# Estimation of individual knot volumes by mixed effects modelling

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Forest Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjfr-2019-0038.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>18-Apr-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Manso, Rubén; Forest Research McLean, J. Paul; Forest Research Ash, Adam; Forest Research Achim, Alexis; Universite Laval</td>
</tr>
<tr>
<td>Keyword:</td>
<td>Knot geometry, Seemingly Unrelated Regression, Sitka spruce, mixed models, BLUPs</td>
</tr>
<tr>
<td>Is the invited manuscript for consideration in a Special Issue? :</td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>
Estimation of individual knot volumes by mixed effects modelling

Rubén Manso¹,*, J. Paul McLean¹, Adam Ash¹, and Alexis Achim²

¹Forest Research, Northern Research Station, Roslin, Midlothian, EH25 9SY, UK

²Département des sciences du bois et de la forêt, Université Laval, 2425 rue de la Terrasse, Québec, QC G1V 0A6, Canada

*e-mail: ruben.manso@forestry.gov.uk, rmgforestal@hotmail.com;

Tlf.: +44 (0)300 067 5979

November 11, 2019
ABSTRACT

We present a new method to estimate individual knot volumes based on a knot geometry model coupled with observations on branch characteristics. X-ray Computer Tomography and image analysis were used to measure the volume and geometry of 424 knots of Sitka spruce. Knot geometry can be described mathematically by deriving functions for relative vertical position, diameter and slope dependent on radial position in the stem. These functions were parameterised using “Seemingly Unrelated Regression” and mixed-modelling techniques. This provided a base model for typical knots. In order to estimate individual-knot volume, we used available data for branch diameter and insertion angle to obtain conditional predictions. We imputed the most likely knot trajectory as relative vertical position cannot be measured on branches. The model explained up to 96% of the variability in knot volume by incorporating the branch measurements in contrast to the 43% using the typical-knot model. Additionally, knot volume assessment based only on conditional predictions of diameter and marginal predictions of vertical position also accounted for 96% of the variability. Therefore measurements of branch diameter alone would be enough to obtain highly precise individual-knot volume predictions. This estimator is a first step towards a knot model to be used for management purposes of Sitka spruce in Great Britain.

Keywords: Knot geometry, Seemingly Unrelated Regression, Sitka spruce, mixed models, BLUPs
The occurrence of knots in the wood of commercial species is a long standing issue for the sawn-wood industries. Knot frequency and size, often jointly referred to as knot area ratio, are considered as major detrimental factors for the quality of sawn wood (Zink-Sharp 2003), which is the end product out of which the highest value of the harvested tree is generally obtained. As a result of their economic impact, learning where, when and how knots occur has been a topic of research for some time, mainly through studies on the branching habit of the species of interest (Achim et al. 2006; Auty et al. 2012; Colin and Houllier 1992; Mäkinen and Colin 1998; Osborne and Maguire 2016, among many others). These studies typically require a lot of manual measurements of branches and the aim is to produce models that describe variables such as branch diameter and angle of insertion/inclination, that can be used to infer knot geometry.

An alternative to the measurement of branches and making inferences about knots is the application of X-ray Computer Tomography to directly measure knots in situ (e.g. Moberg 2000, 2001). Knot volume can now be measured from X-ray images of logs (Pinto et al. 2003) and used to model a multitude of parameters describing the shape of knots (Moberg 2006), which have recently been modelled for various North-American conifer species (Duchateau et al. 2015, 2013). The models by Duchateau et al. (2013) included a model for the z coordinate of knot trajectory (hereafter referred to as relative vertical position) and a knot diameter one. Their models can be described as specific to individual knots, as they provide predictions of knot geometry at the knot level over the stem radius (i.e. from pith to bark, or bark to pith) with an expansion of model parameters as a function of several external tree- and branch-related predictors. Among the latter, the more relevant ones are branch angle, height and diameter at the point of insertion into the
stem. This is useful as knot geometry can consequently be better predicted from measurements of trees that can be obtained in the forest. The downsides are that quite a range of branch measurements are required and that the combined effects of these branch measurements on the model parameters have a composite effect, and cannot be easily discerned on either a mathematical or a biological basis. Therefore an alternative empirical approach that can more parsimoniously account for such variability is desirable.

