The effect of tree (and cambium) age on genomic prediction for solid wood properties in Norway spruce

CURRENT STATUS: UNDER REVIEW

Linghua Zhou  linghua.zhou@slu.se
Sveriges lantbruksuniversitet
Corresponding Author
ORCID: 0000-0002-9814-7662

Zhiqiang Chen
Sveriges lantbruksuniversitet

Lars Olsson
RISE Research Institutes of Sweden AB

Thomas Grahn
RISE Research Institutes of Sweden AB

Bo Karlsson
skogforsk

Harry Wu
Sveriges lantbruksuniversitet

Sven-olof Lundqvist
indic

Maria Rosario Garcia Gil
Sveriges lantbruksuniversitet

DOI:
10.21203/rs.2.22694/v1

SUBJECT AREAS
Epigenetics & Genomics
Abstract

Genomic selection (GS) or genomic prediction is considered as a promising approach to accelerate tree breeding and increase genetic gain by shortening breeding cycle. We investigated the predictive ability (PA) of GS based on 484 progeny trees from 62 half-sib families in Norway spruce (Picea abies (L.) Karst.) for wood density, modulus of elasticity (MOE) and microfibril angle (MFA) measured with SilviScan, as well as for measurements on standing trees by Pilodyn and Hitman instruments. GS predictive abilities (PA) were comparable with those based on pedigree-based selection. The highest PAs were reached with at least 80-90% of the dataset used as training set. Use of different statistical methods had no significant impact on the estimated PAs. We also compared the abilities to predict density, MFA and MOE of 19 year old trees with use of models trained on data from coring at different ages and to different depths into the stem. The comparison indicated that close to the maximal PAs can be reached at age 10-12 by drilling only half way (ringwise) towards the pith, thereby reducing the impact on the tree.

Introduction

Norway spruce is one of the most important conifer species in Europe in relation to economic and ecological aspects. Breeding of Norway spruce started in the 1940s with phenotypic selection of plus-trees, first in natural populations and later in even-aged plantations. Norway spruce breeding cycle is approximately 25-30 years long, of which the production of seeds and the evaluation of the trees take roughly one-half of that time.

Genomic prediction using genome-wide dense markers or genomic selection (GS) was first introduced by Meuwissen. The method modelling the effect of large numbers of DNA markers covering the entire genome and subsequently predict the genomic value of
individuals that have been genotyped, but not phenotyped. As compared to the phenotypic mass selection based on a pedigree-based relationship matrix (A matrix), genomic prediction relies on constructing a marker-based relationship matrix (G matrix). The superiority of the G-matrix is the result of a more precise estimation of genetic similarity based on Mendelian segregation that not only captures recently pedigree but also the historical pedigree \(^{5-7}\), and corrects possible errors in the pedigree \(^{8,9}\).

There are multiple factors affecting genomic prediction accuracy such as the extent of linkage disequilibrium (LD) between the marker loci and the quantitative trait loci (QTL), which is determined by the density of markers and the effective population size \((N_e)\). An increase in accuracy with dense marker has been reported in simulation \(^{10}\) and empirical studies in multiple forest tree species including Norway spruce \(^{11-14}\), but see Kainer et al \(2018\)), whereas SNP position showed no significant effect \(^{15-17}\). Simulation \(^{10}\) and empirical \(^{18}\) studies also agree on the need of a higher marker density in populations with larger effective size \((N_e)\) in order to cover more QTLs under low LD in contributing to the phenotypic variance.

In forest tree species the accuracy of the genomic prediction model has been mainly tested in cross-validation designs where full-sibs and/or half-sibs progenies within a single generation are subdivided into training and validation sets \(^{10,19-22}\). Model accuracy was reported to increase with larger training to validation set ratios \(^{11,17,23}\), while the level of relatedness between the two sets is considered as a major factor \(^{10,15-17,19,24}\). When genomic prediction is conducted across environments, the level of genotype by environment interaction (GxE) of the trait determines its efficiency \(^{11,20,21,25}\). The number of families and progeny size have also been shown to affect model accuracy \(^{11,15}\).
As compared to the previously described factors, trait heritability and specially the trait genetic architecture are intrinsic characteristics to the studied trait in a given population. Those two factors can also be addressed by choosing an adequate statistical model depending on the expected distribution of the marker effects. Despite theory and some results indicate that complex genetic structures obtain better fit with models that assume equal contribution of all markers to the observed variation, traits like disease-resistance are better predicted with methods where markers are assumed to have different variances. However, results in forestry so far indicate that statistical models have little impact on the GS efficiency.

In this study, we conducted a genomic prediction study for solid wood properties based on data from 23-year old trees from open-pollinated (OP) families of Norway spruce. We focused on wood density, microfibril angle (MFA) and modulus of elasticity/wood stiffness (MOE) analyzed with SilviScan in the lab, and on estimates of these 3 traits based on data from measurements on standing trees of Pilodyn penetration depth and Hitman velocity of sound, estimates here named Pilodyn, Velocity and MOE$_{\text{ind}}$, respectively.

The specific aims of the study were: (i) to compare narrow-sense heritability ($h^2$) estimation, predictive ability (PA) and prediction accuracy (PC) of the pedigree-based (ABLUP) models with marker-based models based on data from measurements with SilviScan on increment cores and from Pilodyn and Hitman measurements on standing trees, (ii) to examine the effects on model PA and PC of different training-to-validation set ratios and different statistical methods, (iii) to compare some practical alternatives to implement early training of genomic prediction model into operational breeding.

