm·mm ⁻¹	0.08	179.1	MOR = 44.42· WN_E + 36.629	0.15**	w-a
$WN_{E1/3}$	mm·mm ⁻¹	0.04	236.5	MOR = 48.15· $WN_{E1/3}$ + 37.978	0.10*	w-a
WN_F	mm·mm ⁻¹	0.05	181.3	MOR = 61.56· WN_F + 37.200	0.11*	w-a
$WN_{F1/3}$	mm·mm ⁻¹	0.03	242.4	MOR = 74.17· $WN_{F1/3}$ + 38.054	0.09*	w-a
WN_M	mm·mm ⁻¹	0.06	176.1	MOR = 54.24· WN_M + 36.706	0.14**	w-a
$WN_{M1/3}$	mm·mm ⁻¹	0.03	235.9	MOR = 59.81· $WN_{M1/3}$ + 37.958	0.10*	w-a
WN_{VOL}	‰	1.17	282.8	MOR = 2.068· WN_{VOL} + 37.615	0.19***	w-a
resin	mm	21.7	225.3	MOR = -0.072·RES + 41.592	0.05 ^{ns}	
$knot_T$	mm	283.4	59.1	MOR = -0.066· $knot_T$ + 58.826	0.51***	k-a
$knot_{1/3}$	mm	77.9	79.1	MOR = -0.169· $knot_{1/3}$ + 53.219	0.45***	k-a
KAR	mm ² ·mm ⁻²	0.25	70.1	MOR = -64.00·KAR + 56.090	0.52***	k-a
MKAR	mm ² ·mm ⁻²	0.25	80.1	MOR = -56.61·MKAR + 54.177	0.53***	k-a
MKAR _C	mm ² ·mm ⁻²	0.23	114.5	MOR = -26.25·MKAR _C + 46.115	0.20***	k-b
MKAR _T	mm ² ·mm ⁻²	0.27	95.2	MOR = -41.34·MKAR _T + 51.070	0.45***	k-a
ρ	kg·m ⁻³	549.5	7.1	MOR = 0.120· ρ - 25.758	0.09*	
Simple Vibration Variables						
V_U	m·s ⁻¹	4672	11.9	MOR = 0.010· V_U - 4.390	0.12**	
peak _U	mV	406.8	25.2	MOR = 0.010·peak _U + 35.974	0.004 ^{ns}	
f_{LV}	Hz	999.9	13.0	MOR = 0.063· f_{LV} - 22.712	0.27***	
f_{FV}	Hz	69.6	12.8	MOR = 0.941· f_{FV} - 25.431	0.29***	
f_{EV}	Hz	99.0	11.6	MOR = 0.749· f_{EV} - 34.080	0.30***	
Compound Variables						
MOE _U	MPa	7683	32.6	MOR = 0.002·MOE _U + 24.252	0.11*	
MOR _U	MPa	15.0	50.7	MOR = 0.708·MOR _U + 29.415	0.12**	
MOE _{LV}	MPa	8887	25.3	MOR = 0.004·MOE _{LV} + 0.799	0.40***	
MOE _{FV}	MPa	8810	25.3	MOR = 0.004·MOE _{FV} + 1.139	0.40***	
MOE _{EV}	MPa	8739	23.6	MOR = 0.005·MOE _{EV} - 3.844	0.44***	
MOE _S	MPa	7701	23.9	MOR = 0.006·MOE _S - 3.216	0.44***	
MOR	MPa	40.0	39.0			

⁽¹⁾Statistical Significant Differences. The same letter (a or b) indicates there are no statistically significant differences between the RMSE values in MOR prediction based on waness (w) or knots (k)

*** significant at 0.001 level, ** significant at 0.01 level, * significant at 0.05 level, ns not significant (F-test)

SLG: slope of grain; RG: rate of growth; BS: blue stain; WN: waness; KAR: knot area ratio; MKAR: margin knot area ratio; ρ : density; V_U : velocity of the ultrasounds; peak_U: attenuation of the ultrasounds; f : frequency