One possible alternative is to use a mixed modelling approach (see Lappi and Bailey 1988). In a mixed modelling approach the fixed effects can be used to produce marginal predictions, which represent the mean of a given process if the model is linear, while the random effects can be used to represent the variability around the mean at a given grouping level. In our case we modelled knot geometry as a function of radial position (the fixed component) and we allowed this to vary by individual knots (the random component and grouping level). Individual knots differ from each other for many reasons that we could not feasibly measure (e.g. light availability of individual branches), but using only a relatively small number of measurements the random components of each knot can be quantified through their Best Linear Unbiased Predictors (BLUPs) (Goldberger 1962; Henderson 1975) and projected onto new individuals in the wider population. We ultimately want to predict knot geometries and volumes from external measurements that can be made on trees: the diameter of branches (knot diameter) and the angle of insertion (knot tangent) can be measured, whereas the relative vertical position of knots cannot. This is because knot relative vertical position represents the difference in vertical height between any point in the knot trajectory and knot origin at tree pith, the latter being obviously invisible from outside the tree. These terms are depicted in Fig 1. However, in order to calculate volume we need diameter and relative vertical position for each given radial position, whereas the tangent plays no direct role in the
calculation of volume. Herein lies the problem. Our proposed solution is based on
the fact that the knot relative vertical position, diameter and tangent deviations from
model marginal predictions are not likely to be independent (e.g. a steeper knot may
tend to reach higher and have a larger diameter as well). In other words, random
effects from the processes that define knot geometry can —and ideally should—
be allowed to correlate in the model fit, which implies a simultaneous parameter
estimation of the models representing those processes. It would then be possible
to exploit the random effect correlation to impute the value of the missing random
effect(s). Here we attempt to use knot diameter and knot tangent for these purposes
and the theoretical approach to attain this is developed in this paper.

(Inset Fig 1 here)

In this study, we will consider the example of planted forests of Sitka spruce
(Picea sitchensis (Bong.) Carr.) in Great Britain. Sitka spruce is the most important
coniferous species in Great Britain in terms of area and harvested volume (50% and
60% of the conifer total, respectively; Forestry Commission (2016)) and plays a
crucial role in the economy, particularly that of rural areas. Consequently, there is
an applied interest to understand knot geometry and volume in the species, both for
purposes of sawn wood production and for evaluation of the potential for emerging
markets such as biorefineries, which use raw materials from forests to produce a
wide range of chemicals (Clark et al. 2006). Some of the desirable chemicals can
be found in considerable quantities in conifer knots as extractives (Holmbom et al.
2003; Kebbi-Benkeder et al. 2017)

The main aim is to test the approach for individual knot geometry and volume
estimation through the prediction of random effects and produce a model that can
be made readily available for management purposes of Sitka spruce stands in Great
Britain. The hypothesis is that volume predictions obtained through this approach will be more precise than marginal predictions where only the typical (i.e. the mean) knot geometry and volume over the stem radius are estimated. The main objective of the present study was to test this hypothesis. In order to do this we were required to: (i) develop a model for knot geometry where error terms of the different sub-models are allowed to correlate, (ii) implement the theoretical methods required to predict individual knot volume, (iii) produce the necessary outputs for comparison between typical knot and individual knot volume predictions, and (iv) test the relevance of each geometry model component (knot relative vertical position, diameter and tangent) in individual knot volume prediction.

2 MATERIALS AND METHODS

2.1 DATA

Six Sitka spruce trees were sampled from two sites in Scotland: two mature trees were sampled from a planted stand on the grounds of the Forest Research’s Northern Research Station, Roslin, and nominally represented specimens of large and small diameter. Four additional trees of varying diameter at breast height (DBH) were sampled from Aberfoyle Forest. The trees from both sites were growing in pure even-age stands at an initial 2 m spacing between plants (2500 stems ha\(^{-1}\)) , which are the typical conditions for Sitka spruce planted forests in Great Britain. All trees were cross-cut into logs nominally 1 m in length, resulting in 88 logs with diameters ranging from 50 to 200 mm. Their internal structure was captured using X-ray CT scanning using the equipment and services of Scotland’s Rural College (SRUC). A series of cross-sectional X-ray images were collected in 1.5 and 5 mm spirals along log length for all logs from Roslin and Aberfoyle sites, respectively. The Roslin samples were measured first and it was determined that 5 mm resolution was
considered enough for the purposes of this study. The Roslin data were simplified accordingly. Following the methods of Duchateau et al. (2013) the scans from each log were visualised as image series and descriptively tagged using the Gourmands plugin (Colin et al. 2010) within the open source ImageJ image analysis software (Rasband 1997). The software enabled us to manually mark the locations of the pith and knots within each image and subsequently turned this data into coordinates describing the paths of each knot and estimates of their diameters as they develop from pith to bark Fig 1. Visualization was also possible by means of Bil3D software (Colin et al. 2010) (Fig 2).