Material And Methods

Plant material
The study was conducted on two open-pollinated (OP) progeny trials: S21F9021146 aka F1146 (Höreda, Eksjö, Sweden) and S21F9021147 aka F1147 (Erikstorp, Tollarp, Sweden). Both trials were established in 1990 with a spacing 1.4 m × 1.4 m. Originally, the experiments contained more than 12 progenies from 524 families at each of site, but after thinning activities in Höreda and Erikstorp in 2010 and 2008, respectively, fewer trees were left. In 2011 and 2012, six trees per site (524 * 12 ~ 6000 trees) were phenotyped and the 524 mothers were genotyped using exome capture (single nucleotide polymorphism, SNP). Standing tree-based measurements with Pilodyn and Hitman were performed on the same trees in 2011 and 2013, respectively, after which further thinning was performed. For this study, in 2018, we generated genomic (SNP) data from 484 remaining progeny trees which belonged to 62 of the OP families (out of the original 524) and eight progenies per family on average. This new genotypic data was combined with available phenotypic data for the same trees were used.

Phenotypic data

The phenotypic data was previously described in. The samples collected from the progenies in 2011 and in 2012 were taken as increment cores of 12 mm diameter from pith to bark. The samples were analyzed for from pith to bark variations in many woods and fiber traits with a SilviScan instrument at Innventia (now RISE), Stockholm, Sweden. This data is referred as increment core-based measurements through the text. The annual rings of all samples were identified, as well as their parts of earlywood, transition wood and latewood, averages were calculated for all rings and their parts and dated with year of wood formation.

The aim of breeding is not for properties of individual rings, but properties of the stem at harvesting target age. Therefore, this study focused on predictions of averages for stem
cross-sections, and we chose tree age 19 years as the reference age, with models trained on trait averages for all rings formed up to different younger ages. Three types of averages were calculated and predictions compared for density, MFA and MOE: 1) area-weighted averages, relating to the cross-section of the stem, 2) width-weighted, relating to a radius or an increment core, and 3) arithmetic averages, where all ring averages are weighted with same weight. For the calculation of area-weighted average we assumed that each growth ring is a circular around the pith, calculated the area of each annual ring from its inner and outer radii, and when calculating the average at a certain age, the trait average for each ring was weighted with the ring’s proportion of the total cross-sectional area at that age. Similarly, for the calculation of the width-weighted average, the trait average for each ring was weighted with the ring’s proportion of the total radius from pith to bark at that age. The similar results were obtained with the three average methods. For this reason, only the estimates based on the area-weighted method (the most relevant for breeding) are shown. Tree age 19 years was used as the reference age. Thus, all the selection methods investigated for density, MFA and MOE, phenotypic and genetic, were compared based on how well they predicted the cross-sectional averages of the trees at this age, with their last ring formed during the vegetation period of 2009.

In addition, estimates of the three solid wood traits were calculated based on data from Pilodyn and Hitman instruments, measured on the standing trees without removing the bark at age 22 and age 24 years, respectively. Pilodyn measures the penetration depth of a needle pressed into the stem, which is inversely correlated with wood density. Hitman measures the velocity of sound in the stem, which correlates with microfibril angle, MFA. MOE is related to wood density and velocity of sound and can therefore be estimated by combining the Pilodyn and Velocity data, which estimates we here name
MOE_{ind} (for standing-tree based). Further details on how this was performed in our study are given in 40. The references show that these standing-tree-based measurements provide useful information and are very time and cost-efficient. However, they do not allow calculation of properties of the tree at younger ages. Therefore, we were not able to investigate from what early ages such data can be uses within genomic selection.

Genotypic data

Genomic DNA was extracted from buds or needles when buds were not available. Qiagen Plant DNA extraction protocol was utilized for DNA extraction and purification and DNA quantification performed using the Qubit® ds DNA Broad Range (BR) Assay Kit (Oregon, USA). Genotyping was conducted at Rapid Genomics, USA, using exom capture methodology. Sequence capture was performed using the 40 018 diploid probes previously designed and evaluated for P. abies 41 and samples were sequenced to an average depth of 15x using an Illumina HiSeq 2500 (San Diego, USA) 31. Variant calling was performed using the Genome Analysis Toolkit (GATK) HaplotypeCaller 42 in Genome Variant Call Format (gVCF) output format. After that, the following steps were performed for filtering: 1) removing indels; 2) keeping only biallelic loci; 3) removing variant call rate (“missingness”) < 90%; 4) removing minor allele frequency (MAF) < 0.01. Beagle v4.0 43 was used for missing data imputation. After these steps, 130,269 SNPs were used for downstream analysis.

Population structure

As a first step, we conducted a principal component analysis to determine the presence of structure in our population. The spectral decomposition of the marker matrix revealed that only about 2% of the variation was captured by the first eigenvector, indicating low population structure. Therefore, population structure was not considered in the design of
cross-validation sets (see Modelling and cross-validation chapter for further details on the cross-validation sets design). Previously, low genotype by environment (GxE) interaction was detected for wood quality traits on these two trials \(^{30}\) and consequently, subsequent analyses were conducted obviating population structure and GxE.

**Narrow-sense heritability (\(h^2\)) estimation**

For each trait, an individual tree model was fitted in order to estimate additive variance and breeding values:

\[
y = X\beta + Z\mu + Wb + e
\]

where \(y\) is a vector of measured data of a single trait, \(\beta\) is a vector of fixed effects including a grand mean, provenance and site effect, \(b\) is a vector of post-block effects and \(\mu\) is a vector of random additive (family) effects which follow a normal distribution \(\mu \sim N(0, \sigma^2_\mu)\) and \(e\) is the error term with normal distribution \(N(0, \sigma^2_e)\). \(X\), \(Z\) and \(W\) are incidence matrices, \(A\) is the additive genetic relationship matrix and \(I\) is the identity matrix. \(\sigma^2_\mu\) equals to \(\sigma^2_a\) (pedigree-based additive variance) when random effect in Eq. 1 is pedigree-based in which case \(\mu \sim N(0, \sigma^2_\mu)\), and \(\sigma^2_\mu\) equals to \(\sigma^2_g\) (marker-based additive variance) when random effect in Eq. 1 is marker-based in which case \(\mu \sim N(0, \sigma^2_\mu)\). The \(G\) matrix is calculated as

\[
G = \frac{(M-P)(M-P)^T}{2 \sum_{i=1}^{q} pi(1-pi)}
\]

where \(M\) is the matrix of samples with SNPs encoded as 0, 1, 2 (i.e., the number of minor alleles), \(P\) is the matrix of allele frequencies with the \(i\)th column given by \(2(pi - 0.5)\),

\[\text{sigma}^2_u = \text{sigma}^2_a\text{sigma}^2_g\]
where $\pi$ is the observed allele frequency of all genotyped samples.

