

RESEARCH ARTICLE

Open Access



# Validation of models predicting modulus of elasticity in Douglas-fir trees, boles, and logs

Christine L. Todoroki<sup>1\*</sup> and Eini C. Lowell<sup>2</sup>

## Abstract

**Background:** Acoustic tools have simplified estimation of wood modulus of elasticity (MOE). Strong relationships between acoustic velocity and MOE of logs have encouraged use of acoustics at earlier points in the value chain, culminating in the development of acoustic harvesting systems. With accurate estimates of MOE of individual trees, improvements in efficiency along the value chain and increased value recovery will result.

Our aim was to quantify the accuracy of MOE estimates at three distinct points: pre-harvest (standing trees), during harvest (merchantable boles), and post-harvest (5-m logs). We hypothesised that: (1) MOE estimated from acoustic velocity and wood density would provide greatest accuracy; and (2) bole estimates with a resonance tool would be more accurate than tree estimates with a time-of-flight tool.

**Methods:** A sample of 168 Douglas-fir (*Pseudotsuga menziesii* [Mirb. Franco]) trees, representing the variability in acoustic velocity of 700 standing 36–51-year-old trees, was harvested from three sites. Prior to harvest, time-of-flight and breast-height diameter were recorded. After felling, resonance velocities of boles and subsequent 5-m logs were recorded. Discs, cut from log ends, were immersed, and green wood density determined. Half the logs were processed into boards, the other half into veneer sheets, and all products (in excess of 6000) non-destructively tested for MOE. MOE of parent trees, boles, and logs was then calculated from the mean MOE of derived products.

Predictive mixed-effects models of tree, bole, and log MOE were developed using data from 139 trees. Fixed effects comprised combinations of velocity squared, wood density, acoustic MOE (derived from the wave equation), diameter, height, taper, and age. Random effects comprised site, plot, and, at the log level, tree. The models were validated using data from the remaining trees and compared using multiple performance metrics.

**Results:** For estimating tree MOE, a model with velocity squared, wood density, and taper as predictors is recommended. For estimating MOE of boles and logs, models with velocity squared and wood density are recommended. The models have an accuracy, as determined by RMSE, of about  $\pm 2$  GPa.

**Conclusions:** For accurate MOE estimation, velocity alone is insufficient. Knowledge of wood density is necessary for improved accuracy.

**Keywords:** Accuracy, Density, Modulus of elasticity, Acoustic velocity, Douglas-fir, Mixed-effects models

\* Correspondence: Christine.Todoroki@scionresearch.com

<sup>1</sup>Scion, Private Bag 3020, Rotorua 3046, New Zealand

Full list of author information is available at the end of the article

## Background

Acoustic tools have greatly simplified the process of assessing the dynamic modulus of elasticity (MOE) of trees and wood products, and extensive reviews of their use are provided by Legg and Bradley (2016) and Wang (2013). In short, a sound wave is transmitted, typically by a mechanical impact with a hammer, between one probe and the cut end (resonance method, Wang 2013) or between two probes (time-of-flight method, Bucur 2006; Senalik et al. 2014). Estimates of MOE are then derived from predictive models that include acoustic velocity,  $V$ , or the square of acoustic velocity,  $V^2$ , as a predictor variable.

Acoustic tools have been used in log production settings for more than a decade (Dickson et al. 2004), but the concept of moving the process closer to the beginning of the value chain and implementing acoustic tools on harvester heads is more recent (Walsh et al. 2014). Amishev and Murphy (2008a) examined the feasibility of using acoustic technology on harvesting equipment for identification of veneer quality Douglas-fir logs (*Pseudotsuga menziesii* [Mirb.] Franco). They also reported on in-wood non-destructive measurements to determine applicability for in-forest sorting of veneer quality logs in second growth Douglas-fir and found it likely to improve recovery of higher quality logs (Amishev and Murphy 2008b). The importance of early application is succinctly stated by Divos (2010); “The earlier well-informed decisions are made within the forest-to-products chain, the greater the potential value addition. Thus if it can be shown that strong relationships exist between tree and product, then those relationships can then be used to generate added value.”

Acoustic tools make use of the often-cited relationship between MOE and  $V^2$ , based on the fundamental wave equation (Eq. 1). The equation states that the dynamic MOE of a homogeneous and isotropic material is equal to the square of the acoustic velocity ( $V^2$ ) multiplied by its density ( $\rho$ ).

$$\text{MOE} = \rho \cdot V^2 \quad (1)$$

Though wood is heterogeneous and orthotropic,  $V^2$  (or  $V$ ), has been found to be a highly significant predictor variable in linear models. Yet the strength of the relationship between MOE and  $V^2$  (or  $V$ ) varies considerably (Achim et al. 2010).

In forestry literature, the strength of the relationship between MOE and  $V^2$  (or  $V$ ) has often been cited in terms of the correlation coefficient,  $r$  (Eq. 2), or the coefficient of determination, denoted  $R^2$ .  $R^2$  (and any variation of  $R^2$ ) explains the variance of a linear model. Amongst the more common variations of  $R^2$  is the adjusted  $R^2$ ,  $R^2_{\text{adj}}$ , which adjusts for the number of variables in the regression equation. For linear mixed-effects models with a hierarchical

(i.e. nested) structure, Nakagawa and Schielzeth (2013) derived two  $R^2$  variates: marginal  $R^2$  and conditional  $R^2$ . The former describes the proportion of variance explained by fixed factors alone, while the latter describes the proportion of variance explained by both fixed and random factors. The method, developed for generalised linear models and applicable to random intercepts models, has since been extended, to accommodate random slope models, by Johnson (2014).

$$r = \frac{n \sum x \cdot y - (\sum x) \cdot (\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \cdot \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (2)$$

where  $x$  is the explanatory variable,  $y$  the response variable, and  $n$  the sample size.

In terms of  $R^2$ , the relationship between MOE and  $V^2$  (or  $V$ ) has been demonstrated to nearly span the range of values; from 0, indicating no relationship, to 0.98, just short of the perfect relationship.

Low values of  $R^2$  were found for relationships between acoustic velocity in standing radiata pine (*Pinus radiata* D.Don) trees, and logs, with machine stress-grades of boards derived from those trees, and logs (Matheson et al. 2002). Relationships with trees propagated from an orchard seedlot recorded an  $R^2$  close to 0 ( $r = 0.01$ ), while the control trees recorded an  $R^2$  equal to 0.11 ( $r = 0.33$ ). The relationship with logs had an  $R^2$  of 0.25 ( $r = 0.5$ ).

Moderate values of the coefficient of determination were obtained by Lowell et al. (2014) for relationships between acoustic velocities in standing Douglas-fir trees with mean veneer MOE derived from those trees ( $R^2 = 0.21$ ). Also with Douglas-fir, Briggs et al. (2008) determined an  $R^2$  of 0.62 (after removal of outliers) for the relationship between mean timber MOE of the lowermost log in a tree and acoustic velocity of that log. A coefficient of determination of 0.53 was obtained by Auty and Achim (2008) for the relationship between standing tree acoustic velocity of Scots pine (*Pinus sylvestris* L.) and MOE of small clear samples.

Achim et al. (2010) found that relationships between static MOE of sawn boards and acoustic velocity in logs varied by species; with  $R^2$  values ranging from a low 0.07 (for black spruce (*Picea mariana* [Mill.] Britton, Sterns & Poggenburg)) to 0.51 (for jack pine (*Pinus banksiana* Lamb.) and trembling aspen (*Populus tremuloides* Michx.)). They also determined, for white spruce (*Picea glauca* [Moench] Voss), that static MOE and dynamic MOE were highly correlated ( $R^2 = 0.69$ ).

With logs grouped into percentile batches, Carter et al. (2005) demonstrated a nearly perfect relationship ( $R^2 = 0.98$ ) between dry timber MOE and log-batch velocity

for a species referred to as Southern pine and radiata pine. However, the result was overly optimistic, because an average trend through averages (linear regression through batches) has a high  $R^2$  due to the removal of much of the original variation.

Predictor variables other than acoustic velocity have been used, either in simple linear models or in multivariable models, to estimate MOE of trees and logs. Relationships with diameter alone have tended to be weak. Farrell et al. (2008) determined an  $R^2$  of 0.09 for *Eucalyptus nitens* (H.Deane & Maiden) logs, and Wang et al. (2013) an  $R^2$  of 0.12 for Douglas-fir logs. For Douglas-fir trees, Lowell et al. (2014) determined an  $R^2$  of 0.21 for the relationship with breast-height diameter ( $D$ ), and a slightly stronger relationship, with  $R^2 = 0.28$ , for tree taper. Relationships with the combination of  $V$  and  $D$  have been reported as stronger than those with  $V$  alone (e.g.  $R^2 = 0.36$ , Farrell et al. 2008, and  $R^2_{\text{adj}} = 0.50$ , Wang et al. 2013). Paradis et al. (2013) determined an  $R^2$  of 0.41 for the relation with  $V^2$ ,  $D$ , and an interaction term between the two variables for black spruce.