Subsequently, data were analysed to detect potential measurement errors and epicormic shoots mistakenly considered as true knots. Both potential errors and epicormics shoots were manually examined and these records removed (111 false knots found). After this process, 424 knots remained to carry out the intended analysis.

(Insert Fig 2 here)

2.2 MODELLING APPROACH

The modelling of knot volume consisted of three stages: (i) modelling knot geometry features, (ii) assessment of correlation between these features, and (iii) knot volume estimation.

In the first step of our analysis, the data described before were used to model the three main features that define knot geometry: relative vertical position of knot pith, knot diameter and tangent of knot pith, all three over the stem radius. Preliminary models were fitted to examine and select suitable functional forms. These responses were modelled as a function of stem radius and different transformations of this variable, such as fractional, power or exponential transformations, allow-
ing for a different behaviour depending on whether the target knot belonged to a
whorl or not. Based on visual evidence, we classified a given knot as a whorl
knot when its vertical distance to the nearest knot taken at tree pith was less than
10 mm. Tree DBH and absolute height up the tree were also tested as covariates,
although no significant effect was found. These tests were carried out through a
visual check to standardized model residuals. Residual variances were modelled
to correct from heteroscedastic residual patterns when sensible. In the following
formulae, the different observations within the knots and the knots themselves are
assigned the indices $i$ and $j$ respectively. Next the indices $m = 1, 2, 3$ are assigned to
the knot relative vertical position, diameter and tangent models, respectively. The
preliminary models were as follows

$$y_{1,ij} = b_1 \frac{r_{ij}}{r_{ij}+1} + b_2 r_{ij} W_{ij} + b_3 r_{ij} (1 - W_{ij}) + e_{1,ij}$$ (1)

$$y_{2,ij} = g_0 (1 - e^{r_{ij}}) + g_2 r_{ij} + e_{2,ij}$$ (2)

$$y_{3,ij} = q_0 + q_1 (1 - q_2)^{r_{ij}} + e_{3,ij}$$ (3)

where $y_{m,ij}$ and $e_{m,ij}$ are the responses and the residual errors respectively for
knot relative vertical position, diameter and tangent models. $r_{ij}$ stands for stem ra-
dius, $W_{ij}$ is a dummy variable that adopts the value 1 when the knot belongs to a
whorl and 0 otherwise. $\beta^T = \beta$, $\gamma^T = \gamma$ and $\theta^T = \theta$ are vectors of estimable pa-
rameters. $e_{m,ij}$ are assumed normally distributed such as $e_{m,ij} \sim N(0, \sigma_m^2)$, where
$\sigma_m^2$ is the model residual variance. In the case of the relative vertical position model,
the residual variance was modelled as a function of $r_{ij}$ to meet the homoscestatic-
ity assumption, such that $\sigma_1^2 = \sigma_0^2 \cdot e^{2\phi r_{ij}}$, with $\phi$ an additional parameter to be
estimated. Note that these functions are empirical and therefore the parameters
themselves hold little meaning.

These models provide typical or mean knot features conditional on a given set of covariates through their parameters that are considered fixed effects. The description of the variance around that mean can be best provided with the incorporation of random effects (Gregoire et al. 1995) in a mixed modelling approach. This is the second stage of our analysis. Our hypothesis is that these random effects would be correlated and therefore we can take this correlation into account through so-called “Seemingly Unrelated Regression” (SUR) (Gallant 1987). The principle of SUR is to couple various models through dummy variables and weight their residual variances according to the response type. Once each of the preliminary models in Eqs. 1, 2 and 3 were determined to behave satisfactorily independently, SUR was used as a second analytical step to carry out a simultaneous fit which will then allow random effects to correlate between models. Different structures of random effects at the knot level were tested, the standardised residuals being further checked at each step. Expressing the covariates of the three models in their matrix form as \( X_m \), the definitive model structure with random effects under the SUR formulation yields

\[
\mathbf{y}_m = q_m^T \begin{pmatrix}
X_1 \beta \\
\gamma(X_2, \gamma) \\
X_3 \theta
\end{pmatrix} + q_m^T \begin{pmatrix}
\frac{r_{ij}}{r_{ij} + 1} b_{1,j} \\
(1 + e^{\gamma r_{ij}}) b_{2,j} \\
(1 - \theta_2)^{r_{ij}} r_{ij} b_{3,j}
\end{pmatrix} + q_m^T \mathbf{e} \tag{4}
\]
\[ y_m = \begin{pmatrix} y_{1,ij} \\ y_{2,ij} \\ y_{3,ij} \end{pmatrix}, \quad q_m = \begin{pmatrix} q_{1,ij} \\ q_{2,ij} \\ q_{3,ij} \end{pmatrix}, \]
\[ \varepsilon = \begin{pmatrix} e_{1,ij} \\ e_{2,ij} \\ e_{3,ij} \end{pmatrix} \overset{\text{iid}}{\sim} N_3(0, R), \]
\[ b_j = [b_{1,j}, b_{2,j}, b_{3,j}]^T \overset{\text{iid}}{\sim} N_3(0, G), \]