Pedigree-based individual narrow-sense heritability ($h_a^2$) and marker-based individual narrow-sense heritability ($h_g^2$) were calculated as

$$h_a^2 = \frac{\sigma_a^2}{\sigma_{pa}^2}, \quad h_g^2 = \frac{\sigma_g^2}{\sigma_{pg}^2}$$

respectively, $\sigma^2_{pa}$ and $\sigma^2_{pg}$ are phenotypic variances for pedigree-based and marker-based models, respectively.

**Selection of the optimal training and validation sets ratio**

Cross-validation was conducted after dividing randomly the whole dataset into a training and a validation set. To find the most suitable ratio between the two, we divided the data into sets with five different ratios between the training and the validation sets: 50, 60, 70, 80 and 90%. 100 replicate iterations were carried out for each tested ratio and trait.

**Statistical method for model development**

In the same context we aimed to find optimal methods. Several statistical methods were compared: pedigree-based best linear unbiased predictions (ABLUP), and four GS methods: genomic best linear unbiased predictions (GBLUP), random regression-best linear unbiased predictions (rrBLUP), BayesB, and reproducing kernel Hilbert space (RKHS).

rrBLUP used a shrinkage parameter lamda in a mixed model and assumes that all markers have a common variance. In BayesB the assumption of common variance across marker effects was relaxed by adding more flexibility in the model. RKHS does not assume linearity so it could potentially capture nonadditive relationships. R package rrBLUP...
was used for GBLUP and rrBLUP, package BGLR \(^{49}\) was used for BayesB and RKHS. The pedigree-based relationship matrix was obtained with the R package pedigree (Coster2013).

**PA and accuracy estimation**

The adjusted phenotypes \(y' = y - X\beta\) were used as model response in the genomic prediction models. Model quality was evaluated by predictive ability (PA), which is the mean of the correlation between the adjusted phenotype and the model predicted phenotypes, \(r(y',\hat{y})\) from 100 times CV. Prediction accuracy (PC) was defined as \(PA/\sqrt{h^2}\) \(^{15, 50}\). In order to investigate whether GS model training can be conducted at earlier age, PA at each tree calendar age and cambial age were estimated. In this case, cross validation was conducted only using area-weighted values at each age, then the trait values at each age were estimated. PA at a specific age was calculated as the correlation between estimated trait values at that age and area-weighted values from pith to the last ring (for cambial age) and last year (for calendar age), respectively.

Genomic selection for well-performing trees with the use of marker information (G matrix) requires access to previously trained GS models. Thus, model training is a necessary part of GS integration into operational breeding. Model training can be conducted in already existing plantations with trees of relatively high ages, as illustrated in this work. It is, however, expected and desired that such model training can be conducted with high PAs also for younger trees. This would be especially useful if maturity (flower production) can be accelerated, to shorten the total breeding cycle.

Operationally, it is also important to develop protocols to assess wood quality in resources at minimum cost and time, and with minimal impact on the trees. Therefore, on coring, it is not only important to know the minimum age at which useful information can be
obtained, but also from how many rings from the bark towards the pith information is required to train models with high predictive ability. To address these two practical questions for operational breeding, we trained prediction models based on data from different sets of rings, in order to mimic and compare PAs obtained when coring at different ages of the trees to different depths into the stem, or more precisely, using data from different numbers of rings, starting next to the bark. All the models were judged on, compared by their ability to predict the cross-sectional average of the trait at age 19 years across all trees in the validation set.

Result

Narrow-sense heritability ($h^2$) of the phenotypic traits, predictive ability (PA) and Predictive accuracy (PC) based on pedigree and maker data

In Table 1, narrow sense heritabilities ($h^2$) and Prediction Abilities (PA) based on ABLUP and GBLUP are compared for density, MFA and MOE based on cross-sectional averages at age 19 years, and for Pilodyn, Velocity and MOE$_{ind}$ based on measurements with the bark at age 22 and 24 years, respectively. For density, MOE and Pilodyn, $h^2$ did not differ significantly between estimates based on the pedigree (ABLUP) and marker-based (GBLUP) methods taking standard error into account. For MFA, the pedigree-based $h^2$ was lower than the GBLUP estimate while for Velocity and MOE$_{ind}$, the pedigree-based $h^2$ was higher. When using pedigree, the order of the traits by $h^2$ agrees with their order by PA estimates, but this was not the case based on markers. Traits with higher $h^2$ tended to show also high PA estimates irrespective of the method. The ABLUP PA estimates were similar to the GBLUP estimates for density and Pilodyn, while for the rest of the traits GBLUP delivered slightly higher PA estimates, and significantly higher for MFA. The relative performances
of ABLUP compared to GBLUP differed for MOE, Velocity and MOE\textsubscript{ind}. The h\textsuperscript{2} estimates for MOE were similar for both methods, while the PA estimate based on marker data was higher for GBLUP. In the case of Velocity and MOE\textsubscript{ind}, a higher h\textsuperscript{2} based on pedigree contrasted with a slightly higher PA estimates based on marker data. Standardization of the PAs with the h values did not change the conclusions on the relative efficiencies of pedigree versus marker data-based estimates.