Relationships with wood density as a predictor have tended to be relatively strong. This comes as no surprise as wood density is one of the key determinants of MOE (Zobel and van Buijtenen 1989), explaining up to 60–70 % of the variation in MOE in clear wood (Rozenberg et al. 1999; Evans and Ilic 2001). Worldwide correlations with data obtained from 161 tree species have also demonstrated strong positive relationships between MOE and density (Niklas and Spatz 2010). The worldwide relationships for conifers had an  $R^2$  of about 0.57. Merlo et al. (2014) observed that, in comparison to models without wood density as a predictor, models which included density had  $R^2$  values that were about 4–9 % higher. More modest improvements were obtained by Dickson et al. (2004) using Pilodyn depth as a surrogate for wood density. Liu et al. (2007) derived models for estimating MOE of 90–100-year-old black spruce trees. One model based on wood density, diameter, and crown length had an  $R^2$  of 0.65 and a mean absolute percentage error (MAPE) of 1.2 GPa; another model based on diameter, crown length, stem taper, and stand density had an  $R^2$  of 0.55 and MAPE of 1.4 GPa. Relatively good accuracy metrics were obtained when the models were validated with external data ( $R^2 = 0.55$ , MAPE = 1.1 GPa for the first model,  $R^2 = 0.43$ , MAPE = 1.1 GPa for the second).

Irrespective of the value of  $R^2$ , models for estimating MOE (or any other response) must be evaluated with external data before they are used in practice (Collins et al. 2014; Altman and Royston 2000). External validation is a necessary step in determining the predictive capability of the model and transportability to other settings. The use of (or lack of) external validation has recently come under scrutiny. In a review of articles, Collins et al. (2014) reported that the difference between observed and predicted external data was one of the key

performance measures for models yet was often omitted from the publication and concluded that “It may, therefore, not be surprising that an overwhelming majority of developed prediction models are not used in practice”.

Unfortunately, none of the aforementioned models for predicting MOE, with the exception of that by Liu et al. (2007), were supported by external validation datasets (though an internal validation technique, cross-validation, was used by Merlo et al. (2014)). Therefore, even though the performance of the aforementioned models may appear to be satisfactory, the results are an optimistic estimate of model performance. The model parameters are optimised, i.e. fine-tuned, to the development dataset, hence tests with new data may demonstrate disappointing results.

External validation provides a stronger test of model performance than internal validation because it addresses model transportability through the use of an independent dataset (Steyerberg and Vergouwe 2014), whereas internal cross-validation can cause bias in error estimation (Varma and Simon 2006).

With external validation, predictions are calculated using the model with new data from an external dataset (i.e. different to that used in model development). The accuracy (or error) between actual and estimated (external) data is then measured using one or more performance metrics. However, many performance metrics exist (Makridakis and Hibon 1995), with no single method designated as the best (Winkler and Murphy 1992). Furthermore, the use of a single error measure may lead to incorrect interpretation.

Standard performance metrics include  $R^2$ , root mean square error (RMSE), and mean absolute percentage error (MAPE). The use of  $R^2$  has been criticised (Willmott 1981), mainly due to its sensitivity to outliers. The RMSE (Eq. 3), while also influenced by outliers, is a convenient metric as it assumes the same units as the data, is representative of the size of a typical error, and is a lower bound on the standard deviation of a prediction error. The MAPE (Eq. 4), on the other hand, while also influenced by outliers, expresses errors as a percentage of the actual data, so is easy to interpret. For binary data, performance measures can be derived from a confusion matrix (Fawcett 2006). The  $2 \times 2$  matrix contains four cells: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). If the purpose of segregation is to identify high MOE potential, then the true positive rate, the proportion of positive cases that are correctly identified (Eq. 5), is of primary importance.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

$$\text{MAPE} = \frac{100 \times \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \quad (4)$$

$$\text{TP rate} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

To achieve our aim of quantifying the accuracy of MOE estimates, we first developed predictive linear mixed-effects models of tree, bole, and log MOE. The models were then judged in terms of multiple accuracy metrics applied across two datasets: (1) the calibration dataset from which the models were derived, and (2) an external validation dataset not used in model development. Accuracy metrics for the latter dataset were of particular interest.

Based on findings in the literature, as noted above, we hypothesised that MOE estimated from the combination of acoustic velocity and wood density would provide the greatest accuracy and that bole estimates with a resonance tool would be more accurate than tree estimates with a time-of-flight tool.

Our focus was on MOE predictions of individual Douglas-fir wood members (standing trees, boles, and logs) from attributes of those individuals (e.g. acoustic velocity, wood density, diameter). To aid transportability of the predictive models to other settings, a wide range in attributes was required. Thus trees were harvested from three sites, each comprising five plots which had been managed with differing silvicultural treatments. Though necessary to obtain variability of attributes, and necessary as structures in the mixed-effects models, our focus was not on modelling MOE as a function of site, nor of silviculture.

## Methods

### Sites

Three long-term Douglas-fir research trials in the Pacific Northwest of America, two in Washington (site 803, at 47° 16' N, 123° 27' W, with an elevation of 175 m; and site 805, at 48° 19' N, 122° 9' W, with an elevation of 168 m) and one in Oregon (site 808, at 44° 57' N, 123° 40' W, with an elevation of 762 m) formed the basis for this study. Trees at sites 803 and 808 were about 50 years old (51 and 46 years for the respective sites) while those at site 805 were 36 years old. The sites were similar in terms of mean annual temperature (9–11 °C), but differed in

mean annual precipitation, with 264 mm per year at site 803, about 110 mm at site 805, and 336 mm per year at site 808.

### Silviculture

Each site comprised five plots with differing thinning treatments; a control (A) which was not thinned, and four other regimes of varying thinning intensities. The timing of thinning was specified by relative density (Curtis 1982), which in turn was dependent on stand growth rates. In some cases, stand growth was insufficient and the prescribed thinning treatment(s) did not occur. For example, on sites 803 and 805, the delayed thinning treatment (treatment C), required a relative density (RD) of 65 for thinning to occur. As this RD was not attained, the thinning operation did not occur. On site 808 under treatment C, though the stand was thinned, it was damaged by a storm, thus eliminating this stand from the study. For treatments D and E with prescribed multiple thinnings, only growth at site 805 under treatment D was sufficient for more than one thinning to occur. The thinning prescriptions (thinning timings and prescribed number of thinnings) and the actual number of thinnings that occurred are listed in Table 1. Stocking levels at harvest for each of the plots is also provided.

### Sample tree selection

At each plot on the three sites, all Douglas-fir trees within an 18-m radius of the plots geometric centre were acoustically assessed. The 50 or so acoustic readings of each plot were ordered and 12 trees selected using a stratified approach: two trees were randomly selected from the lowest 10 %, two from the highest 10 %, and eight from the middle 80 %. If a sample tree had defects (such as forks, crooks, pistol butt, rot) that would compromise the objectives of the study, a replacement tree was randomly selected from the appropriate acoustic class. Examination of the effect of such defects on acoustic velocity relationships was beyond the scope of the study.

Half of the trees within each acoustic group on each plot were allocated for peeling into veneer, the other half for sawing into boards. There were no bucking decisions concerning which log in a subject tree went to which

**Table 1** Thinning prescriptions, specified by relative density (RD) following Curtis (1982), and actual number of thinnings with final stocking at the three study sites (803, 805, 808)

Treatment	Prescribed treatment		Actual thinnings			Stocking at harvest (sph)		
	No.	Description	803	805	808	803	805	808
A	0	No thinning (control)	0	0	0	514	692	494
B	1	Minimal thinning: RD55 → RD30	1	1	1	287	425	292
C	1	Delayed thinning: RD65 → RD35	0	0	1	558	351	<sup>a</sup>
D	2+	Repeated thinning: RD55 → RD30, all subsequent RD50 → RD30	1	2	1	326	217	242
E	3+	Repeated thinning: RD55 → RD35, RD55 → RD40, all subsequent RD60 → RD40	1	1	1	410	400	351

<sup>a</sup>Damaged by storm and eliminated from study



product as all logs from a tree went to the same product, i.e. all logs from a single tree were processed to the same product type.

#### Acoustic measurements of trees, boles, and logs

Standing tree acoustic velocity and breast-height diameter were measured at harvest. A total of nine acoustic velocity hits were recorded with the Fakopp TreeSonic tool (Fakopp Bt. Agfalva, Hungary) for each tree. Three observations were obtained (as recommended by the manufacturers (Fakopp 2006)) at each of three locations, approximately 120° apart around the base of the tree. Observations were averaged, providing a mean acoustic velocity reading for each tree. Trees were felled, delimited, and topped to give a merchantable bole with a top diameter of 15 cm inside the bark for sawmill trees, and 13 cm inside the bark for peeler trees. Acoustic velocity of the merchantable bole was obtained with the Director HM-200 (Fibre-gen, New Zealand), a resonance-type tool. The boles were crosscut at 10-m intervals then crosscut again at the mid-section and the resulting 5-m-long logs acoustically assessed with the Director HM-200 tool.

#### Wood density

Each log included mill trim plus an additional length to allow for the cutting of discs. The discs, approximately 5 cm thick, were removed from the butt end of the bole, and from the top of each log, and processed for green weight and green volume, to derive a measure of wood density using the Archimedean immersion method. Mean wood density of each bole (and each tree) was calculated as the diameter-weighted mean of all discs from each tree, and mean wood density of each log was calculated as the diameter-weighted mean of the two log-end discs.