where \( q_m \) is a vector of dummy variables that alternatively adopt the value 1 for a given \( m \) and 0 otherwise; \( b_{1,j}, b_{2,j}, \) and \( b_{3,j} \) are the elements of the vector \( b_j \) of the random effects at the knot level and \( R \) and \( G \) are the variance-covariance matrices of the distribution of the residuals and the random effects, respectively. In this case, \( R \) can be factorized as \( \sigma_m^2 W^2 \), where \( W \) is a diagonal matrix of the variance weights for the different models with elements \( w_m \) (Pinheiro and Bates 2000). Constant weights for each model corresponding to the variance function \( g(w_m) = \delta_m \) are compulsory in SUR, \( \delta_m \) being a vector of additional parameters to estimate. Elements of \( w_1 \) are conventionally set to 1 (thus \( \delta_1 \) as well) so that \( \sigma_m^2 = \sigma_1^2 \delta_m \). Additionally, we still set \( \sigma_1^2 = \sigma_1^2 \cdot e^{2\phi_{ij}} \) in order to deal with heteroscedasticity in the first model. All parameters were estimated through the maximization of model likelihood.

The third analytical step focused on the use of parameter estimates from the relative vertical position and diameter models in Eq. 4 to assess the total volume of a given knot \( (V_j) \). Each knot was considered as a series of consecutive and tapered, cylinders. The cross sectional area of the cylinder was considered as the cross sectional area at the mid point between two consecutive predicted sections. This cross sectional area was assumed elliptical. Although vertical and horizontal diameters of knots measured in a vertical section can be assumed identical in the
INDIVIDUAL KNOT VOLUME ESTIMATION

species (Pearson correlation coefficient 0.99), relative to the major knot axis, the vertical diameter is greater than the horizontal diameter (Achim et al. 2006). Hence the cross sections perpendicular to knot axis must be an ellipse. The minor axis of the ellipse $R_{\text{min}}$ is equal to the diameter that was measured (equation 5a) and the major axis $R_{\text{max}}$ must be calculated geometrically (equation 5b). The length of the section $l$ is the length of the line between the centre of the two consecutive sections (equation 5c). Cylinder volumes can then be numerically integrated to produce knot volume $V_j$ (equation 5d)

\[
R_{\text{min},i+1,j} = \frac{d_{i+1,j} + d_{i,j}}{2} \quad (5a)
\]
\[
R_{\text{max},i+1,j} = R_{\text{min},i+1,j} \cos \eta_{i,i+1,j} \quad (5b)
\]
\[
l_{i,i+1,j} = \sqrt{(h_{i+1,j} - h_{i,j})^2 + (r_{i+1,j} - r_{i,j})^2} \quad (5c)
\]
\[
V_j = \sum_{i=1}^{K-1} \pi R_{\text{min},i+1,j} R_{\text{max},i+1,j} l_{i,i+1,j} \quad (5d)
\]

where $d_{ij}$, $h_{ij}$ and $r_{ij}$ are the $i^{th}$ observation on diameter, relative vertical position and location at the stem radius in knot $j$, respectively. $\eta_{i,i+1,j}$ is the angle of the slope between sections $i$ and $i + 1$ of knot $j$, and $K$ is the number of segments in which the knot is divided. $h_{ij}$ and $d_{ij}$ were estimated using Monte Carlo simulation from the relative vertical position and diameter models. The theoretical distribution of the volume of a typical knot can be obtained through a sufficient number of Monte Carlo simulations (10,000 in this case). The expected value for knot volume at a given stem radius would be the mean of such distribution. All sources of uncertainty were considered, i.e. fixed effects, random effects and residual variance. Fixed-effect and random-effect realizations were held constant within knots. Models in this section were fitted with package nlme (Pinheiro et al. 2013) in R (R Core Team).
2015).