With respect to the variables included in the group of features and biological degradations, it was found that knot_T and RG provided the lower RMSE values in MOR prediction (Table 2). Various authors have also found in different conifers that knottiness is a good MOR predictor, better than density (Conde García *et al.* 2007) and better than SLG (Zhou and Smith 1991). According to Zhou and Smith (1991), RG was also a good MOR predictor. However, other studies conducted with different pine species found that density was a better MOR predictor than other characteristics such as knottiness and RG (Carballo *et al.* 2009; França *et al.* 2019). The RMSE values obtained for RG (12.8 MPa) and for density (14.9 MPa) are higher than those obtained by Hautamäki *et al.* (2014) in *Pinus sylvestris* (9.8 MPa and 10.2 MPa, respectively). These differences may be due to the heterogeneity of timber from such distant sources. It is notable, however, that the RMSE of knot_T (10.6 MPa) was practically the same as that obtained by Hautamäki *et al.* (2014) for KAR (10.4 MPa).

Table 2. RMSE of the Simple Linear Regressions for MOR Prediction (10-fold cross-validation, 5 repetitions)

Variable	RMSE (MPa)
Features Group	
SLG	14.07 ^b
RG	12.78 ^{ab}
WN _M	14.54 ^b
knot_T	10.59 ^a
ρ	14.86 ^b
Simple Vibration Variables Group	
V_U	14.40 ^a
f_{LV}	13.18 ^a
f_{FV}	13.10 ^a
f_{EV}	12.97 ^a
Compound Variables Group	
MOE _U	14.54 ^c
MOR _U	14.44 ^{bc}
MOE _{LV}	11.95 ^{ab}
MOE _{FV}	12.11 ^{ab}
MOE _{EV}	11.66 ^a
MOE _S	11.71 ^a

Within each group of variables, the same letter indicates there are no statistically significant differences (Kruskal-Wallis test, Dunn's test and Bonferroni adjustment).
 SLG: slope of grain; RG: rate of growth; WN: waness; ρ : density; V_U : velocity of the ultrasounds; f : frequency

In the simple vibration variables for MOR prediction, no statistically significant differences were found between the RMSE values obtained from V_U and the three resonance frequencies (f_{LV} , f_{FV} , and f_{EV}). In the compound variables, it was found that the ultrasound-based variables (MOE_U and MOR_U) had the highest errors, while MOE_S and MOE_{EV} gave the best MOR predictions. The RMSE value obtained for MOR prediction with MOE_S (11.71 MPa) was similar to those obtained by Pommier *et al.* (2013) in *Pinus pinaster* (13.9 MPa), Yang *et al.* (2017) in southern pine (14.65 MPa), and Hautamäki *et*

al. (2014) in *Pinus sylvestris* (8.20 MPa). The RMSE values for MOE_{dyn} calculated on the basis of SLR with the resonance tests (MOE_{LV} , MOE_{FV} , and MOE_{EV}) were between 11.66 MPa and 12.11 MPa, which were similar to those obtained by Pommier *et al.* (2013) and by Yang *et al.* (2017) in different pine species (12.5 MPa and 14.65 MPa, respectively).

Multiple Linear Regression

Including different variables should improve the fit of the model. A comparison of the RMSE values obtained in the univariate and multivariate linear regressions is shown in Table 3. To show overfitting, the RMSE calculated with the whole dataset, without cross-validation, is also provided in the table. It can be seen that this overfitted RMSE decreases constantly as the number of variables increases. However, this decrease is the result of the model being fitted to the noise of the dataset and therefore becoming an overfitted model. This can be verified in the fact that the RMSE calculated with 10-fold cross-validation initially decreases as variables are added to the model, but then subsequently increases. These results confirm that the cross-validation method is a good strategy to avoid overfitting.

