Probabilistic sensitivity analysis to understand the influence of micromechanical properties of wood on its macroscopic response

P.O. Hristov[a,*], F.A. DiazDelaO[a], E.I. Saavedra Flores[b], C.F. Guzmán[b], U. Farooq[c]

[a] *Institute for Risk and Uncertainty, School of Engineering, University of Liverpool, Liverpool L69 7ZF, United Kingdom*
[b] *Departamento de Ingeniería en Obras Civiles, Universidad de Santiago de Chile, Av. Ecuador 3659, Estación Central, Santiago, Chile*
[c] *Parker Hannifin Manufacturing (UK) Ltd*

Abstract

This paper investigates the influence of the uncertainty in different micromechanical properties on the variability of the macroscopic response of cross-laminated timber plates, by means of a probabilistic sensitivity analysis. Cross-laminated timber plates can be modelled using a multiscale finite element approach which although suitable, suffers from high computational cost. Investigating parametric importance can incur considerable time penalty since conventional sensitivity analysis relies on a large number of code evaluations to produce accurate results. In order to address this issue, we build a statistical approximation to the code output and use it to perform sensitivity analysis. We investigate the effect of a collection of parameters on the density and Young's moduli of wood. Additionally, the influence on the response of cross-laminated timber plates subject to bending, in-plane shear

*Corresponding author. Tel. +44(0)747 461 42 87
Email address: p.hristov@liv.ac.uk (P.O. Hristov)

and compression loads is investigated due to its relevance within the engineering community. The presented results provide a practical insight into the importance of each micromechanical parameter, which allows research effort to be focused on the important wood properties.

Keywords: Probabilistic sensitivity analysis, Gaussian process emulation, Cross laminated timber, Multi-scale analysis, Finite elements

1. Introduction

In recent years, considerable attention has been paid to the investigation of wood at multiple length scales. At microscopic levels, wood shows remarkable features, such as its highly organised hierarchical design, its ability to deflect microcracks which results in an increased fracture toughness, and its lightweight and excellent thermal and acoustic insulation characteristics due to its porous microstructure. At very large scales, the above properties, combined with its reduced environmental impact, make wood to be an ideal candidate for building applications. In particular, cross-laminated timber (CLT) has been increasingly spreading in Europe and North America over the last decade as a novel prefabricated building system [1]. CLT panels are composite structures made up of several layers of boards stacked crosswise and glued together on their faces, as can be seen in Figure 1.

[Figure 1 about here.]

The main advantages of CLT are its fast and efficient on-site installation, its favourable seismic performance, its ability to self-protect against fire and its excellent strength [2].

Despite the above advantages, the computational modelling of CLT, and in general timber structures, still represents a very challenging task. This can be attributed to the highly heterogeneous macroscopic properties of wood. One possible approach to tackle this problem is to predict the macroscopic response using the mechanical information coming from its microstructure. This task can be achieved by means of the finite element (FE) based multiscale modelling technique [3]. Considerable effort has been devoted to the computational modelling of timber structures [4, 5, 6, 7], but the complete understanding of the mechanical properties of this material at small spatial scales is still an open issue.

The micromechanical properties of wood can be uncertain due to the lack of knowledge or because of measurement errors at such small length scales. As these properties are crucial to develop reliable predictive models, the uncertainty in their values must be taken into account. Recently, Saavedra Flores et al. [8, 9] considered the uncertainty in the micromechanical parameters of a multiscale model for wood. This uncertainty was propagated to the macro-scale, giving rise to uncertain macroscopic properties. In this new paper, we continue the line of development started in the above references [8, 9]. By means of a probabilistic sensitivity analysis, we investigate the influence of uncertainty in different microscopic properties on the variability of the macroscopic response of wood. Probabilistic sensitivity analysis (SA) relies on a large number of expensive code evaluations to produce accurate results. In order to tackle the high computational cost associated with the analysis, we build a more affordable surrogate of the code and use it to perform the numerical studies. There are a number of methodologies for approximating the output of expensive codes, see *e.g.* [10]. One particular approach is Gaussian process emulation (GPE), which builds a statistical approximation to the output of the code. Using this technique we investigate the extent to which different micromechanical parameters influence the macroscopic response of wood. Due to its relevance within the engineering community, we also explore the influence on the response of CLT plates subject to bending, in-plane shear and compression loads. Once the relative importance of different parameters is known, the information can be used to either reduce the computational cost of the model by fixing the least influential variables, or to maximize the reduction in response uncertainty by conducting research on the important wood properties.