2.3 **Volume estimation of individual knots through prior relative vertical position sampling**

In the previous subsection, methods were developed to predict knot volume for a given stem radius. At this stage we can consider that we have the ability to describe typical knots in a stem. However, we also want to quantify how observations of branches will affect the volume of individual knots. Therefore, the next step was to develop methods for individual knot volume that incorporate observed branching parameters, i.e., insertion diameter and insertion angle of branches. Those methods rely upon the prediction of the random effects in Eq. 4 for new observations.

It is well known that the prediction of random effects in a linear context can be carried out through the BLUPs approach as

\[
\hat{b} = GZ^T\hat{V}^{-1}(y - X\hat{\beta})
\]

(6)

where \(Z\) is the matrix of the covariates affected by the random effects, \(\hat{V}\) is the estimated matrix of the variance of the model, which can be calculated as \(\hat{V} = Z\hat{G}Z^T + \hat{R}\), \(y\) is a vector with new observations and \(X\hat{\beta}\) are the marginal model predictions for the individuals that \(y\) belongs to. Ideally, we could use the branch geometry at the point of branch insertion into the stem as new observations \(y\). In our case, we considered knot geometry at the stem periphery from our own dataset as a proxy (see Table 2 for some basic statistics).

Again, it is not feasible to obtain knot relative vertical position from measuring dimensions on the tree. Therefore, the relative vertical position was systematically sampled over its observed range, each simulation being used as an observation itself. As a result, a set of predicted random effects was obtained for each combi-
nation of observed diameter and tangent, plus sampled relative vertical position. Each of these sets can be seen as a sample of the probability density distribution of the random effects. The set of maximum density corresponds to that of the most likely relative vertical position given the observations on knot diameter and knot tangent at the stem periphery. We thereby used these optimal BLUPs as a means to obtain conditional predictions on individual knot relative vertical position and knot diameter over the stem radius. Subsequently, knot volume was estimated following the Monte Carlo procedure introduced in the previous section. In order to test the robustness of the approach, we additionally computed individual knot volume by imputing the most likely relative vertical position with the diameter-related random effects only (i.e. ignoring the tangent), and simply combining conditional knot diameter predictions and marginal relative vertical position predictions.

(Insert Table 2 here)

2.4 Evaluation of Typical and Individual Knot Volume Estimator

The different versions of the SUR were compared through their AICs (Akaike Information Criterion). Knot volume estimates based on the marginal predictions from the SUR model were evaluated by examining the empirical coverage of their confidence intervals. Confidence intervals for the prediction of each knot were calculated through the percentile rank method out of 10,000 Monte Carlo simulations. The empirical coverage was assessed as the proportion of the knots whose observed volume fell within those confidence intervals for different levels of significance $1 - \alpha$ or “nominal coverage”. For example, if $\alpha = 0.05$, 95% of their corresponding confidence intervals should contain the observed volume. If the predictions are unbiased, then nominal and empirical coverage should coincide. The observed vol-

https://mc06.manuscriptcentral.com/cjfr-pubs
volume was computed through Eq. 5d using the measurement data from the X-ray CT
scans and it was assumed to be a good estimate of the true volume.

Additionally, absolute and relative estimation bias ($E$, $E_r$), root mean square error ($RMSE$) and estimation efficiency ($EF$, Soares et al. (1995)) were calculated by comparing the expected volume of each individual knot in our dataset ($\hat{V}_j$) with their respective observed volume ($V_j$). $\hat{V}_j$ was computed out of 10 000 Monte Carlo simulations. In order to determine the relative merit of each of the introduced approaches, these metrics were computed for typical knot volume predictions (i.e. using the procedure in subsection 2.2) and for the predictions from the different versions of the individual knot volume estimator (i.e. following the steps in subsection 2.3).

3 RESULTS

The fitting process of the model for knot geometry under SUR (Eq. 4) successfully converged. The resulting parameter estimates are shown in Table 1. In light of the correlations among the random effects, it is possible to conclude that knots with a steeper profile (i.e. higher relative vertical position and tangent than the average) tend to exhibit larger diameters. Also, deviations in relative vertical position vary positively with deviations in tangent as stem radius increases. These results mirror the empirical correlations found among geometry components at the stem periphery (Table 2).

In order to illustrate the expected evolution of knot relative vertical position, diameter and tangent along with stem radius, simulations based on the model were run for a hypothetical knot that belongs to a whorl occurring at a stem radius of 150 mm (Figs. 3, 4 and 5).
The empirical coverage probability closely matched their nominal values (Fig. 6). This indicates that model predictions and their confidence intervals are reasonably unbiased.