Table 3. RMSE for MOR Prediction through MLR and SLR with 10-fold cross-validation and 5 Repetitions

Number of Variables	Variables in the model	RMSE Overfitted (MPa)	RMSE (MPa)	Δ RMSE ⁽¹⁾ (%)
15	All 15 variables shown in Table 2	(7.07)	10.29 ^a	-2.8
7	$-0.581 \cdot \mathbf{SLG} - 5.92 \cdot \mathbf{RG} - 20.0 \cdot \mathbf{WN}_m - 0.0249 \cdot \mathbf{knot}_T - 0.033 \cdot \mathbf{V}_U + 0.0923 \cdot \mathbf{f}_{LV} + 0.0047 \cdot \mathbf{MOE}_{EV} + 91.07$	(7.52)	8.77 ^b	-17.2
6	$-5.19 \cdot \mathbf{RG} - 21.9 \cdot \mathbf{WN}_m - 0.031 \cdot \mathbf{knot}_T - 0.0306 \cdot \mathbf{V}_U + 0.0852 \cdot \mathbf{f}_{LV} + 0.0049 \cdot \mathbf{MOE}_{EV} + 81.58$	(7.70)	8.56 ^b	-19.2
5	$-4.33 \cdot \mathbf{RG} - 0.0301 \cdot \mathbf{knot}_T - 0.0289 \cdot \mathbf{V}_U + 0.0892 \cdot \mathbf{f}_{LV} + 0.004 \cdot \mathbf{MOE}_{EV} + 73.28$	(7.92)	8.60 ^b	-18.8
4	$-0.0405 \cdot \mathbf{knot}_T - 0.0242 \cdot \mathbf{V}_U + 0.0725 \cdot \mathbf{f}_{LV} + 0.0045 \cdot \mathbf{MOE}_{EV} + 53.11$	(8.20)	8.69 ^b	-17.9
3	$-0.0414 \cdot \mathbf{knot}_T - 0.014 \cdot \mathbf{V}_U + 0.0063 \cdot \mathbf{MOE}_{EV} + 62.29$	(8.55)	9.04 ^{ab}	-14.6
2	$-0.0464 \cdot \mathbf{knot}_T + 0.003 \cdot \mathbf{MOE}_{EV} + 27.24$	(9.63)	9.96 ^{ab}	-5.9
1	$-0.0663 \cdot \mathbf{knot}_T + 58.83$	(10.88)	10.59 ^a	-

⁽¹⁾RMSE increase with respect to the SLR model; the RMSE obtained with the whole dataset (RMSE overfitted) is shown in brackets
 SLG: slope of grain; RG: rate of growth; WN: waness; V_U : velocity of the ultrasounds; f : frequency

Starting with a SLR with \mathbf{knot}_T , the inclusion of a new variable in the model resulted in important RMSE decrease in MOR prediction, reaching 17.9% in the model with 4 variables. However, the models with 5 or 6 variables added small decreases in the RMSE that were below 1% (with no statistically significant differences). With 7 or more variables the RMSE values began to increase indicating that these new variables added to the model were redundant. The model that obtained the lowest RMSE value was the MLR with 6 variables (8.56 MPa). However, the MLR model with 4 variables (\mathbf{knot}_T , V_U , f_{LV} , and MOE_{EV}) was the simplest model that presented an RMSE (8.69 MPa) significantly lower than the SLR.

Other authors have also found improvements in MLR compared to SLR models in MOR prediction in pine. Arriaga *et al.* (2012, 2014) found that MOR prediction on the basis of MOE_{LV} or MOE_{EV} improved when including in the model a knottiness variable. The results obtained by Grazide *et al.* (2015) showed that multivariate models based on knot variables and the MOE_S increased the R² value 27% compared to the SLR based only on the MOE_S. Similarly, while working with RG, MOE_{LV}, and knottiness, França *et al.* (2019) found that the univariate MOR prediction models obtained lower R² values than the MLR models that included two of these variables. The best result was obtained with those three variables for the MLR. Villasante *et al.* (2019) found that the lowest RMSE value in MOR prediction was obtained with a MLR that included 5 variables (velocity of resonant frequency, concentrated knot diameter ratio, attenuation of the ultrasonic wave, penetration depth, and withdrawal force of a screw).