Marker-based PA and PC between increment core-based and standing-base wood quality traits

The marker-based PAs were generally 25–30% higher for traits density, MFA and MOE measured with SilviScan than for their respective standing tree-based method which measured with Pilodyn and Hitman. Concordantly, the h\textsuperscript{2} values were 46%, 65% and 55% higher based on Silviscan methods, respectively. However, if we compare PC of the increment core- and standing tree-based methods, they were similar, and PC of MOE\textsubscript{ind} was even higher than that for MOE using GBLUP.

Effects on PAs of the GS models ratios between the training and validation sets, and from the statistical method used

For density and Pilodyn, all five methods resulted in very similar PA estimates across the ratios, while rrBLUP and GBLUP seemed superior for the rest of the traits, and mostly so for Velocity and MOE (Fig. 1). The rest of the analysis were conducted based on the GBLUP modelling method.

PAs on estimation of traits at reference age with models trained on data available at earlier ages

For density, the estimated PAs showed a rising trend within a span of about 0.25–0.30 for
the models based on both age types, after the first years. But the year-to-year fluctuations were more intense for models based on data organized on tree age. As MFA typically develops from high values at the lowest cambial ages via a rapid decrease to lower and more stable values from cambial age 8-12 years and on, one may expect that models trained on data from only low ages would have difficulties to predict properties at age 19 years. This was also confirmed. We even obtained some negative PA values at early ages, such as years 1995 and 1996, and the PAs for cambial age-based models started from very low values, then increasing. The curves for MOE showed PAs developing at values in between those for density and MFA. This is logical, as MOE is influenced by both density and MFA, with particularly negative effects from the high MFAs at low cambial ages. At cambial age 13, MFA and MOE showed a drop in the cambial age-based PA estimates. Generally, the Figure indicates that genomic selection for density could be conducted at an earlier age than for MFA and MOE.

Search for optimal sampling and data for training of GS prediction models

Figure 2 showed estimated PAs of models trained on data from sampling different years, using data from all rings available at that age (except for the innermost ring). In this section we instead estimate how much PAs can be reached if the increment core is not drilled all the way to the pith, but to different shorter depths, reducing the injury to the tree, Fig. 3a-d. This analysis was preformed based on tree age data only, as the cambial age of a ring can only be precisely known if the core is drilled to the pith, allowing all rings to be counted.

Each row of the figures represents a tree age when cores are samples, starting at age 3 years when the first 60 trees formed a ring at breast height, ending at the bottom with the reference age 19 years with 17 rings. Each column represents a depth of coring, counted in numbers of rings. As one more ring is added each year, thus also to the
maximum possible depth on coring, the tables are diagonal. The uppermost diagonal represents models trained on data from the 60 (12%) trees which had reached breast height at age 3. The diagonal next below represents models based on the 243 (51%) trees with rings at age 4, etc. The PAs shown below the three uppermost diagonals represent models trained of data from more than 90% of the trees. The PAs were calculated from the cross-validation, based on data from the trees on which the respective models were trained. This means that the PAs of the three uppermost diagonals are based only on fast-growing trees not fully representative for the trials. Many of the highest PAs found occur along these diagonals. Due to their trees’ special growth, only PAs based on more than 90% of the trees will be further commented.

For wood density, Fig. 3b, the variations in predictability show an expected general pattern: The PAs increased with the tree age on coring, and also with the depth, the number of rings from which the cross-sectional averages were calculated and exploited on training of the prediction models. The highest values, 0.29, are obtained at age 19 years, but then also data from the reference year are included on training the prediction model. An example of quite high PAs at lower ages and depths: For coring at tree ages 10–12 years and using data from the 3–5 outermost rings, all alternatives gave PA values of 0.26–0.29.

For MFA, a trait with low heritability, the PA values are low as already shown in Fig. 2 and the pattern in Fig. 3c is not easy to interpret. Here, the same set of alternatives of samples at tree ages 10–12 and depths 3–5 gave PA values of 0.15–0.18, compared to the maximum of 0.19 among all alternatives using 90% of the trees. The values are lower at the highest ages. Streaks of higher and lower values can be imagined along the diagonals. The pattern for MOE is similar to that of MFA, but on higher level. Training on data from coring at ages and to depths as above gave PA values of 0.20–0.23, compared to the
corresponding maximum of 0.25.

Discussion

We have conducted a genomic prediction study for solid wood properties assessed on increment cores from Norway spruce trees with SilviScan derived data from pith to bark, using properties of annual rings formed up to tree age 19 years as the reference age. We also compared this with genomic prediction of proxies for density, MFA and MOE estimated with data from same trees measured at the bark of the standing trees with Pilodyn and Hitman instruments. The study was conducted on 62 open-pollinated (OP) families. On Norway spruce operational breeding, the use of OP families is preferable because it does not require expensive control crosses. The only action required is to collect cones where progenies are typically assumed to be half-sibs. Thus, OP families permit evaluation of large numbers of trees at lower cost and efforts than structured crossing designs. We investigated narrow-sense heritability estimation with ABLUP and marker-based GBLUP and the effect on predictive ability (PA) from using different training-to-validation set ratios, as well as different statistical methods. Further, we investigated what precision of GS can be reached when training the models with data from trees of different ages, also comparing results for the solid wood properties to those for their proxies. We also estimated PAs reached when coring to different depths from the bark at different tree ages, in order to find cost-effective methods for GS with minimum impact on the trees on the acquisition of data for training the prediction models.