#### Board and veneer production

Each log was uniquely labelled to enable full tracking of products back to the parent bole, tree, plot, and site. Sawmill logs were sawn into boards (nominally 5 cm thick), kiln-dried (and after transportation to the laboratory where E-rating occurred, had a moisture content (MC) of about 24 %), planed, measured (for thickness, width, and length), graded under the supervision of a grading inspector, and E-rated using a Metriguard E-computer (Metriguard Inc., Pullman, WA, USA) following standard test methods (ASTM 2003). More detail of the procedure for determining board MOE is provided by Wang et al. (2013). Peeler logs were steamed, peeled into veneer sheets (with dry target sizes of 2.59 m in length, 1.32 m in width, and 3.8 mm in thickness), dried (using schedules determined by the cooperating mill in accordance with their normal practices), weighed, and

measured. Veneer sheet density was derived as the quotient of mass and volume, and all sheets were non-destructively tested for MOE using a Metriguard 2600™ veneer tester (Metriguard Inc., Pullman, WA, USA). More detail of the procedure for determining veneer MOE is provided by Lowell et al. (2014). Following data consolidation, a total of 1639 boards and 4645 veneer sheets were available to this study.

#### Mean MOE

Mean MOE of logs, boles, and trees was calculated as the volume-weighted mean MOE of constituent products. Therefore, mean bole MOE was identical to mean tree MOE, thus allowing comparison of the accuracy of time-of-flight and resonance methods.

#### Data consolidation

Numerous checks were made throughout the data collection, processing, and data entry processes. Further verification exercises were performed to investigate outliers and to examine consequences of missing data. Much of this was performed graphically, noting the location of the missing product within the parent log and bole. After cleaning the data, a total of 162 trees was available for this study; of which 139 trees (and their products) were used in model development and the remaining 23 trees (and associated products) reserved for model validation. Acoustic velocity measurements of two boles were missing, hence there were two fewer boles than trees. All validation trees and products were from treatment C. This enabled the model development phase to be as close to a balanced design as possible, while also meeting requirements for external validation.

#### Modelling approach

Models estimating MOE of trees, boles, and logs were developed using the calibration dataset. Fixed effects included velocity squared, diameter, height, taper, wood density, age, and acoustic MOE (equivalent to  $V^2 \cdot \rho$ ; Eq. 1). Interaction terms between  $V^2$  and  $D$  (following Paradis et al. 2013) were also evaluated. Random effects were site and plot (plot nested within site). For log-level models, random effects included trees (nested within plots, and plots within sites) and a first-order autocorrelation structure (Pinheiro and Bates 2000), with ring number as a covariate, included in the models to address the correlated nature of the multiple logs cut from each tree. The fixed and random variables for the tree, bole, and log level models are summarised in Table 2 while summary statistics (mean, minimum, and maximum) for the tree, bole, and log calibration and validation datasets are provided in Table 3.

Diameter, denoted  $D$ , at the tree level corresponded to breast-height diameter (1.4 m), while at the bole and log

**Table 2** Summary of fixed and random effects used in models predicting mean MOE

Model level	Fixed effects	Random effects <sup>a</sup>
Tree	$V^2, D, H, T, \rho, A, E$	Site/plot
Bole	$V^2, D, H, T, \rho, A, E$	Site/plot
Log	$V^2, D, H, T, \rho, A, E$	Site/plot/tree

$V$  = acoustic velocity (time-of-flight for trees, resonance for boles and logs,  $\text{km s}^{-1}$ );  $D$  = diameter (breast-height for trees, at base for bole and logs, cm);  $\rho$  = mean weighted green wood density ( $\text{kg m}^{-3}$ );  $A$  = age (harvest age for trees, ring counts for boles and logs, years);  $H$  = height (total height of trees, or length of boles (m), or relative height of log in tree);  $T$  = taper (i.e.  $D/H$ ,  $\text{cm m}^{-1}$ );  $E$  = acoustic MOE (i.e.  $V^2 \cdot \rho$ , following Eq. 1)

<sup>a</sup>w/x/z = z nested within x within w

levels, diameter corresponded to the diameter at the base. Velocity, denoted  $V$ , at the tree level corresponded to that derived from the “time-of-flight” measured on the standing tree, in contrast to resonance velocity at the other two levels. Height, denoted  $H$ , corresponded to total tree height at the tree level, to bole length at the bole level, and to relative height of the log within the tree (measured from the base of the log) at the log level. The same variable notations were used at each level, being easily discernible from the context.

Linear models were formulated using the Laird and Ware (1982) model (Eq. 6)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \tag{6}$$

where  $\mathbf{y}$  is the response vector,  $\mathbf{X}$  and  $\mathbf{Z}$  are matrices of explanatory variables corresponding to fixed and random effects, respectively,  $\boldsymbol{\beta}$  and  $\mathbf{u}$  are the corresponding vectors of parameters for the respective fixed and random effects, and  $\boldsymbol{\varepsilon}$  is a vector of random errors. The models were developed using the linear and nonlinear mixed-effects model package, “nlme”, (Pinheiro et al. 2013) within the *R* environment (R Core Team 2013). For multivariable models, which have the possibility of dependency amongst variables, variance inflation factors (VIF, Davis et al. 1986) were computed. Following O’Brien (2007), arbitrary thresholds (e.g. 4 or 10) used to eliminate a variable were not applied, and the variance of the regression coefficients were viewed in context. The significance of explanatory variables was evaluated using a 0.05 level of significance.

Accuracy of models were compared using multiple accuracy metrics ( $R^2_{\text{marg}}$ ,  $R^2_{\text{cond}}$ , AIC, RMSE, MAPE, and TP rate), to avoid incorrect interpretation which could result from a single metric. Calculation of  $R^2_{\text{marg}}$  (for fixed effects only) and  $R^2_{\text{cond}}$  (for both fixed and random effects) for nested models followed Johnson (2014). Because  $R^2$  tends to increase with model complexity, the Akaike information criterion (AIC, Akaike 1974), which penalises complex models to determine a balance between model fit and model simplicity, was also used for model selection. A threshold value of 11.0 GPa (based on the tree breeding value recommended by Jayawickrama et al. 2009) was

**Table 3** Summary statistics of the calibration and validation datasets (mean, min-max)

Variable <sup>a</sup>	Trees		Boles		Logs	
	Calibration $N = 139$	Validation $N = 23$	Calibration $N = 137$	Validation $N = 23$	Calibration $N = 481$	Validation $N = 78$
$V$ ( $\text{km s}^{-1}$ )	3.86 2.53–4.44	3.87 2.92–4.50	3.59 3.08–4.12	3.74 3.37–4.12	3.53 2.77–4.18	3.68 3.24–4.02
$D$ (cm)	39.0 20.8–60.7	33.3 21.6–46.0	41.5 21.2–62.9	37.1 25.5–48.8	33.8 17.3–62.9	30.1 18.5–48.8
$H$ (m) <sup>b</sup>	32.1 21.9–43.6	33.3 27.1–39.7	25.0 16.5–35.4	24.9 14.3–32.6	0.23 0.00–0.73	0.23 0.01–0.56
$T$ ( $\text{cm m}^{-1}$ )	1.22 0.74–2.07	1.00 0.73–1.17	1.68 1.01–2.91	1.49 1.20–1.78	1.12 0.00–3.92	0.92 0.21–2.18
$\rho$ ( $\text{kg m}^{-3}$ )	805 674–902	809 708–878	804 674–902	809 708–878	805 634–986	801 666–952
$A$ (years)	44 36–51	44 36–51	41 33–48	40 34–47	32 13–48	33 17–47
$E$ (GPa)	12.1 5.3–16.4	12.2 6.9–16.4	10.4 7.3–13.7	11.4 8.9–13.4	10.0 5.8–13.6	10.8 8.4–12.7
MOE (GPa)	10.4 6.1–15.1	10.9 8.5–12.8	10.4 6.1–15.1	10.9 8.5–12.8	10.3 4.6–16.7	10.8 6.1–14.5

<sup>a</sup>All variables are as defined in Table 2

<sup>b</sup>For logs,  $H$  represents relative height and is unit-less

assumed in the classification procedure defining TP rate, and standard scientific rounding practices applied (e.g. an estimate of 10.95 GPa, rounds to 11.0 GPa, therefore a true positive score would result).

The values of  $R^2_{\text{marg}}$ ,  $R^2_{\text{cond}}$ , and AIC were calculated for the calibration models. Diagnostics of the calibration models also included plots of residuals against fitted values and plots of observed values versus fitted values. The RMSE, MAPE, and TP rate were calculated for both calibration and validation datasets, for comparative purposes; however, those of the validation datasets were of greatest interest because they provide insight to the practicality of the models and parameters beyond the dataset upon which they were derived. All predictions and subsequent calculations of RMSE, MAPE, and TP rate, were made at the population level, because, in practice, contributions due to random effects are unknown.

### Results

The best models, in terms of the accuracy metrics, for the calibration dataset were not necessarily the best amongst the validation dataset, particularly for models estimating MOE of trees. In general, there was better agreement in performance metrics of the validation and calibration datasets for log-level models than for tree-level models.

The combination of  $V^2$  and  $\rho$ , in the linear mixed-effects models, generally improved estimates of MOE (in

comparison to estimates based on  $V^2$  alone), but did not always result in the best model for estimating MOE. Models in acoustic MOE ( $E = V^2, \rho$ ) generally had similar performance metrics to models in  $V^2 + \rho$ . Interaction terms were not significant.

### Models estimating tree MOE

Performance metrics of all valid tree models (i.e. models for which all variables were significant,  $p < 0.05$ ) are shown in Table 4. In actuality, all model variables listed in Table 4 were highly significant ( $p < 0.001$ ) and had low dependencies (VIF less than 2.2).