Concerning individual knots, volume estimations carried out by combining conditional knot diameter predictions and marginal knot relative vertical position led to a massive improvement with respect to the mean knot volume predictions (96% and 43% of the variability explained, respectively). Observed and predicted values from both approaches are shown in Figs. 7 and 8. In contrast, additionally using the BLUPs from the diameter model or the tangent model to impute the most likely trajectory did not lead to further improvements of individual knot volume predictions. This information is summarized in Table 3.

4 DISCUSSION

A novel approach that allows for individual knot volume estimation was introduced. First, it was possible to simultaneously fit mixed models for knot geome-
try that eventually served to estimate the volume of typical knots. The process of simultaneous fitting allowed the assessment of between-model random effect correlations, which helped describe how individual knot geometry and consequently volume deviate from the mean predicted values. This information was combined with pseudo external branch observations (i.e. derived from knot measurements at stem periphery) to predict model random effects and, in turn, individual knot volume with a high degree of precision. All of these aspects contribute to the originality of this paper and to the state of the art in the field of knot modelling.

The most complete empirical work on knot geometry based on in situ measurements so far (Duchateau et al. 2013) sought to describe relative vertical position and diameter of individual knots as functions of stem radius. The parameters of these functions were in turn made dependent on different tree and branch level covariates, which may not be always available. We circumvented this problem through the aforementioned mixed-modelling approach, which drastically reduced the number of variables needed to model individual knot geometry and volume. We consider this reduction as a major advantage over previous empirical approaches in this field.

The functional form used to model knot relative vertical position and diameter in Duchateau et al. (2013) was defined so that some boundary conditions were met (e.g. relative vertical position and diameter at the pith had to be zero and relative vertical position at the point of branch insertion into the stem should match that observed). In our case, preliminary tests with these equations and others aimed at imposing those conditions resulted in either unacceptable patterns in model residuals or convergence failure of the fitting algorithm. Consequentially, a simpler and more flexible approach was trialled that still proved satisfactory in terms of the model assumptions. These unbounded functional forms may not be adequate for other species or growing conditions, therefore this is an aspect that requires further research.
As far as we are aware, simultaneous parameter estimation of the models defining knot geometry had not previously been attained. The merit of this approach has already been studied in the context of branch geometry, although with no success (Mäkinen and Colin 1998). Simultaneous fitting allows between-model error correlations to be taken into account. In our case, these correlations allowed the prediction of random effects even when a critical external observation parameter (relative vertical position) was omitted to represent the fact that it is physically unobtainable. This ultimately led to more realistic predictions of knot characteristics. Neglecting between-model error correlations can be of further importance when it comes to error propagation to obtain knot volume. For example, here knots steeper than the average were observed to have the tendency to have also larger cross-sections than the average. Simulations carried out through uncorrelated random effects could therefore result in predictions where relative vertical position increases notably while diameter for the same knot remains at low levels. The estimated volume and volume prediction intervals of such a knot would be definitely unrealistic.

Unsurprisingly, incorporating branch observations to estimate individual knot volume largely outperformed the simple application of the relative vertical position and diameter models to obtain typical knot volume, in line with what has been observed in other areas of forest science (see, for example Calama and Montero 2004; Mehtätalo 2004). Interestingly, imputing the most likely relative vertical position by means of the correlation of the random effects yielded similar results to simply using the conditional knot diameter predictions plus the marginal knot relative vertical position. An evaluation of the predicted most likely relative vertical position shows a relative improvement in terms of explained variability when knot diameter observations were used (45%, no imputation; 55%, imputation) but not further improvement when the knot tangent observations were additionally considered (57%).
The reason why the 10% improvement when knot diameter observations were used did not translate into a better volume prediction may be due to how and where knot volume is allocated, plus how the random effects affect knot trajectory. Relative vertical position increases linearly over the whole stem radius except in the very close proximity to the stem pith, where knot diameter is nearly negligible and therefore so is the share of volume. Further, here the single random effect included in the relative vertical position model only affects the parameter linked to the term $r/(r + 1)$, which basically governs the height at which knot trajectory evolves in a linear fashion. In consequence the individual knot trajectories, defined by that random effect, do not have a major impact in terms of volume as they do not differ in the linear part, which accounts for most of the volume. This is in contrast to the conditional predictions of knot diameter. As it can be seen in Fig. 4, the uncertainty around the expected value for a typical knot dramatically increases already in the proximity to the tree pith. Most of this uncertainty is due to the variance of the associated random effect. This means that the between-knot variability in terms of knot diameter is considerable all along the knot trajectory. In consequence, the conditional predictions of knot diameter have a great impact on knot volume, as observed in the simulations. In summary, while knot diameter greatly increases the accuracy of our predictions of knot volume, we do not require the sampling approach to infer a more precise knot trajectory. This technique may be useful for other criteria where knot trajectories are relevant, like the proportion of the height of the stem affected by knots. The degree of improvement to be expected is limited though.