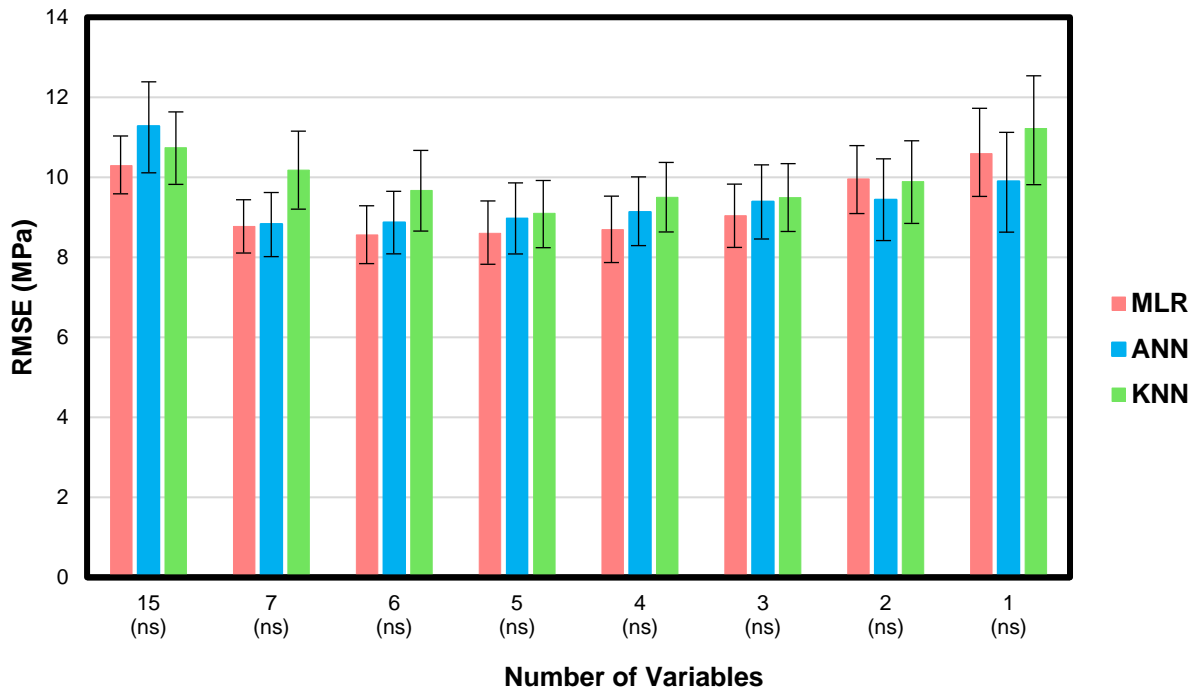


Fig. 4. RMSE for MOR prediction through MLR, ANN, and KNN algorithms with 10-fold cross-validation and 5 repetitions; mean values and 95% confidence intervals (ns: non-significant differences)

Nonlinear Multivariate Models

It was found that the ANN models with a lower number of variables gave a lower error when using fewer neurons. The same was found by Tetko *et al.* (1995). The best predictions were obtained with an ANN based on a single hidden layer of 3 to 5 neurons. No statistically significant differences were found in MOR prediction between the MLR- and ANN-based models for any of the eight groups of variables studied (Fig. 4). These results concur with those of Villasante *et al.* (2019) with *Pinus sylvestris* and Tanaka *et al.* (1996) with *Cryptomeria japonica*. However, other studies have reported better predictions with ANN than MLR, including Mansfield *et al.* (2007) with *Tsuga heterophylla* and Fathi *et al.* (2020) with poplar. These discrepancies may be attributable to the fact that both these

studies worked with small clear samples whereas the present study was conducted using structural size samples.