The strength of the relationships for estimating tree MOE, as assessed by  $R^2_{\text{marg}}$ , ranged from 0.18 (for two single-variable models;  $D$  and  $\rho$ ) to 0.51 (for two multivariable models;  $E + T$  and  $V^2 + \rho + T$ ). RMSE ranged from 1.1 to 1.6 GPa, and MAPE from 7 to 13 %. The best (lowest) AIC was associated with the model in  $V^2, \rho$ , and  $D$ .

The relationship between  $V^2$  and MOE was not strong and considerable variation remained unexplained. The term  $V^2$  explained only 36 % of the variance, while random factors explained a further 9 %. With  $V^2$  as the sole predictor, RMSE and MAPE of the validation dataset (1.5 GPa and 11 %, respectively) were the worst of all models, and segregation accuracy, as defined by TP rate, was just 46 %.

Diameter, as a predictor variable, recorded good accuracy metrics amongst the validation dataset. Prediction

**Table 4** Performance metrics of linear models estimating MOE of standing trees. Italicised values indicate the recommended model; bold values indicate the best value(s) in each column

Fixed effects <sup>a</sup>	Calibration dataset						Validation dataset			Comments
	$R^2_{\text{marg}}$	$R^2_{\text{cond}}$	AIC	RMSE (GPa)	MAPE (%)	TP rate (%)	RMSE (GPa)	MAPE (%)	TP rate (%)	
$V^2$	0.36	0.45	465	1.3	10	58	1.5	11	46	
$D$	0.18	0.49	493	1.6	12	31	1.2	<b>7</b>	55	
$\rho$	0.18	0.44	487	1.5	11	36	1.2	8	27	
$T$	0.32	0.62	483	1.5	13	53	1.2	9	82	Best 1-variable
$E$	0.43	0.54	443	1.2	9	67	1.4	10	73	
$V^2 + \rho$	0.44	0.56	438	1.2	9	67	1.2	9	64	Similar to $E$
$V^2 + T$	0.45	0.58	449	1.2	10	60	1.3	10	82	
$V^2 + D$	0.37	0.50	457	1.3	10	56	1.3	10	64	
$D + H$	0.33	<b>0.64</b>	480	1.5	12	56	1.3	9	82	Similar to $T$
$D + \rho$	0.31	0.61	449	1.4	10	42	1.2	9	55	
$H + \rho$	0.21	0.59	479	1.6	13	29	1.4	10	36	
$E + D$	0.45	0.57	434	1.2	9	64	1.3	9	64	
$E + T$	<b>0.51</b>	0.60	429	<b>1.1</b>	9	69	1.3	10	<b>91</b>	
$E + D + H$	<b>0.51</b>	0.61	428	<b>1.1</b>	9	69	1.3	10	<b>91</b>	
$V^2 + D + H$	0.46	0.60	447	1.2	10	60	1.4	11	82	
$V^2 + \rho + D$	0.46	0.61	<b>422</b>	<b>1.1</b>	<b>8</b>	64	<b>1.1</b>	8	73	Best AIC
$V^2 + \rho + T$	<b>0.51</b>	<i>0.61</i>	423	<b>1.1</b>	<b>8</b>	<b>71</b>	<b>1.1</b>	8	<b>91</b>	Best validation

<sup>a</sup>All variables are as defined in Table 2

accuracy with  $D$  was better than prediction accuracy with  $V^2$ . However, performance metrics amongst the calibration dataset for  $D$ , particularly the explained variance, was poor ( $R_{\text{marg}}^2$  0.18).

The model in  $E$  (where  $E = V^2 \cdot \rho$ ) recorded the best apparent performance amongst single-variable models and would have been selected as the best single-variable model had results been based solely on the calibration dataset with no subsequent validation. However, better validation performance metrics were recorded for the model with taper as the sole predictor.

The model in taper recorded the best TP rate (82 %) amongst the single-variable models. The rate was considerably higher than TP rates of other single-variable models (validation dataset). However, model performance and accuracy metrics of the calibration dataset were less satisfactory. Similar accuracy metrics were found for the model with  $D + H$ . The TP rate was greatest (91 %) for two models:  $E + T$  and  $V^2 + \rho + T$ . However, the latter model recorded better performance metrics overall so was selected in favour of the former for further analysis.

Plots of actual versus predicted tree MOE of four selected models ( $V^2$ ,  $T$ ,  $V^2 + \rho + D$ ,  $V^2 + \rho + T$ ) are provided in Fig. 1. The improved performance of two models ( $V^2 + \rho + D$ ,  $V^2 + \rho + T$ ) is clearly visible, as is the reason for the high TP rate (refer Eq. 5) of the taper model with the validation dataset.

**Recommended tree model**

Based on the performance statistics of Table 4, and scatter plots of Fig. 1, the model with  $V^2 + \rho + T$  as predictor variables is recommended for estimating tree MOE. Fixed factors of the recommended model explained 51 % of the variance and random factors an additional 10 %. Parameter estimates of the recommended model, together with standard errors, are given in Eq. 7.

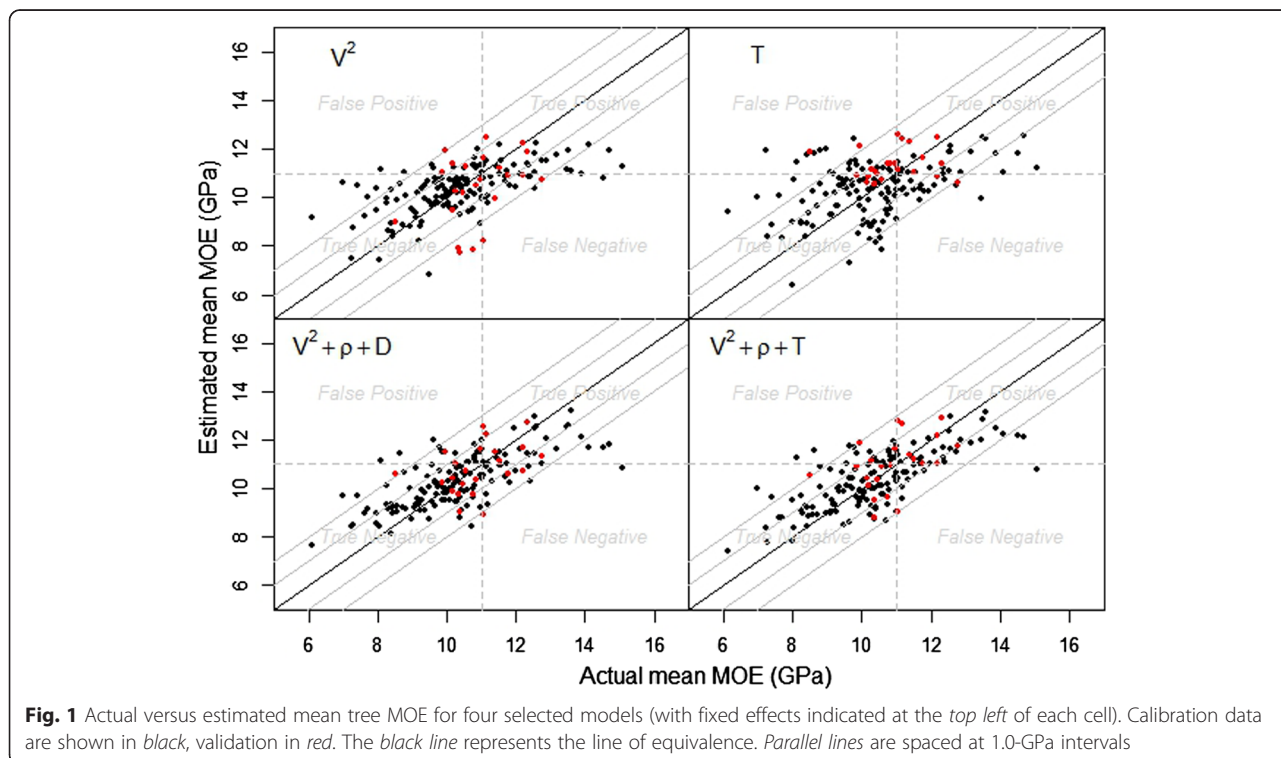
$$\text{MOE} = -1.363 + 0.2783V^2 + 0.01294\rho - 2.348T \quad (7)$$

Std. error (2.167) (0.0453) (0.00233) (0.554)

Estimates of the random effects associated with the recommended model for predicting mean tree MOE were 0.23, -0.32, and 0.09 GPa for sites 803, 805, and 808, respectively. Estimates of the random effects associated with individual plots (nested within sites), ranged from -0.37 to 0.58 GPa, and are given in Table 5. The magnitude of random effects estimates is relatively small in comparison to the precision of the model of  $\pm 2.2$  GPa ( $= 1.96 \times 1.1$  GPa).

**Models estimating bole MOE**

Performance metrics for all valid bole models are shown in Table 6. With the exception of three models for which  $p$  values were less than 0.02 ( $V^2 + H$ ,  $T + A + \rho$  and  $E + D + H$ ), all other model variables listed in Table 6 were highly significant ( $p < 0.001$ ). The model in  $E + D + H$  recorded the highest degree of dependency amongst



**Fig. 1** Actual versus estimated mean tree MOE for four selected models (with fixed effects indicated at the top left of each cell). Calibration data are shown in black, validation in red. The black line represents the line of equivalence. Parallel lines are spaced at 1.0-GPa intervals



**Table 5** Random effects estimates for the recommended model (Eq. 7) for estimating mean tree MOE (GPa)

	Site 803	Site 805	Site 808
	0.23	-0.32	0.09
Treatment	Treatment within site effects		
A	-0.27	-0.37	-0.32
B	0.04	-0.17	0.28
D	-0.12	0.08	0.10
E	0.58	0.11	0.04

variables (VIF 1.9 for  $D$ ). All other models had lower dependencies (VIF less than 1.6).