Our findings also showed that diameter behaviour over stem radius of a given knot can accurately be obtained from a single diameter measurement if the model is unbiased and the random effects are properly specified. It is important to note that our model may not meet these requirements under a more active management, such as intensive thinning, which could temporarily change the rate of growth of
the branch. Otherwise, the possibility of using a single measurement under the conditions for which the model was fitted is most convenient as branch diameter at the point of branch insertion into the stem is well known to be an excellent proxy of knot diameter at the periphery of the stem (e.g. Duchateau et al. 2013). Contrastingly branch insertion angle and knot angle at the stem periphery may exhibit a poor correlation (Duchateau et al. 2013; Lemieux et al. 2001). In spite of this limitation, modelling insertion angle is still a major goal in many studies on knot features, some of them dealing with knot volume (e.g. Duchateau et al. 2013; Fredriksson 2012; Mäkinen and Colin 1998; Osborne and Maguire 2016, among many others).

Here we have demonstrated that insertion angle can safety be omitted when predicting knot volume, at least in the type of planted Sitka spruce forests investigated, which represents a worthwhile advantage over other modelling alternatives. Thus investigating whether this behaviour occurs in other species and/or under other management regimes may prove highly valuable.

From a practical perspective, we believe our approach is highly applicable in practical forest management. Sitka spruce is a species of high economic interest in Great Britain, with a wide range of uses where knot occurrence is seen either as a disadvantage (e.g. structural timber) or as an opportunity (e.g. chemical extractive production, Moore 2011). In either case, a prior assessment of knot geometry and volume in standing timber may prove useful. As shown, both features can be ascertained if branch diameter at insertion is known. This metric could be obtained by a direct manual measurement at breast height, or there is potential that this may be derived for a larger portion of the tree bole using laser scanning technology (Dassot et al. 2012; Hackenberg et al. 2015; Liang et al. 2016). Consequently knot assessment is possible at the forest inventory stage and our model makes it possible to consider knots in future planning. Alternatively, branch diameter at insertion could be obtained from existing models that were developed to predict the number, distri-
bution and diameter of branches in Sitka spruce in Great Britain (Achim et al. 2006; Auty et al. 2012), in an attempt to integrate branches in dynamic growth models. Further, these models could be coupled with that proposed here to estimate the impact of different silvicultural choices on knot volume. This would help maximize the profit at the end of the rotation. In the case of structural timber, an optimal silvicultural scheme may lead to a reduction in the number and size of knots, which in turn would enhance timber stiffness. This mechanical property is well-known to be the limiting factor when grading Sitka spruce timber (Moore 2011). Consequently, addressing this issue may significantly increase the value of British timber.

In the case of chemical extractive production, the goal would be to promote knots. Given the current development of the bioeconomy business, this could rise the benefits from poor or exposed sites where the grades for structural timber cannot be reached.

While the parameterisation presented here could provisionally be used for the aforementioned purposes, a more robust model calibration based on a larger sample of knots from different locations would be desirable. Additionally, it is important to note that the sample trees used in this study did not originate from the British selective breeding programme for Sitka spruce, aimed at improving growth rates and stem form including branching. According to Lee (2001) a considerable reduction in the number and size of branches has been observed in improved individuals. Given that improved material has been used in the vast majority of the new plantations of the species in Great Britain over the last 25 years (Rook 1992), an extended sample for model recalibration would need to include improved trees as well.
AKNOWLEDGEMENTS

The authors wish to express their gratitude to Madeleine Murtagh for her contribution with image processing, the staff at SRUC for X-ray CT scanning and Frédéric Mothe (Unité Mixte de Recherche Silva) for use of his software. Special thanks to Tom Connolly, Tommaso Locatelli, Gustavo López and Helen McKay (Forest Research) for their useful comments on an earlier version of this paper.