For the KNN-based models, the number of neighbor values employed were between 4 and 6. The inverse of the distance was used as distance weighting method. No improvement was observed in MOR prediction when using the KNN algorithm, with no statistically significant differences detected between the MLR and KNN results (Fig. 4).

CONCLUSIONS

1. The variable with the lowest modulus of rupture (MOR) prediction error was knottiness. Despite the complexity required to obtain the knot area ratio (KAR) value, this method was unable to significantly decrease the error obtained when using the simpler method described in Annex A of EN 1309-3 (2018).
2. The prediction error of the best univariate model, based on knottiness, was clearly improvable through the incorporation of other variables. The model with 4 variables (knot_T, V_U , f_{LV} , and MOE_{EV}) was the most suitable, as adding more variables only decreased the error slightly and non-significantly.
3. Performing tests in different vibration modes proved to be better than when using a single vibration mode test, as the best model combined a transverse and a longitudinal variable.
4. The usage of a cross-validation method also proved to be a suitable strategy to avoid overfitting- in multivariate models based on non-destructive testing (NDT) to determine bending strength in timber.
5. The nonlinear artificial neural network (ANN)- or k-nearest neighbors (KNN)-based models obtained similar results to those with multiple linear regression (MLR). No statistically significant differences were detected between these models in terms of MOR predictive capacity.

REFERENCES CITED

- Aira, J. R., Villanueva, J. L., and Lafuente, E. (2019). "Visual and machine grading of small diameter machined round *Pinus sylvestris* and *Pinus nigra* subsp. *salzmannii* wood from mature Spanish forests," *Materials and Structures* 52(2), article no. 32. DOI: 10.1617/s11527-019-1330-4
- Alexander, D. L. J., Tropsha, A., and Winkler, D. A. (2015). "Beware of R²: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models," *Journal of Chemical Information and Modeling* 55(7), 1316-1322. DOI: 10.1021/acs.jcim.5b00206
- Arriaga, F., Íñiguez-Gonzalez, G., Esteban, M., and Divos, F. (2012). "Vibration method for grading of large cross-section coniferous timber species," *Holzforschung* 66(3), 381-387. DOI: 10.1515/HF.2011.167
- Arriaga, F., Monton, J., Segues, E., and Íñiguez-Gonzalez, G. (2014). "Determination of the mechanical properties of radiata pine timber by means of longitudinal and