The strength of the relationships, as assessed by  $R_{\text{marg}}^2$ , for estimating bole MOE ranged from 0.12 (for the model in  $A$ ) to 0.69 ( $V^2 + \rho + D + A$ ), a greater range than that observed for the tree-level models. RMSE ranged from 0.9 to 1.6 GPa, and MAPE from 6 to 12 %, similar to the tree-level models. The best model, in terms of AIC, was that in  $V^2 + \rho + D + A$ .

Performance metrics of both validation and calibration datasets indicated that  $V^2$  was a moderately good predictor

of bole MOE. Unlike the tree-level model, the apparent performance of the model for estimating mean MOE of the bole with  $V^2$  as the sole predictor was in agreement with the performance of the model with the validation dataset. The addition of  $\rho$  to the model (either as  $E = V^2 \cdot \rho$  or as  $V^2 + \rho$ ) increased TP rate for both datasets and, for the validation dataset, TP increased from 73 % ( $V^2$  alone) to 91 % ( $V^2$  and  $\rho$ ). With a marginally better AIC (372 in comparison to 373), the model in  $V^2 + \rho$  was determined to be better than the model in  $E (V^2 \cdot \rho)$ .

Overall, the model in  $V^2$ ,  $\rho$ ,  $D$ , and  $A$  recorded the best performance metrics amongst the calibration dataset, and relatively good metrics when validated; however, the simpler model in  $V^2$  and  $\rho$  recorded better validation metrics.

Five models, all of which included  $V^2$  and  $\rho$  (either as a linear combination or as a product, i.e.  $E$ ), recorded the highest TP rate of 91 %. All five models also recorded the same RMSE (1.3 GPa) and MAPE (9 %) for the validation dataset. Accuracy metrics of the calibration dataset were very similar for  $E$  and  $V^2 + \rho$ . Slightly improved accuracy metrics were obtained with the three more complex models ( $E + T$ ,  $E + D + H$ ,  $V^2 + \rho + T$ ) all with similar calibration accuracy metrics.

**Table 6** Performance metrics of linear models estimating MOE of boles. Italicised values indicate the recommended model; bold values indicate the best value(s) in each column

Fixed effects <sup>a</sup>	Calibration dataset						Validation dataset			Comments
	$R_{\text{marg}}^2$	$R_{\text{cond}}^2$	AIC	RMSE (GPa)	MAPE (%)	TP rate (%)	RMSE (GPa)	MAPE (%)	TP rate (%)	
$V^2$	0.45	0.60	432	1.2	9	60	1.3	9	73	
$D$	0.14	0.43	493	1.6	12	27	1.1	<b>7</b>	36	
$\rho$	0.18	0.43	481	1.5	11	36	1.2	8	27	
$A$	0.12	0.16	511	1.5	11	44	1.1	9	36	
$T$	0.27	0.53	487	1.5	12	51	<b>1.0</b>	<b>7</b>	46	
$E$	0.60	0.73	373	1.0	7	64	1.3	9	<b>91</b>	High TP rate
$V^2 + \rho$	<i>0.60</i>	<i>0.73</i>	372	<i>1.0</i>	7	64	<i>1.3</i>	9	<b>91</b>	<i>High TP rate</i>
$V^2 + H$	0.50	0.65	428	1.2	9	64	1.4	8	82	
$D + A$	0.31	0.34	484	1.4	10	47	1.2	10	55	
$D + H$	0.27	0.53	485	1.5	12	56	<b>1.0</b>	<b>7</b>	46	
$D + \rho$	0.31	0.59	444	1.4	10	42	1.2	9	55	
$H + \rho$	0.20	0.56	477	1.6	12	33	1.3	10	46	
$T + \rho$	0.34	0.54	454	1.3	11	56	<b>1.0</b>	8	36	
$E + T$	0.61	<b>0.78</b>	367	1.0	8	64	1.3	9	<b>91</b>	High TP rate
$T + A + \rho$	0.45	0.50	450	1.2	9	64	<b>1.0</b>	8	64	
$E + D + H$	0.61	<b>0.78</b>	368	1.0	8	67	1.3	9	<b>91</b>	High TP rate
$V^2 + D + H$	0.53	0.76	418	1.3	10	<b>69</b>	1.4	9	82	
$V^2 + \rho + D$	0.59	0.75	367	1.0	7	67	1.3	10	82	
$V^2 + \rho + T$	0.61	0.77	365	1.0	8	67	1.3	9	<b>91</b>	High TP rate
$V^2 + \rho + D + A$	<b>0.69</b>	0.74	<b>364</b>	<b>0.9</b>	<b>6</b>	<b>69</b>	1.3	10	73	Best AIC

<sup>a</sup>All variables are as defined in Table 2

Plots of actual versus predicted bole MOE of four selected models ( $V^2$ ,  $V^2 + \rho$ ,  $V^2 + \rho + T$ ,  $V^2 + \rho + D + A$ ) are provided in Fig. 2. Improvements due to augmenting the  $V^2$  model with  $\rho$  are clearly visible. Added benefits of further augmentation of this model with  $T$ , or  $D$  and  $A$ , are less apparent.

**Recommended bole model**

On the basis of the performance statistics of Table 6 and scatter plots of Fig. 2,  $V^2 + \rho$ , with a precision of about  $\pm 2.0$  GPa ( $=1.96 \times 1.0$  GPa), is the recommended model for estimating bole MOE. Fixed factors of the recommended model explained 60 % of the variance and random factors an additional 13 %. Parameter estimates, together with standard errors, of the recommended model in  $V^2 + \rho$ , for estimating bole MOE are given in Eq. 8.

$$\text{MOE} = -13.30 + 0.7852V^2 + 0.01681\rho \quad (8)$$

Std. error (1.71) (0.0600) (0.00191)

Estimates of the random effects associated with the recommended model for predicting mean bole MOE were 0.34, -0.53, and 0.19 GPa for sites 803, 805, and 808, respectively. Estimates of the random effects due to treatment nested within site are given in Table 7.

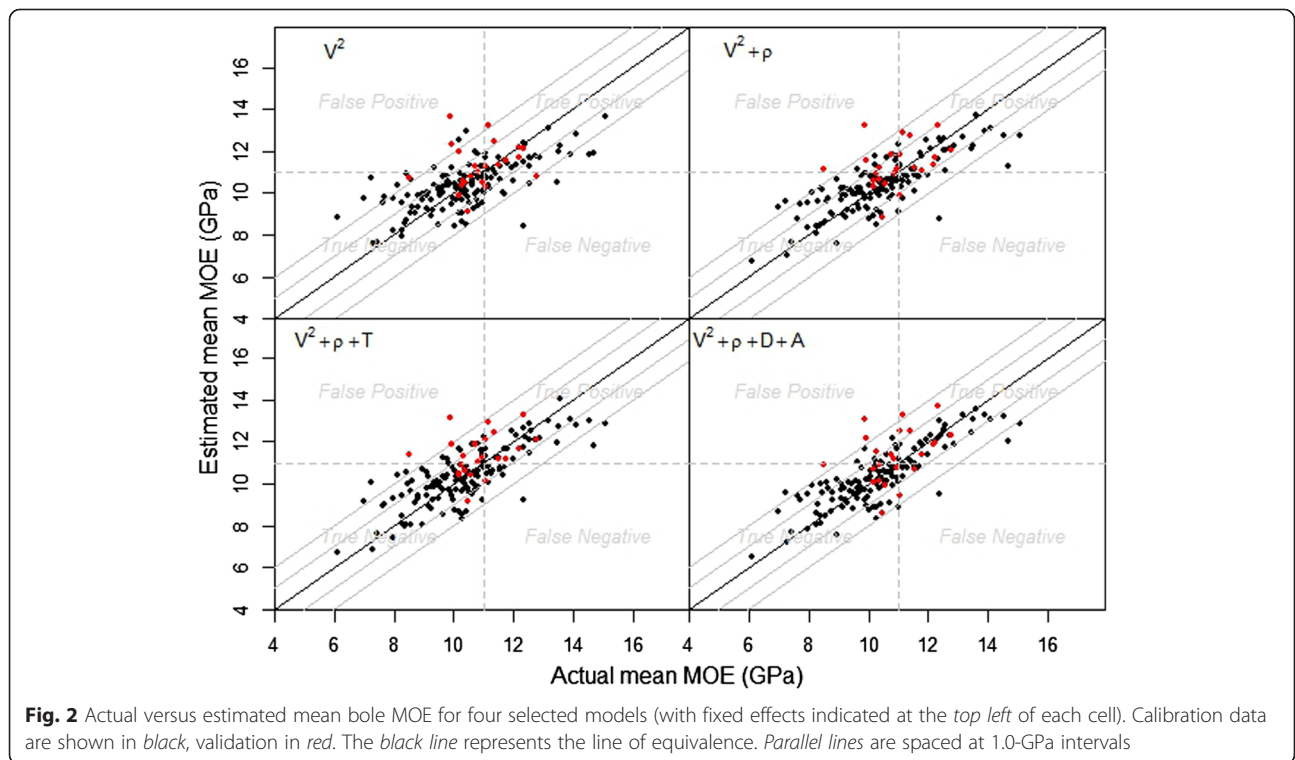
**Table 7** Random effects estimates for the recommended linear model (Eq. 8) for estimating mean bole MOE (GPa)

	Site 803	Site 805	Site 808
Treatment	Treatment within site effects		
A	-0.14	-0.19	-0.10
B	-0.06	-0.13	0.10
D	-0.25	0.00	0.00
E	0.57	0.12	0.06

**Models estimating log MOE**

Performance metrics of all valid log models are shown in Table 8. All models comprised highly significant predictor variables ( $p < 0.001$ ). Values of VIF were less than 2.5 for all but three models. The variable  $A$  had a VIF of 3.9 in the model in  $D + \rho + A$ , and 4.2 in the model in  $V^2 + \rho + A + D$ , while relative height and base diameter ( $H + D$ ) each had a VIF of 3.1. However, model behaviour was not compromised. Estimated MOE increased with increasing  $V^2$ ,  $\rho$ , and  $A$ , and decreased with increasing  $D$ ,  $H$ , and  $T$ .