Narrow-sense heritability ($h^2$)

In our study, PA estimates for both pedigree and marker-based methods were consistent with their respective $h^2$ estimates. A conifer literature review indicates that the level of consistency varies across studies $^8, 18−20$. When comparing between the ABLUP and GBLUP
methods, our estimates for density, MOE and Pilodyn the $h^2$ were on similar levels for Velocity and MOE$_{ind}$ higher based on ABLUP and for MFA higher for GBLUP. Using GBLUP the estimated $h^2$ values were clearly lower for the standing tree-based methods than for measured density, MFA and MOE: 54%, 35% and 45% lower, respectively. In a previous study conducted on full-sib progenies in Norway spruce, however, the ABLUP-based $h^2$ were reported higher in all three standing-tree-based measurements $^{11}$. Instead, other conifer studies based on full- or half-sib progenies reported a comparable performance of A-matrix and G-matrix based methods in Pinus taeda $^{18, 23}$, Douglas-fir $^{29}$ and Picea mariana $^{15}$ for growth related traits and wood properties. Moreover, ABLUP accuracies were lower for growth, form and wood quality in Eucalyptus nitens $^{24}$. Experimental design factors such as number progenies and their level of coancestry, statistical method and the traits and pedigree errors under study may account for the apparent inconsistency in the relative performance of both methods $^{51}$.

Our results indicate that for more heritable traits ABLUP and GBLUP capture similar levels of additive variance, whereas for traits with very low heritability using ABLUP, such as MFA, the markers are able to capture additional genetic variance probably in the form of historical pedigree reflected in the G matrix. Less obvious is the case for Velocity and MOE$_{ind}$ where GBLUP seems to capture lower values of additive variance. It is possible that at intermediate values of $h^2$ the benefits of capturing historical consanguinity is overcome by possible confounding effects caused by markers which are identical by state (IBS) or simply due to genotyping errors. The $h^2$ values obtained with ABLUP and GBLUP is the result of a balance between multiple factors such as the genetic structure of the trait, the historical pedigree, and the possible model overfitting to spurious effects or genotyping
Effects on GS model predictive ability (PA) of training-to-validation sets ratios and statistical methods

In conifers and Eucalyptus cross-validation is often performed on 9/1 training to validation sets ratio \(^8, 12, 15, 16, 28\). This coincide with the general conclusion from the present study, with exception for MFA and MOE, for which the best results were obtained at ratio 8/2. It has been suggested that when the trait has large standard deviation, more training data is needed to cover the variance in order to get high predictive ability \(^52\). So, for density, Pilodyn and Velocity, PA kept increasing with the size of the training set. But for other traits with smaller standard deviation, (4.44 and 2.28 for MFA and MOE), PA decreased when increasing the training set from 80–90%, which may indicate that too much noise was introduced during model training.

The fact that the estimated PAs for all the solid wood properties as measured in the lab instrument are 25–30% higher than their proxies estimated from measurements of penetration depths and sound velocity at the bark may reflect the indirect nature of their proxies: the correlations calculated for the almost 6000 trees initially sampled were between \(-0.62\) Pilodyn and density, \(-0.4\) between Velocity and MFA and 0.53 between \(\text{MOE}_{\text{ind}}\) and MOE \(^40\).

In the conifer literature it has more often been reported similar performance of different marker-based statistical models for wood properties \(^11, 12, 18, 28, 53\). This general conclusion agrees with our findings for all our traits with the exception of Velocity and to a less extent \(\text{MOE}_{\text{ind}}\). For these, GBLUP and rrBLUP performed better than the other GS methods, which could be the result of a highly complex genetic structure where a large number of genes of similar and low effect are responsible of the control of the trait. For
traits affected by major genes the variable selection methods, for example BayesB or LASSO, have been reported to perform better \(^{18}\), whereas for additive traits the use of nonparametric models may not yield the expected accuracy \(^{54}\).

Comparison of PA and PC from methods based on pedigree (ABLUP) and markers

Generally, pedigree-based PA estimates in conifer species have been reported to be higher or comparable to marker-based models \(^{11, 15, 16, 19, 20, 23}\), but there are also some studies reporting marker-based PA estimates to be higher \(^{13, 24, 55}\). Our results for density and Pilodyn follow the general finding in forest trees, whereas for MFA, a low heritability trait, the PA estimate based on GBLUP model is substantially higher (0.16) compared to the ABLUP model (0.04). As we already discussed in relation to \(h^2\), additional genetic variance is being captured by the G kinship matrix. When PA is standardized with \(h\), the predictive accuracies of the methods become more similar across traits, indicating that proportionally similar response to GS can be expected for all traits.

Use of tree age versus cambial age (ring number)

From a quick look at Fig. 2, one may get the impression that breeding based on cambial age data allows earlier selection than using tree age data. That would however be a too rushed conclusion. At tree age 3 years, after the vegetation period of 1993, only 12.5% of the trees had formed a first annual ring at breast height. Not until tree age 6 years, more than 90% of the trees had done so, that is by number 5 years “higher age”. In contrast, all the trees had obviously a ring of cambial age 1. But if aiming for 90% representation, one must wait several years more for more rings were formed at breast height, more precisely, from 1993 to end of growth season 1996, tree age 6. And to train models based on data from 90% of the trees for cambial age say, 6 at breast height, samples cannot be collected until the end of growth season at tree age 11 years, or if a representation of
80% is judged as satisfactory, at tree age 10 years. This has to be considered if selection efficiencies are calculated based on cambial age data, which is common. Such results have for instance been published based on the almost 6000 trees sampled 2011 and 2012.

Correctly compared based on minimum 90% of the trees, the estimated PAs shown in Fig. 2 are similar between the age alternatives, or slightly better for use of tree age. For example, the PA for MOE using cambial age data shows a smooth increase, reaching above 0.2 at cambial age (ring number) 7, which needs data from the tree of age 12. The corresponding curve from using tree age passed above 0.2 already at age 8 years. However, curves based on tree age often show larger year-to-year variation. This is most likely an effect of the fact that the rings of same cambial age represent wood formed across a span of years with different weather. Thus, cambial age data reflect annual weather across a range of years, which does not happen when using tree age data. On the other hand, from a practical point of view, methods based on use of tree age may be easier to apply in operational breeding, especially in light of results as in Fig. 3, indicating that high PAs can be reached without coring all the way to the pith. To number the rings for precise cambial age, you need to find the innermost ring at the pith, but that may not be necessary for good results.