The rest of the paper is organized as follows: Section 2 outlines the relevant wood properties and corresponding modelling strategies. Section 3 introduces the basics of probabilistic sensitivity analysis and Gaussian process emulation. Section 4 presents the micro-macro study, in which the effect of the micro parameters on each macro parameter is measured. Section 5 provides an interpretation of the results from the previous section and finally Section 6 draws the main conclusions of the presented work.

2. Multi-scale modelling

The multiscale modelling of timber is described in this section. A computational homogenisation approach is adopted to capture the hierarchical nature of wood at different length scales. Here, four different spatial scales are considered. These are the nanometer, micrometer, millimeter and the structural scales. For further details on the present homogenisation approach, we refer, for instance, to Saavedra Flores et al.[11].

2.1. Relevant wood micromechanical parameters

At nanometer levels, wood contains three basic constituents: cellulose, hemicellulose and lignin [12]. These three fundamental constituents form the wood cell-wall composite material whose basic unit building block is called microfibril. This composite comprises reinforcing cellulose fibrils oriented mainly in a single direction (in almost the whole cell-wall volume) periodically embedded in a softer matrix.

The specific angle of the microfibrils with respect to the longitudinal axis of the wood cell is typically called microfibril angle, *MFA*. The volume fraction of cellulose f_c is defined as the volume of cellulose with respect to the total volume of the cell-wall composite. Similarly the volume fraction of hemicellulose f_h relates the volume of hemicellulose. The reinforcing cellulose is made up of periodic alternations of crystalline and amorphous fractions. The degree of cellulose crystallinity f_{cc} is defined as the volume fraction of the crystalline portion of cellulose with respect to the total volume of cellulose. As the cellulose is a long and stiff polymeric fibre, the length of the crystalline fraction is termed here L_{cc}.

The matrix of the cell-wall composite is made of hemicellulose and lignin polymers. Hemicellulose is built up of sugar units and has little strength, with mechanical properties highly sensitive to moisture changes. Lignin is an amorphous and hydrophobic polymer and its main purpose is to cement the individual wood fibres together and to provide inter-fibre shear strength.

At the micrometer scale, the material can be represented by a periodic arrangement of long slender tubular micro-fibres (or wood cells), oriented nearly parallel to the axis of the stem. The cross-sections of each micro-fibre is (normally) hexagonal, and can be defined by means of four geometric parameters. These are the tangential and radial dimensions of the hexagonal cross-section, denoted here as T and R (along the tangential and radial directions of wood), respectively, the thickness of the cell-wall, t_c, and the angle θ (whose value can be, for instance, 0^o for a rectangular cross-section, or 30^o for a regular hexagonal shape). In softwoods, wood fibres can be divided into *early-wood* and *late-wood*. The *early-wood* fibres are characterised by large diameters and thin cell-walls, whereas *late-wood* fibres are composed of narrow diameters with much thicker cell-walls. In order to avoid confusion, we use in this paper the following terminology to differentiate both types of cells. The variables T_p, R_p and t_{cp} refer to the tangential and radial dimensions and thickness of *early-wood* fibres. Similarly, T_v, R_v and t_{cv} refer to the tangential, radial and thickness dimensions of *late-wood* fibres. Given the little information reported on the distinction of θ between *early-wood* and *late-wood*, such an angle is assumed to be the same for both types of cells.

At the scale of a few millimeters wood is represented by the growth rings, typically found in the cross-section cut through the trunk of a tree. Within a growth ring, the volume fraction of *early-wood* fibres with respect to the total volume of growth ring is denoted as P_{ew}. For further information about the morphology and composition of wood at microscopic levels, we refer, for instance, to [13, 14]. The final macroscopic or structural scale is represented by the periodic repetition of the growth rings which form the base material.

Summarizing, 13 micromechanical parameters are defined. Four at the nanometer scale (MFA, f_c , f_h and f_{cc}), eight at the micrometer scale (t_c,

θ, T_p, R_p, t_{cp}, T_v, R_v and t_{cv}) and one at the millimeter scale (P_{ew}).

2.2. Macroscopic parameters

The general procedure consists of building a material model for wood by homogenising the three material scales described in the previous section (at the level of the microfibril, wood fibres and growth rings). With this model at hand, we can predict the response of any (macroscopic) timber structure (that is, the structural scale).