The strength of the relationships, as assessed by  $R^2_{\text{margin}}$ , for estimating log MOE ranged from 0.02 (for the model in  $D$ ) to 0.64 ( $V^2 + \rho + A + D$ ). RMSE ranged from 1.1 to 1.9 GPa (the latter associated with the model in  $D$ ), and MAPE from 9 (for several models) to 15 % (the latter also associated with the model in  $D$ ).



**Fig. 2** Actual versus estimated mean bole MOE for four selected models (with fixed effects indicated at the top left of each cell). Calibration data are shown in black, validation in red. The black line represents the line of equivalence. Parallel lines are spaced at 1.0-GPa intervals

**Table 8** Performance metrics of linear models estimating MOE of logs. Italicised values indicate the recommended model; bold values indicate the best value(s) in each column

Fixed effects <sup>a</sup>	Calibration dataset						Validation dataset			Comments
	$R^2_{\text{marg}}$	$R^2_{\text{cond}}$	AIC	RMSE (GPa)	MAPE (%)	TP rate (%)	RMSE (GPa)	MAPE (%)	TP rate (%)	
$V^2$	0.43	0.63	1513	1.3	10	61	<b>1.4</b>	11	74	Best 1-variable
$D$	0.02	0.46	1695	1.9	15	2	1.7	14	0	Worst
$H$	0.06	0.58	1647	1.8	14	0	1.7	13	0	
$A$	0.09	0.55	1649	1.7	14	21	1.6	13	17	
$E$	0.51	0.57	1501	1.2	10	64	<b>1.4</b>	<b>10</b>	62	
$V^2 + \rho$	<i>0.53</i>	<i>0.61</i>	<i>1485</i>	<i>1.2</i>	<b>9</b>	<i>67</i>	<b>1.4</b>	<b>10</b>	<b>74</b>	Best validation
$V^2 + T$	0.44	0.64	1498	1.3	10	62	<b>1.4</b>	11	71	
$V^2 + A$	0.43	0.63	1509	1.3	10	61	<b>1.4</b>	11	71	
$E + D$	0.53	0.65	1471	1.2	<b>9</b>	64	<b>1.4</b>	<b>10</b>	60	
$E + T$	0.54	0.59	1480	1.2	<b>9</b>	64	<b>1.4</b>	11	57	
$E + A$	0.62	0.69	<b>1424</b>	<b>1.1</b>	<b>9</b>	71	<b>1.4</b>	<b>10</b>	69	Best equal AIC
$E + H$	0.56	0.69	1436	1.2	<b>9</b>	64	<b>1.4</b>	<b>10</b>	67	
$A + T$	0.12	0.53	1636	1.7	14	33	1.5	12	36	
$A + \rho$	0.12	0.56	1645	1.7	13	27	1.6	13	21	
$H + \rho$	0.08	0.62	1638	1.8	14	22	1.7	13	12	
$D + A$	0.25	0.54	1621	1.6	13	39	1.6	13	45	
$D + H$	0.13	0.60	1626	1.7	14	19	1.6	13	21	
$D + \rho + H$	0.17	0.62	1610	1.7	13	27	1.7	13	26	
$D + \rho + A$	0.31	0.54	1611	1.5	12	45	1.6	13	50	
$V^2 + \rho + D$	0.54	0.67	1469	1.2	<b>9</b>	67	<b>1.4</b>	<b>10</b>	69	
$V^2 + \rho + T$	0.56	0.67	1459	1.2	<b>9</b>	68	<b>1.4</b>	11	<b>74</b>	Similar to $V^2$
$V^2 + \rho + A$	0.63	<b>0.70</b>	1425	<b>1.1</b>	<b>9</b>	<b>73</b>	<b>1.4</b>	11	69	
$V^2 + \rho + H$	0.55	0.69	1439	1.2	<b>9</b>	65	<b>1.4</b>	11	69	
$V^2 + \rho + A + D$	<b>0.64</b>	<b>0.70</b>	<b>1424</b>	<b>1.1</b>	<b>9</b>	70	1.5	11	69	Best equal AIC

<sup>a</sup>All variables are as defined in Table 2

From the above statistics, it is apparent that  $D$  is a poor predictor of log MOE, even though it is highly significant as a predictor variable. Furthermore, when validated, all estimates of MOE were less than 11.0 GPa, resulting in a TP rate of 0 %. Accuracy statistics for relative height,  $H$ , were similarly poor, and when in combination with  $V^2$ , neither  $H$  (in the model  $V^2 + H$ ) nor  $D$  (in the model  $V^2 + D$ ) was significant.

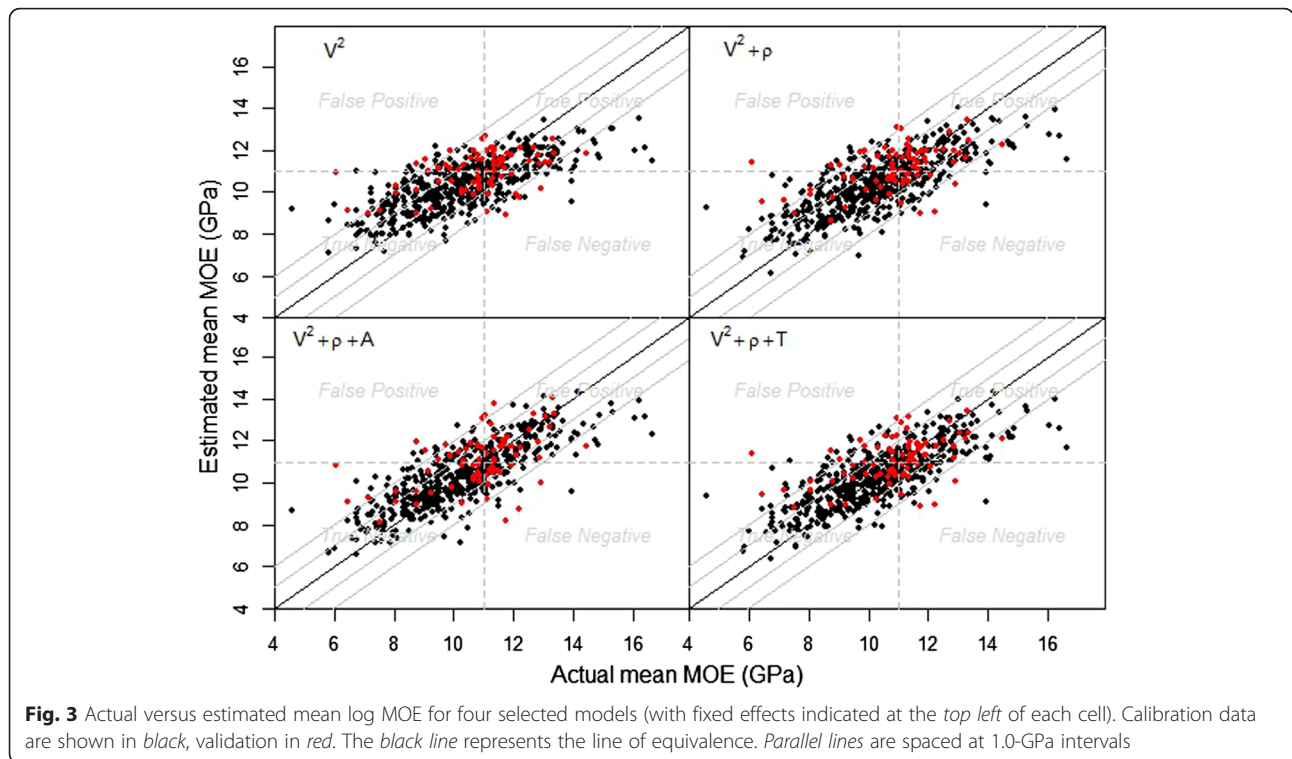
In contrast, model performance metrics indicated that  $V^2$  was a moderately good predictor of mean log MOE (RMSE 1.4 GPa, MAPE 11 %, TP rate 74 %, validation dataset). With the addition of  $\rho$ , the performance of the calibration model in  $V^2 + \rho$  improved, and a small improvement in MAPE of the validation dataset was also noted. In terms of the validation dataset, no other model recorded better performance metrics. Plots of actual versus predicted log MOE of four selected models ( $V^2$ ,  $V^2 + \rho$ ,  $E + A$ ,  $V^2 + \rho + T$ ) are provided in Fig. 3. Differences in model performance are difficult to discern.