Implementation of GS for solid wood into operational breeding

The results indicate that GS can result in similar early selection efficiency or even higher than traditional pedigree-based breeding and offers further possibilities. Previously, in loblolly pine it was reported that models developed for diameter at breast height (DBH) and height with data collected on 1 to 4-year old trees had limited accuracy in predicting phenotypes at age 6-year old 21. In British Colombia Interior spruce, the predictive
accuracy for tree height of models trained at ages 3 to 40 years, at certain intervals, and validated at 40 years revealed less opportunities for early model training, since the plateau was not reached until 30 years \(^{28}\).

In our study, the highest PA values were obtained for the subsets of fast-growing trees which had reached breast height already at tree age 3 and 4 years, 12\% and 51\% of the total number of trees, representing a limited number of the OP families included in the analysis. The trees of this fast-growing group are affected by high intensity of selection for alleles accelerating growth within each OP family. Also, on cross-validation the prediction abilities for this group were calculated based on the trees within the same group. In this elite group different factors could account for a higher PA value, such as lower phenotypic variance, decreased number of alleles of minor effect could also facilitate identification of major effects and/or higher consanguinity between those families which may share alleles for growth. These models are shown for completeness, but as they cannot be used for operational breeding they are not further discussed.

Models for genetic selection are useful in different steps of a breeding program. A first type of prediction models, here illustrated with Table 1, can be trained from existing trials, preferably based on trees of as old age as available, as the aim of the breeding is to predict tree qualities at age of harvesting when the major part of the stem will be dominated by mature wood. Training the models in older trees for wood properties also allows considering other properties which cannot be easily observed from trees of very young age, such as stem straightness and health. For wood density, the results indicate that models can be built without coring very deep into the stem. It may be expected that this is valid also for instance for tracheid dimensions which in combination determines the wood density \(^{34}\).
Due to the juvenility of the trees used in our study, the wood includes considerable proportions of both high MFA core wood and of lower MFA outer wood, with a pronounced shift in between. The results indicate that for such trees, information from the area around where the shift occurs is more useful for training of the GS model than information from the outermost rings, albeit resulting in generally low prediction abilities. This is an unexpected result. It may be expected that models for MFA and MOE trained on older trees would rely more on data from the mature wood outside the shift, similar to the models for density. Apart from models predicting traits of mature wood, ideally in trees of harvesting age, there is also an interest for models focusing on local features which can be detrimental for quality of wood-based products, such as traits of the inner core with high MFA and low MOE, and the size of this core with problematic properties for many solid wood products.

As illustrated in this work, two aspects of incorporating wood properties into operational GS breeding programs can be addressed with the same set of data. Firstly, as mentioned above, models for cost-effective selection based on genomic information from existing trees. In that case, models from data at old ages would normally be preferred, for example for wood density some model at bottom line of Fig. 3a. Secondly, models providing guidance on at what age it is reasonable to approach young trees for training of GS models for specific traits: a) trees in existing juvenile trials, or b) trees of new generations with different pools of genetics. As an example, the same Fig. 3b for wood density suggests GS model training at tree ages 10 to 12 on the third to fifth outermost rings to reduce costs and the negative impact on the tree.

Conclusions

1. In comparison with phenotypic selection, Genomic selection methods showed similar to higher prediction abilities (PAs) for both increment core- and standing tree-based
phenotyping methods. This indicates that the standing tree-based measurements may be a cost-effective alternative method for GS, but higher PAs were obtained based on increment core-based wood analyses.

2. Different genomic prediction statistical methods provided similar PA. At least 80% data should be included in the training set in order to reach the highest levels of PA

3. This study represents the first published investigation of the efficiency of GS with prediction models trained on data acquired from sampling/coring trees at different ages, combined with sampling/coring to different depths, to optimize the operational breeding for the combination of length of breeding cycle, cost and impact on the trees. The results indicate that similar efficiency can be obtained without coring all the radius to the pith.

Declarations

Contributions

LZ analysed data and drafted the manuscript. ZC designed sampling strategy, coordinated field sampling and edited the manuscript. BK participated in the selection of the breeding populations, providing access to field experiments and edited the manuscript. LO, TG conducted SilviScan measurements and edited the manuscript. HW conceived and designed the study and edited manuscript. SOL and RRG provided ideas and revised manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to acknowledge the UPSC Vinnova Center of Forest Biotechnology. We also acknowledge the Swedish Research Program Bio4Energy, the Swedish Foundation for Strategic Research (SSF) and RISE for their support in phenotypic and genotypic data collection.

Ethics approval and consent to participate
The plant materials analysed for this study comes from common garden experiments that were established and maintained by the Forestry Research Institute of Sweden (Skogforsk) for breeding selections and research purposes. Three tree breeders in Sweden were co-authors in this paper. They agreed to access the materials.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**References**

1. Hannrup, B., et al., *Genetic parameters of growth and wood quality traits in Picea abies*. Scandinavian Journal of Forest Research, 2004. 19(1): p. 14–29.

2. Skogforsk, *Strategi för framtida skogsträdsförädling och framställning av förädlat skogsodlingsmaterial i Sverige*. 1995.

3. Karlsson, B. and O. Rosvall, *Progeny testing and breeding strategies*. Proceedings of the Nordic group for tree breeding. 1993: Edinburgh.

4. Meuwissen, T.H., B.J. Hayes, and M.E. Goddard, *Prediction of total genetic value using genome-wide dense marker maps*. Genetics, 2001. 157(4): p. 1819-29.