- transverse vibration methods,” *Holzforschung* 68(3), 299-305. DOI: 10.1515/hf-2013-0087
- ASTM E1876-15 (2015). “Standard test method for dynamic Young’s modulus, shear modulus, and Poisson’s ratio by impulse excitation of vibration,” ASTM International, West Conshohocken, PA, USA.
- Brunetti, M., Burato, P., Cremonini, C., Negro, F., Nocetti, M., and Zanuttini, R. (2016). “Visual and machine grading of larch (*Larix decidua* Mill.) structural timber from the Italian Alps,” *Materials and Structures* 49(7), 2681-2688. DOI: 10.1617/s11527-015-0676-5
- Carballo, J., Hermoso, E., and Fernández-Golfín, J. I. (2009). “Mechanical properties of structural maritime pine sawn timber from Galicia (*Pinus pinaster* Ait. ssp. *atlantica*),” *Forest Systems* 18(2), 152-158. DOI: 10.5424/fs/2009182-01058
- Conde García, M., Fernández-Golfín Seco, J. I., and Hermoso Prieto, E. (2007). “Mejora de la predicción de la resistencia y rigidez de la madera estructural con el método de ultrasonidos combinado con parámetros de clasificación visual [Improving the prediction of strength and rigidity of structural timber by combining ultrasound techniques with visual grading parameters]” *Materiales de Construcción* 57(288), 49-59. DOI: 10.3989/mc.2007.v57.i288.64
- Dahlen, J., Montes, C., Eberhardt, T. L., and Auty, D. (2018). “Probability models that relate nondestructive test methods to lumber design values of plantation loblolly pine,” *Forestry* 91(3), 295-306. DOI: 10.1093/forestry/cpy001
- EN 408:2011+A1 (2012). “Timber structures. Structural timber and glued laminated timber. Determination of some physical and mechanical properties,” European Committee for Standardization, Brussels, Belgium.
- EN 1309-3 (2018). “Round and sawn timber – Methods of measurements - Part 3: Features and biological degradations,” European Committee for Standardization, Brussels, Belgium.
- EN 13183-1 (2002). “Moisture content of a piece of sawn timber. Part 1: Determination by oven dry method,” European Committee for Standardization, Brussels, Belgium.
- EN 13183-2 (2002). “Moisture content of a piece of sawn timber – Part 2: Estimation by electrical resistance method,” European Committee for Standardization, Brussels, Belgium.
- Fathi, H., Nasir, V., and Kazemirad, S. (2020). “Prediction of the mechanical properties of wood using guided wave propagation and machine learning,” *Construction and Building Materials* 262, article ID 120848. DOI: 10.1016/j.conbuildmat.2020.120848
- Faydi, Y., Brancheriau, L., Pot, G., and Collet, R. (2017). “Prediction of oak wood mechanical properties based on the statistical exploitation of vibrational response,” *BioResources* 12(3), 5913-5927. DOI: 10.15376/biores.12.3.5913-5927
- França, F. J. N., Seale, R. D., Shmulsky, R., and França, T. S. F. A. (2019). “Modeling mechanical properties of 2 by 4 and 2 by 6 southern pine lumber using longitudinal vibration and visual characteristics,” *Forest Products Journal* 68(3), 286-294. DOI: 10.13073/FPJ-D-17-00069
- Frank, E., Hall, M. A., and Witten, I. H. (2016). *The WEKA workbench. Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques,”* Morgan Kaufmann, (https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf), Accessed 9 Dec 2021