**Recommended log model**

On the basis of the performance metrics of the validation dataset, the model in  $V^2 + \rho$ , with a precision of about  $\pm 2.4$  GPa ( $= 1.96 \times 1.2$  GPa), is the recommended model for estimating log MOE. Fixed factors of the recommended model explained 53 % of the variance and random factors an additional 8 %. Parameter estimates of the recommended model, together with standard errors are given in Eq. 9. The parameter estimate for the auto-correlation structure was 0.847.

$$\begin{aligned} \text{MOE} &= -5.267 + 0.8126V^2 + 0.006693\rho \\ \text{Std. error} &(1.201) (0.0415) (0.001060) \end{aligned} \tag{9}$$

Estimates of the random effects associated with the recommended model ( $V^2 + \rho$ ) for predicting mean log MOE were 0.26, -0.49, and 0.23 GPa for sites 803, 805, and 808, respectively. Estimates of the random effects due to treatment nested within site were minimal, ranging from -0.07



to 0.06, while those due to individual trees were considerably greater. Within-plot tree-to-tree variation in MOE ranged from 0.8 to 2.5 GPa. Random effects estimates are given for site and plot in Table 9. Random effects due to individual trees were too numerous to list, hence minimum, mean, and maximum values are provided in the table.

**Discussion**

Random effects of the recommended model for estimating log MOE (with  $V^2$  and  $\rho$  as predictors) indicated high levels of tree-to-tree variation within plots (with differences of up to 2.5 GPa; refer Table 9, plot D, site 803), minimal levels of

**Table 9** Random effects estimates for the recommended model (Eq. 9) for estimating mean log MOE (GPa)

	Site 803	Site 805	Site 808
	0.26	-0.49	0.23
Treatment	Treatment effects		
A	-0.03	-0.07	-0.05
B	0.06	0.01	0.03
D	-0.06	0.01	0.03
E	0.05	0.00	0.01
	Tree effects (min, mean, max)		
A	(-0.58, -0.08, 0.91)	(-1.03, 0.17, 1.22)	(-0.97, -0.15, 0.70)
B	(-0.57, 0.12, 0.72)	(-0.84, -0.19, 0.25)	(-0.99, 0.03, 0.68)
D	(-1.62, 0.03, 0.86)	(-0.42, 0.01, 0.40)	(-0.47, -0.13, 0.31)
E	(-0.49, 0.06, 0.43)	(-0.38, 0.08, 0.43)	(-0.98, 0.03, 1.17)

variation between plots (all close to zero), and modest levels of variation between sites (0.3, -0.5, and 0.2 GPa for sites 803, 805, and 808, respectively). This suggests, though not conclusively, that the models may be highly transportable across a variety of thinning treatments, and to a lesser extent, transportable across sites. The former suggestion is also supported by the accuracy metrics of the validation dataset, for which the logs were derived from plots thinned under a different regime. The latter suggestion needs to be investigated further because the sites used in this study were relatively similar in terms of mean annual temperature (9–11 °C), but differed in terms of mean annual precipitation (with a range of 110 to 336 mm per year).

All three recommended models for estimating tree, bole, and log MOE, included  $V^2$  and  $\rho$  as predictor variables. The tree-level model also included taper. As taper increases (as can be caused by heavy thinning regimes), tree MOE is predicted to decrease (Eq. 7). While the addition of taper to the recommended bole-level model improved the apparent performance of the model (refer Table 6), there was no improvement in external accuracy metrics. Further validation with additional external datasets is required to determine if taper improves model performance.

Fixed effects of the three recommended models explained 51–60 % of the variation in MOE, and random effects a further 8–10 %. The unexplained variation could be due to other factors, such as knots, bark, and other wood defects not examined here. Moisture content may also have been a contributing factor. Moisture content of Douglas-fir timber



when green is about 45 % (Rietz 1999) and values of 24 % are typical for air-dried timber. Though the latter moisture content, which was used in this study, may appear to be high, it should be noted that “Douglas-fir is unique among all softwood species in that it is naturally dimensionally stable, having the ability to season well in position. Many builders prefer to cut, nail and fasten Douglas-fir in the “green” or unseasoned condition, allowing it to air dry during construction” (WWPA 2002). Therefore, the MC values used in this study are accepted values, and appropriate. Beyond wood-related factors, further variation could be due to human factors associated with application of the acoustic tools. For example, Paradis et al. (2013) demonstrated that acoustic velocities are sensitive to probe depth, with increased acoustic velocity (of about  $0.25 \text{ km s}^{-1}$ ) associated with increased penetration depth (from 1.5 cm to 3.0 cm); therefore, with a team of people facilitating data collection, it is highly possible that there was variation amongst operators. However, the somewhat lower acoustic velocity of logs in this study, in comparison to those reported by Wang et al. (2013), are thought to be due to age differences of the study trees, with velocity, and MOE, being lower for younger trees.

The RMSE values obtained indicate that MOE estimates will be within about  $\pm 2$  GPa of the actual mean value, 95 % of the time. The MAPE values indicate that prediction accuracy will be, on average, out by about 10 %. The TP rates indicate that, of those trees, boles, and logs that meet or exceed the 11.0 GPa threshold, at least 64 % will be correctly segregated as high MOE wood products. There will, however, also be some low MOE material that will be incorrectly classified as high MOE. From Figs. 1, 2, and 3, we note that these, the “False Positives”, were reassuringly, relatively few in number. There were also few “False Negatives” (refer figures).

The TP rate threshold value of 11.0 GPa was somewhat arbitrary although based on the recommended tree breeding value of 11 GPa. This begets the question of how to deal with differing threshold levels. One approach would be to extend the confusion matrix approach to differing threshold values, and, for each threshold, plot the resulting receiver operating characteristics (ROC) graph depicting the relative trade-offs between benefits (true positives) and costs (false positives) (Fawcett 2006). Comparison of the ROC graphs may then provide further insight regarding model applicability and transportability.

The apparent performance of standing-tree time-of-flight velocity predictions of MOE was optimistic, and tests with the external validation dataset were disappointing. External validation indicated that predictions based on time-of-flight were less accurate than predictions based on diameter. Standing-tree time-of-flight velocity predictions were also less accurate than resonance velocity predictions of bole MOE.  $R_{\text{marg}}^2$  was 0.45 and 0.60, and TP rate of the

validation dataset 46 and 73 % for the respective models. However, augmentation of standing tree time-of-flight by wood density and tree taper improved accuracy metrics ( $R_{\text{marg}}^2$  was 0.51 and TP rate 91 %) to those comparable to the better bole-level models.

In log-level models, MOE decreased with increasing relative height. This is consistent with results found by Wang et al. (2013) for log position. However, with predictions at the population level, we found  $R_{\text{marg}}^2$  to be just 0.06, while  $R_{\text{cond}}^2$  was 0.58, similar to the value of 0.59 found by Wang et al. (2013). We note that if MOE had been predicted from relative height at the optimal level (i.e. not the population level) and a simple linear relationship formed between actual and predicted MOE, a  $R^2$ , and  $R_{\text{adj}}^2$ , of 0.74 would result. These pseudo- $R^2$  values are clearly overly optimistic, thus for random effects models we recommend the use of the two  $R^2$  variates (marginal  $R^2$  and conditional  $R^2$ ) recently developed by Nakagawa and Schielzeth (2013) and extended by Johnson (2014).

## Conclusions

A model with a good accuracy metric does not necessarily equate to good forecast accuracy as models tend to perform better on the data on which the model was developed rather than on new data. Therefore, due consideration must be given to the importance of external validation.

For accuracy in MOE estimation and tree segregation, knowledge of velocity alone is insufficient. Knowledge of wood density is also required for improved accuracy. Taper is also important to predicting tree MOE, and this study suggests that it may possibly be important to accurate prediction of bole MOE. Further validation with new data is required.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

CLT developed statistical models, analysed the Douglas-fir data, and made major contributions towards the writing of this manuscript. ECL facilitated the Douglas-fir trial and data collection processes throughout the many phases of this study and provided input to concepts culminating in this manuscript. Both authors read and approved the final manuscript.

## Acknowledgements

This research effort was supported by a Joint Venture Agreement between the New Zealand Forest Research Institute Limited, trading as Scion, and the USDA Forest Service. Support was also provided by the Sustainable Forest Productivity component of Agenda 2020, a joint effort of the USDA Forest Service Research & Development and the American Forest and Paper Association. We are grateful and wish to express our gratitude to the companies and individuals who contributed to the processing and collection of data underpinning this study. We also wish to thank the reviewers of this manuscript for their insightful comments and suggestions.

## Author details

<sup>1</sup>Scion, Private Bag 3020, Rotorua 3046, New Zealand. <sup>2</sup>USDA Forest Service, Portland Forestry Sciences Laboratory, 620 SW Main, Suite 400, Portland, OR 97205, USA.