5. Bouvet, J.-M., et al., *Modeling additive and non-additive effects in a hybrid population using genome-wide genotyping: prediction accuracy implications*. Heredity, 2016. 116(2): p. 146 %@ 1365-2540.

6. El-Dien, O.G., et al., *Implementation of the realized genomic relationship matrix to open-pollinated white spruce family testing for disentangling additive from nonadditive genetic effects*. G3: Genes, Genomes, Genetics, 2016. 6(3): p. 743-753 %@ 2160-1836.

7. El-Kassaby, Y.A., et al., *Breeding without breeding: is a complete pedigree necessary
for efficient breeding? PLoS One, 2011. 6(10): p. e25737 %@ 1932-6203.

8. Munoz, P.R., et al., Genomic relationship matrix for correcting pedigree errors in breeding populations: impact on genetic parameters and genomic selection accuracy. Crop Science, 2014. 54(3): p. 1115-1123 %@ 0011-183X.

9. Tan, B., et al., Genomic relationships reveal significant dominance effects for growth in hybrid Eucalyptus. Plant science, 2018. 267: p. 84-93 %@ 0168-9452.

10. Grattapaglia, D. and M.D.V. Resende, Genomic selection in forest tree breeding. Tree Genetics & Genomes, 2011. 7(2): p. 241-255 %@ 1614-2942.

11. Chen, Z.-Q., et al., Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. BMC genomics, 2018. 19(1): p. 946 %@ 1471-2164.

12. Isik, F., et al., Genomic selection in maritime pine. Plant Science, 2016. 242: p. 108-119 %@ 0168-9452.

13. Kainer, D., et al., Accuracy of Genomic Prediction for Foliar Terpene Traits in Eucalyptus polybractea. G3: Genes, Genomes, Genetics, 2018. 8(8): p. 2573-2583 %@ 2160-1836.

14. Zapata-Valenzuela, J., et al., SNP markers trace familial linkages in a cloned population of Pinus taeda—prospects for genomic selection. Tree Genetics & Genomes, 2012. 8(6): p. 1307-1318 %@ 1614-2942.

15. Lenz, P.R.N., et al., Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (Picea mariana). BMC genomics, 2017. 18(1): p. 335 %@ 1471-2164.

16. Müller, D., P. Schopp, and A.E. Melchinger, Persistency of Prediction Accuracy and Genetic Gain in Synthetic Populations Under Recurrent Genomic Selection. G3: Genes|Genomes|Genetics, 2017. 7(3): p. 801-811.
17. Tan, B., et al., *Evaluating the accuracy of genomic prediction of growth and wood traits in two Eucalyptus species and their F1 hybrids*. BMC plant biology, 2017. 17(1): p. 110 %@ 1471-2229.

18. Resende, M.F.R., Jr., et al., *Accuracy of genomic selection methods in a standard data set of loblolly pine (Pinus taeda L.).* Genetics, 2012c. 190(4): p. 1503-1510.

19. Beaulieu, J., et al., *Accuracy of genomic selection models in a large population of open-pollinated families in white spruce*. Heredity, 2014a. 113(4): p. 343-352.

20. El-Dien, O.G., et al., *Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing*. BMC genomics, 2015. 16(1): p. 370 %@ 1471-2164.

21. Resende Jr, M.F.R., et al., *Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments*. New Phytologist, 2012b. 193(3): p. 617-624 %@ 0028-646X.

22. Resende, M.D.V., et al., *Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees*. New Phytologist, 2012a. 194(1): p. 116-128 %@ 1469-8137.

23. Zapata-Valenzuela, J., et al., *Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine*. G3: Genes, Genomes, Genetics, 2013. 3(5): p. 909-916 %@ 2160-1836.

24. Suontama, M., et al., *Efficiency of genomic prediction across two Eucalyptus nitens seed orchards with different selection histories*. Heredity, 2019. 122(3): p. 370 %@ 1365-2540.

25. Beaulieu, J., et al., *Genomic selection accuracies within and between environments and small breeding groups in white spruce*. BMC genomics, 2014b. 15(1): p. 1048.

26. Daetwyler, H.D., et al., *Genomic prediction in animals and plants: simulation of data,
validation, reporting, and benchmarking. Genetics, 2013. 193(2): p. 347-365 @ 0016-6731.

27. de Almeida Filho, J.E., et al., The contribution of dominance to phenotype prediction in a pine breeding and simulated population. Heredity, 2016. 117(1): p. 33 @ 1365-2540.

28. Ratcliffe, B., et al., A comparison of genomic selection models across time in interior spruce (Picea engelmannii x glauca) using unordered SNP imputation methods. Heredity, 2015. 115(6): p. 547 @ 1365-2540.

29. Thistlethwaite, F.R., et al., Genomic selection of juvenile height across a single-generational gap in Douglas-fir. Heredity, 2019. 122(6): p. 848-863.

30. Chen, Z.-Q., et al., Inheritance of growth and solid wood quality traits in a large Norway spruce population tested at two locations in southern Sweden. Tree Genetics & Genomes, 2014. 10(5): p. 1291-1303.

31. Baison, J., et al., Genome-wide association study identified novel candidate loci affecting wood formation in Norway spruce. The Plant Journal, 2019. 100(1 @ 0960-7412): p. 83-100.

32. Zhou, L., et al., Genetic analysis of wood quality traits in Norway spruce open-pollinated progenies and their parent plus trees at clonal archives and the evaluation of phenotypic selection of plus trees. Canadian Journal of Forest Research, 2019. 49(7): p. 810-818.

33. Evans, R., Rapid Measurement of the Transverse Dimensions of Tracheids in Radial Wood Sections from Pinus radiata, in Holzforschung - International Journal of the Biology, Chemistry, Physics and Technology of Wood. 1994. p. 168.