- García Esteban, L., García Fernández, F., and de Palacios, P. (2009). “MOE prediction in *Abies pinsapo* Boiss. timber: Application of an artificial neural network using non-destructive testing,” *Computers & Structures* 87(21-22), 1360-1365. DOI: 10.1016/j.compstruc.2009.08.010
- Grazide, C., Cointe, A., Coureau, J. L., Morel, S., and Dumail, J. F. (2015). “Wood heterogeneities and failure load of timber structural elements: A statistical approach,” *Wood Science and Technology* 49(2), 421-440. DOI: 10.1007/s00226-015-0706-z
- Halabe, U. B., Bidigalu, G. M., GangaRao, H. V. S., and Ross, R. J. (1997). “Nondestructive evaluation of green wood using stress wave and transverse vibration techniques,” *Materials Evaluation* 55(9), 1013-1018.
- Hashim, U. R., Hashim, S. Z. M., and Muda, A. K. (2016). “Performance evaluation of multivariate texture descriptor for classification of timber defect,” *Optik* 127(15), 6071-6080. DOI: 10.1016/j.ijleo.2016.04.005
- Hassan, K. T. S., Horáček, P., and Tippner, J. (2013). “Evaluation of stiffness and strength of Scots pine wood using resonance frequency and ultrasonic techniques,” *BioResources* 8(2), 1634-1645. DOI: 10.15376/biores.8.2.1634-1645
- Hautamäki, S., Kilpeläinen, H., and Verkasalo, E. (2014). “Factors and models for the bending properties of sawn timber from Finland and north-western Russia. Part II: Scots pine,” *Baltic Forestry* 20(1), 142-156.
- Íñiguez González, G., Arriaga Martitegui, F., Esteban Herrero, M., and Argüelles Álvarez, R. (2007). “Los métodos de vibración como herramienta no destructiva para la estimación de las propiedades resistentes de la madera aserrada estructural [Vibration methods as non-destructive tool for structural properties assessment of sawn timber]” *Informes de la Construcción* 59(506), 97-105. DOI: 10.3989/ic.2007.v59.i506.512
- Lam, F., Barrett, J. D., and Nakajima, S. (2005). “Influence of knot area ratio on the bending strength of Canadian Douglas fir timber used in Japanese post and beam housing,” *Journal of Wood Science* 51(1), 18-25. DOI: 10.1007/s10086-003-0619-6
- Lever, J., Krzywinski, M., and Altman, N. (2016). “Points of significance: Model selection and overfitting,” *Nature Methods* 13(9), 703-704. DOI: 10.1038/nmeth.3968
- Mansfield, S. D., Iliadis, L., and Avramidis, S. (2007). “Neural network prediction of bending strength and stiffness in western hemlock (*Tsuga heterophylla* Raf.),” *Holzforschung* 61(6), 707-716. DOI: 10.1515/HF.2007.115
- Martins, C. E. J., Dias, A. M. P. G., Marques, A. F. S., and Dias, A. M. A. (2017). “Non-destructive methodologies for assessment of the mechanical properties of new utility poles,” *BioResources* 12(2), 2269-2283. DOI: 10.15376/biores.12.2.2269-2283
- Pommier, R., Breysse, D., and Dumail, J. F. (2013). “Non-destructive grading of green Maritime pine using the vibration method,” *European Journal of Wood and Wood Products* 71(5), 663-673. DOI: 10.1007/s00107-013-0727-y
- Tanaka, T., Nagao, H., and Kato, H. (1996). “A preliminary investigation on evaluation of strength of soft wood timbers by neural network,” in: *10th International Symposium on Nondestructive Testing of Wood*, Lausanne, Switzerland, pp. 323-329.
- Tetko, I. V., Livingstone, D. J., and Luik, A. I. (1995). “Neural network studies. 1. Comparison of overfitting and overtraining,” *Journal of Chemical Information and Computer Sciences* 35(5), 826-833. DOI: 10.1021/ci00027a006
- Vega, A., Dieste, A., Guaita, M., Majada, J., and Baño, V. (2012). “Modelling of the mechanical properties of *Castanea sativa* Mill. structural timber by a combination of

- non-destructive variables and visual grading parameters,” *European Journal of Wood and Wood Products* 70(6), 839-844. DOI: 10.1007/s00107-012-0626-7
- Villasante, A., Íñiguez -Gonzalez, G., and Puigdomenech, L. (2019). “Comparison of various multivariate models to estimate structural properties by means of non-destructive techniques (NDTs) in *Pinus sylvestris* L. timber,” *Holzforschung* 73(4), 331-338. DOI: 10.1515/hf-2018-0103
- Walker, J. (1993). *Primary Wood Processing: Principles and Practice*, Chapman & Hall, London, United Kingdom.
- Weaver, W., Timoshenko, S., and Young, D. H. (1990). *Vibration Problems in Engineering*, Wiley, New York, NY, USA, pp.367-426.
- Yang, B. Z., Seale, R. D., Shmulsky, R., Dahlen, J., and Wang, X. (2017). “Comparison of nondestructive testing methods for evaluating no. 2 southern pine lumber: Part B, modulus of rupture,” *Wood and Fiber Science* 49(2), 134-145.
- Zhou, H., and Smith, I. (1991). “Factors influencing bending properties of white spruce lumber,” *Wood and Fiber Science* 23(4), 483-500.

Article submitted: July 9, 2021; Peer review completed: November 30, 2021; Revised version received and accepted: December 15, 2021; Published: December 17, 2021.
DOI: 10.15376/biores.17.1.1106-1119