Received: 16 March 2016 Accepted: 23 May 2016

Published online: 14 June 2016

## References

- Achim, A., Paradis, N., Salenikovich, A., & Power, H. (2010). Using acoustic tools to improve the efficiency of the forestry wood chain in eastern Canada. 'The Future of Quality Control for Wood & Wood Products', 4-7th May 2010, Edinburgh. The Final Conference of COST Action E53. <http://www.coste53.net/downloads/Edinburgh/Edinburgh-Presentation/59.pdf> Accessed 26 April 2016.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. doi:10.1109/TAC.1974.1100705.
- Altman, D. G., & Royston, P. (2000). What do we mean by validating a prognostic model? *Statistics in Medicine*, 19(4), 453–473.
- Amishev, D., & Murphy, G. E. (2008a). Implementing resonance-based acoustic technology on mechanical harvesters/processors for real-time wood stiffness assessment: opportunities and considerations. *International Journal of Forest Engineering*, 19(2), 48–56.
- Amishev, D., & Murphy, G. E. (2008b). In-forest assessment of veneer grade Douglas-fir logs based on acoustic measurement of wood stiffness. *Forest Products Journal*, 58(11), 42–47.
- ASTM. (2003). *Standard test methods for non-destructive evaluation of wood-based flexural members using transverse vibration* (D6874-03). West Conshohocken, PA: American Society for Testing and Materials.
- Auty, D., & Achim, A. (2008). The relationship between standing tree acoustic assessment and timber quality in Scots pine and the practical implications for assessing timber quality from naturally regenerated stands. *Forestry*, 81, 475–487.
- Briggs, D., Thienel, G., Turnblom, E., Lowell, E., Dykstra, D., Ross, R. J., Wang, X., & Carter, P. (2008). Influence of thinning on an acoustic velocity of Douglas-fir trees in western Washington and western Oregon. In *Proceedings of the 15th International Symposium on Nondestructive Testing of Wood* (pp. 113–123). Duluth, Minnesota, USA: Western Forestry and Conservation Association Wood Quality Workshop.
- Bucur, V. (2006). *Acoustics of wood*. Berlin, Heidelberg, Germany: Springer.
- Carter, P., Briggs, D., Ross, R. J., & Wang, X. (2005). Acoustic testing to enhance western forest values and meet customer wood quality needs. PNW-GTR-642. In C. A. Harrington & S. H. Schoenholtz (Eds.), *Productivity of western forests: a forest products focus* (pp. 121–129). Portland, USA: USDA Forest Service, Pacific Northwest Research Station.
- Collins, G. S., de Groot, J. A., Dutton, S., Omar, O., Shanyinde, M., Tajar, A., Voysey, M., Wharton, R., Yu, L.-M., Moons, K. G., & Altman, D. G. (2014). External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Medical Research Methodology*, 14, 40. <http://www.biomedcentral.com/1471-2288/14/40>.
- Curtis, R. O. (1982). A simple index of stand density for Douglas-fir. *Forest Science*, 28, 92–94.
- Davis, C. E., Hyde, J. E., Bangdiwala, S. I., & Nelson, J. J. (1986). An example of dependencies among variables in a conditional logistic regression. In S. H. Moolgavkar & R. L. Prentice (Eds.), *Modern statistical methods in chronic disease epidemiology* (pp. 140–147). New York: Wiley.
- Dickson, R. L., Joe, B., Harris, P., Holtorf, S., & Wilkinson, C. (2004). Acoustic segregation of Australian-grown *Pinus radiata* logs for structural board production. *Australian Forestry*, 67(4), 261–266.
- Divos, F. (2010). Acoustic tools for seedling, tree and log selection. The Future of Quality Control for Wood & Wood Products, 4-7th May 2010, Edinburgh. Scotland. The Final Conference of COST Action E53. <http://www.coste53.net/downloads/Edinburgh/Edinburgh-Presentation/51.pdf>. Accessed 11 May 2016.
- Evans, R., & Illic, J. (2001). Rapid prediction of wood stiffness from microfibril angle and density. *Forest Products Journal*, 51, 53–57.
- Fakopp (2006). *TreeSonic Microsecond Timer User's Guide*. Agfalva, Hungary: Fakopp Bt. [http://www.fakopp.com/site/downloads/TreeSonic\\_Guide.pdf](http://www.fakopp.com/site/downloads/TreeSonic_Guide.pdf) Accessed 6 May 2016.
- Farrell, R., Innes, T., & Nolan, G. (2008). Sorting plantation *Eucalyptus nitens* logs with acoustic wave velocity. Project No PN07.3018. Victoria, Australia: Forest and Wood Products Australia Limited. [www.fwpa.com.au](http://www.fwpa.com.au).
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874.
- Jayawickrama, K. J. S., Ye, T. Z., Gupta, R., & Cherry, M. L. (2009). *Including wood stiffness in tree improvement of coastal Douglas-fir in the US Pacific Northwest: a literature review and synthesis* (Research Contribution, Vol. 50). Corvallis, OR, USA: Forest Research Laboratory, Oregon State University.
- Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's  $R^2$  GLMM to random slopes models. *Methods in Ecology and Evolution*, 5, 944–946. doi:10.1111/2041-210X.12225.
- Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics*, 38, 963–974. doi:10.2307/2529876.
- Legg, M., & Bradley, S. (2016). Measurement of stiffness of standing trees and felled logs using acoustics: a review. *Journal of the Acoustic Society of America*, 139(2), 588. <http://dx.doi.org/10.1121/1.4940210>.
- Liu, C., Zhang, S. Y., Cloutier, A., & Rycabel, T. (2007). Modeling lumber bending stiffness and strength in natural black spruce stands using stand and tree characteristics. *Forest Ecology and Management*, 242, 648–655.
- Lowell, E. C., Todoroki, C. L., Dykstra, D. D., & Briggs, D. G. (2014). Linking acoustic velocity of standing Douglas-fir trees to veneer stiffness: a tree-log-product study across thinning treatments. *New Zealand Journal of Forestry Science*, 44, 1. <http://www.nzforestryscience.com/content/44/1/1>.
- Makridakis, S., & Hibon, M. (1995). Evaluating accuracy (or error) measures. INSEAD Working Paper Series 95/18/TM. Fontainebleau, France: INSEAD <http://sites.insead.edu/facultyresearch/research/doc.cfm?did=46875>. Accessed 26 May 2016.
- Matheson, A. C., Dickson, R. L., Spencer, D. J., Joe, B., & Illic, J. (2002). Acoustic segregation of *Pinus radiata* logs according to stiffness. *Annals of Forest Science*, 59, 471–477. doi:10.1051/forest:2002031.
- Merlo, E., Alvarez-Gonzalez, J. G., Santaclara, O., & Riesco, G. (2014). Modelling modulus of elasticity of *Pinus pinaster* Ait. in northwestern Spain with standing tree acoustic measurements, tree, stand and site variables. *Forest Systems*, 23(1), 153–166. <http://dx.doi.org/10.5424/fs/2014231-04706>.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining  $R^2$  from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142.
- Niklas, K. J., & Spatz, H.-C. (2010). Worldwide correlations of mechanical properties and green wood density. *American Journal of Botany*, 97(10), 1587–1594. doi:10.3732/ajb.1000150.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, 41, 673–690. doi:10.1007/s11135-006-9018-6.
- Paradis, N., Auty, D., Carter, P., & Achim, A. (2013). Using a standing-tree acoustic tool to identify forest stands for the production of mechanically-graded lumber. *Sensors*, 13, 3394–3408. doi:10.3390/s130303394.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & the R Development Core Team. (2013). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-111.
- R Core Team. (2013). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rietz, R. C. (1999). *Air drying of lumber*. General Technical Report FPL-GTR-117. Madison, WI, USA: USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin.
- Rozenberg, P., Franc, A., Mamdy, C., Launay, J., Schermann, N., & Bastien, J. C. (1999). Genetic control of stiffness of standing Douglas-fir; from the standing stem to the standardised wood sample, relationships between modulus of elasticity and wood density parameters. Part II. *Annals of Forest Science*, 56(2), 145–154.
- Senalik, C. A., Schueneman, G., & Ross, R. J. (2014). *Ultrasonic-based nondestructive evaluation methods for wood: a primer and historical review*. General Technical Report FPL-GTR-235. Madison, WI, USA: USDA Forest Service, Forest Products Laboratory.
- Steyerberg, E. W., & Vergouwe, Y. (2014). Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European Heart Journal*, 35, 1925–1931. doi:10.1093/eurheartj/ehu207.
- Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7, 91. doi:10.1186/1471-2105-7-91.
- Walsh, D., Strandgard, M., & Carter, P. (2014). Evaluation of the Hitman PH330 acoustic assessment system for harvesters. *Scandinavian Journal of Forest Research*, 29(6), 593–602. doi:10.1080/02827581.2014.953198.
- Wang, X. (2013). Acoustic measurements on trees and logs: a review and analysis. *Wood Science and Technology*, 47, 965–975. doi:10.1007/s00226-013-0552-9.
- Wang, X., Verrill, S., Lowell, E., Ross, R. J., & Herian, V. L. (2013). Acoustic sorting models for improved log segregation. *Wood and Fiber Science*, 45(4), 343–352.
- Willmott, C. J. (1981). On the validation of models. *Physical Geography*, 2(2), 184–194. doi:10.1080/02723646.1981.10642213.

- Winkler, R. L., & Murphy, A. H. (1992). On seeking a best performance measure or a best forecasting method. *International Journal of Forecasting*, 8(1), 104–107.
- WWPA (2002). *Douglas Fir & Western Larch Species Facts*. Portland, Oregon: Western Wood Products Association. <http://www2.wwpa.org/species/products/douglasfir/tabid/405/default.aspx> Accessed 6 June 2016.
- Zobel, B. J., & van Buijtenen, J. P. (1989). *Wood variation its causes and control*. Springer series in wood science. New York, USA: Springer-Verlag New York Inc.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---