34. Lundqvist, S.-O., et al., Age and weather effects on between and within ring variations of number, width and coarseness of tracheids and radial growth of young
35. Downes, G.M., et al., Relationship between wood density, microfibril angle and stiffness in thinned and fertilized Pinus radiata. Iawa Journal, 2002. 23(3): p. 253-265 %@ 0928-1541.

36. Lenz, P., et al., Genetic improvement of white spruce mechanical wood traits—early screening by means of acoustic velocity. Forests, 2013. 4(3): p. 575-594.

37. Haines, D.W. and J.-M. Leban, Evaluation of the MOE of Norway spruce by the resonance flexure method. Forest products journal, 1997. 47(10): p. 91 %@ 0015-7473.

38. Knowles, R.L., et al., Evaluation of non-destructive methods for assessing stiffness of Douglas fir trees. New Zealand Journal of Forestry Science, 2004. 34(1): p. 87-101.

39. Lindström, H., P. Harris, and R. Nakada, Methods for measuring stiffness of young trees. Holz als Roh-und Werkstoff, 2002. 60(3): p. 165-174 %@ 0018-3768.

40. Chen, Z.-Q., et al., Estimating solid wood properties using Pilodyn and acoustic velocity on standing trees of Norway spruce. Annals of Forest Science, 2015. 72(4): p. 499-508.

41. Vidalis, A., et al., Design and evaluation of a large sequence-capture probe set and associated SNPs for diploid and haploid samples of Norway spruce (Picea abies). bioRxiv, 2018: p. 291716.

42. McKenna, A., et al., The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res, 2010. 20(9): p. 1297-303.

43. Browning, S.R. and B.L. Browning, Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized
haplotype clustering. Am J Hum Genet, 2007. 81(5): p. 1084-97.

44. VanRaden, P.M., Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science, 2008. 91(11): p. 4414-4423.

45. Whittaker, J.C., R. Thompson, and M.C. Denham, Marker-assisted selection using ridge regression. Genetics Research, 2000. 75(2): p. 249-252 %@ 1469-5073.

46. Hayes, B.J. and M.E. Goddard, Prediction of total genetic value using genome-wide dense marker maps. Genetics, 2001. 157(4): p. 1819-1829 %@ 0016-6731.

47. Heslot, N., et al., Genomic Selection in Plant Breeding: A Comparison of Models. Crop Science, 2012. 52: p. 146-160.

48. Endelman, J.B., Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. The Plant Genome, 2011. 4: p. 250-255.

49. Pérez, P. and G. de los Campos, Genome-wide regression and prediction with the BGLR statistical package. Genetics, 2014. 198(2): p. 483-495.

50. Dekkers, J.C.M., Prediction of response to marker-assisted and genomic selection using selection index theory. Journal of Animal Breeding and Genetics, 2007. 124(6): p. 331-341.

51. Vela-Avitúa, S., et al., Accuracy of genomic selection for a sib-evaluated trait using identity-by-state and identity-by-descent relationships. Genetics Selection Evolution, 2015. 47(1): p. 9.

52. Isidro, J., et al., Training set optimization under population structure in genomic selection. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik, 2015. 128(1): p. 145-158.

53. Thistlethwaite, F.R., et al., Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. BMC Genomics, 2017. 18(1): p. 930.
54. Desta, Z.A. and R. Ortiz, *Genomic selection: genome-wide prediction in plant improvement*. Trends in Plant Science, 2014. 19(9): p. 592-601.

55. El-Dien, O.G., et al., *Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (Picea glauca x engelmannii)*. Molecular Breeding, 2018. 38(3): p. 26.

**Table**

Table 1 Trait heritability, predictive ability (PA) and predictive accuracy (PC) Predictive accuracy (PC) for density, MFA and MOE cross-sectional averages at tree age 19 years, for their proxies on the stems without removing the bark at tree ages 21 and 22 years. Standard errors are shown in within parenthesis.

| Trait     | Narrow-sense heritability (standard error) ($h^2$) | Predictive ability (standard error) (PA) | Predictive accuracy (PC) |
|-----------|--------------------------------------------------|-----------------------------------------|--------------------------|
|           | ABLUP    | GBLUP   | ABLUP    | GBLUP   | ABLUP   |
| density   | 0.70 (0.18) | 0.69 (0.15) | 0.30 (0.01) | 0.29 (0.03) | 0.36     |
| MFA       | 0.04 (0.08) | 0.17 (0.13) | 0.04 (0.01) | 0.16 (0.02) | 0.20     |
| MOE       | 0.27 (0.14) | 0.31 (0.15) | 0.15 (0.01) | 0.22 (0.02) | 0.29     |
| Pildyn    | 0.35 (0.15) | 0.32 (0.14) | 0.22 (0.01) | 0.20 (0.01) | 0.37     |
| Velocity  | 0.16 (0.12) | 0.11 (0.10) | 0.10 (0.01) | 0.13 (0.01) | 0.25     |
| MOEind    | 0.31 (0.14) | 0.17 (0.13) | 0.17 (0.01) | 0.19 (0.01) | 0.31     |

ABLUP=pedigree-based Best Linear Unbiased Predictor (BLUP); GBLUP= genomic-based BLUP.

**Figures**
Predictive ability obtained with different ratios of training set and validation set, using different statistical methods.
Estimated Predictive abilities (PA) for prediction of cross-sectional averages at tree age 19 years, based on cross-sectional averages at different tree ages (upper graphs) and cambial ages (lower graphs) from pith to bark.
Figure 3

Predictive ability from bark to pith at different tree ages (y-axis) and an increasing number of rings included in the estimation (x-axis). 3a) Number of trees at each tree age with different number of rings. 3b) PA of density at each tree age with different number of rings. 3c) PA of MFA at each tree age with different number of rings. 3d) PA of MOE at each tree age with different number of rings.