REFERENCES


INDIVIDUAL KNOT VOLUME ESTIMATION


INDIVIDUAL KNOT VOLUME ESTIMATION


INDIVIDUAL KNOT VOLUME ESTIMATION


### TABLES

Table 1. Maximum likelihood parameters estimates and standard errors of the SUR model. For explanation of parameters refer to equations 1-4 in the text.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fixed effects estimate</th>
<th>Standard error</th>
<th>Variance components estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>17.7659</td>
<td>0.8878</td>
<td>$\sigma_1^2$ 1.3424</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.5041</td>
<td>0.0035</td>
<td>$\phi$ 0.0250</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.4602</td>
<td>0.0053</td>
<td>$\delta_2$ 1.2617</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>31.3331</td>
<td>1.0341</td>
<td>$\delta_3$ 0.5147</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.0213</td>
<td>0.0005</td>
<td>$\sigma_{21}^2$ 330.3888</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.0957</td>
<td>0.0069</td>
<td>$\sigma_{22}^2$ 48.9116</td>
</tr>
<tr>
<td>$\theta_0$</td>
<td>0.4740</td>
<td>0.0129</td>
<td>$\sigma_{3}^2$ 4.7067</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>2.1417</td>
<td>0.1100</td>
<td>$\sigma_{b1}b_2$ 33.9414</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.0765</td>
<td>0.0015</td>
<td>$\sigma_{b1}b_3$ 37.146849</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\sigma_{b2}b_3$ 4.1118</td>
</tr>
</tbody>
</table>
Table 2. Basic statistics of knot relative vertical position ($y_1$, mm), diameter ($y_2$, mm) and tangent ($y_3$, mm/mm) data at the stem periphery.

<table>
<thead>
<tr>
<th>Response</th>
<th>mean</th>
<th>standard dev.</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1$</td>
<td>55.8873</td>
<td>27.7044</td>
<td>1</td>
<td>0.5310</td>
<td>0.2921</td>
</tr>
<tr>
<td>$y_2$</td>
<td>16.5543</td>
<td>6.2427</td>
<td>1</td>
<td>0.3225</td>
<td></td>
</tr>
<tr>
<td>$y_3$</td>
<td>0.4820</td>
<td>0.2237</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Evaluation metrics for knot volume predictions. $E$ (mm$^3$) and $E_r$ are the absolute and relative estimation bias, respectively, $RMSE$ (mm$^3$) is the root mean square error and $EF$ stands for estimation efficiency.

<table>
<thead>
<tr>
<th>Predictions used</th>
<th>$E$</th>
<th>$E_r$</th>
<th>RMSE</th>
<th>$EF$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean knot</td>
<td>887</td>
<td>0.0627</td>
<td>9754</td>
<td>0.4395</td>
</tr>
<tr>
<td>Diameter (conditional),</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative vertical position (marginal)</td>
<td>643</td>
<td>0.0455</td>
<td>2644</td>
<td>0.9588</td>
</tr>
<tr>
<td>Diameter (conditional),</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative vertical position (imputed from diameter model)</td>
<td>651</td>
<td>0.0460</td>
<td>2632</td>
<td>0.9592</td>
</tr>
<tr>
<td>Diameter (conditional),</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative vertical position (imputed from diameter and tangent models)</td>
<td>708</td>
<td>0.0500</td>
<td>2656</td>
<td>0.9584</td>
</tr>
</tbody>
</table>
FIGURE CAPTIONS

Fig. 1. The terms used to describe knot geometry. The longitudinal section of the tree is shown in the xz plane, while the cross section of the tree is shown in the xy plane. This cross section of the tree is taken at the line ab. The diameter of the knot denoted by cd is measured in the xy plane and we assume that it is equal to the diameter in the xz plane. The tangent (line ef) of the knot trajectory is located at the intersection of the knot midpoint (the centre of cd) with the cross section (line ab), this intersection of cd and ab on the xz plane is the vertical position.

Fig. 2. Left: knot display by means of Bil3D software, with whorls, inter-whorls and epicormics (the latter are the two thin flat lines at about one third of the image; eventually removed in the analysis); Right: X-ray image analysis through Gourmand software.

Fig. 3. Predictions of knot relative vertical position over stem radius and 95% percentile rank confidence intervals for a 150-mm long whorl knot.

Fig. 4. Predictions of knot diameter over stem radius and 95% percentile rank confidence intervals for a 150-mm long whorl knot.

Fig. 5. Predictions of knot tangent over stem radius and 95% percentile rank confidence intervals for a 150-mm long whorl knot.

Fig. 6. Nominal coverage probability over empirical coverage probability. The straight line represents a perfect agreement between observed and expected volume distribution.

Fig. 7. Observed vs predicted mean knot volume. The solid line represents a perfect agreement between observations and predictions. The dashed one is the regression line between them.

Fig. 8. Observed vs predicted individual knot volume. The solid line represents a perfect agreement between observations and predictions. The dashed one is the
regression line between them.
FIGURES

Figure 1

Figure 2

Figure 3
**Figure 4**

**Figure 5**
INDIVIDUAL KNOT VOLUME ESTIMATION

Figure 8