In this study, we choose two types of structural configurations. First, we analyse a timber plate of length of 2.4 m (parallel direction to wood fibres), width of 1.2 m (perpendicular direction to wood fibres), and thickness of 4 cm. We note that the general dimensions of 1.2 m by 2.4 m belong to a standard geometry adopted typically for the experimental testing of structural panels [15, 16, 17]. The plate is subject to four-point bending along the length and width of the panel. From these analyses we obtain the longitudinal and transverse Young's moduli for wood, E_0 and E_{90}, respectively. Second, we analyse a CLT plate. The motivation of choosing CLT for this study is because of its increasing use worldwide as a promising prefabricated construction system [11]. The CLT plate consists of three layers of boards stacked crosswise and glued together on their faces. Each layer is 4 cm of thick, with a length of 2.4 m and a width of 1.2 m (that is, the first configuration described above). Thus, the total thickness of the CLT plate is 12 cm. The outer layers are made of timber members oriented in the long direction of the panel (that is, the strong direction). The central layer is made of members oriented in the short (or weak) direction. The CLT plate is subject to three-point bending along the strong direction, in-plane shear loading and compression parallel to wood fibres in the outer layers. From these analyses, we obtain the bending stiffness K_{bend}, the in-plane shear stiffness K_{sh} and the axial stiffness K_{comp} of the CLT plate. For further details on these stiffness components, we refer to [11]. In addition, we compute the macroscopic density of the material ρ. The above six macroscopic parameters (ρ, E_0, E_{90}, K_{bend}, K_{comp} and K_{sh}) are selected because of their relevance in the day-to-day practice of analysis and design of timber structures, particularly in the context of CLT structures.

2.3. Modelling of macro and micro-scales

Multi-scale models enable specifying the relationships between physical variables observed at different length scales. These are of particular importance in the study of heterogeneous materials with hierarchical microstructures in which the macroscopic response of the material can be predicted from the information coming from the microscopic (or lower) level.

In the present multiscale constitutive theory, each material scale is associated with a microstructure whose most statistically relevant features are incorporated within a representative volume element (RVE). This RVE is assumed to have a (microscopic) characteristic length much smaller than the macro-continuum, and at the same time, a size large enough to capture the microscopic heterogeneities in an averaged sense. This multiscale methodology has proven to be successful to reproduce the mechanical behaviour of materials at several length scales. As described at the beginning of Section 2, four spatial scales can be identified. Three of them represent material scales, and a fourth is associated with the structural scale.

Depending on the kinematic constraints imposed in the RVE, several classes of multiscale models can be defined. Here, we choose the periodic boundary displacement fluctuations multiscale model [18], which is typically used to model periodic media, like wood micro structures and several other natural materials. The type of wood species chosen for this investigation is radiata pine grown in Chile, which has several applications in building and engineering structures.

Each spatial scale was modelled using the FE model with meshes depicted in Figure 2. Note that all the FE meshes used in our computational simulations were obtained after a preliminary convergence study. The results (omitted here for brevity) did not indicate a significant change of the simulation outcomes for increasing mesh densities. Additionally, the same mesh was used in previous works [2, 19] We also note that the first two material scales (Figure 2(a), Figure 2(b), Figure 2(c)) have already been described in [11] and therefore, we skip the details about their modelling. A typical finite element mesh of the RVE chosen to describe the mechanical response of the growth ring is shown in Figure 2(d). It consists of 288 nodes and 165 hexahedral elements. The turquoise colour represents the portion of material calculated by the computational homogenisation of the *early-wood* RVE shown in Figure 2(c), whereas the light brown colour shows the material obtained by the homogenisation of the *late-wood* RVE shown in Figure 2(b). The periodic repetition of the growth rings forms the base material for the macroscopic or structural scale (in this case, the CLT panels). This scale is modelled using the finite element mesh depicted in Figure 2(e) (for the four-point bending) and 2(f) (for the three-point bending). The discretisation is the same in both figures, with 379 093 nodes and 345 600 SOLID45 elements. The computational homogenisation procedure described in this section is implemented in the commercial software ANSYS [20].

[Figure 2 about here.]

3. Probabilistic sensitivity analysis

Being a representation of a complex natural process, it is expected that the FE code will also be complex. Complexity of computer codes is mainly associated with their high computational cost and the lack of an analytical expression of the input/output mapping i.e., the model is treated as a *black box*. The multiscale FE model used in this paper is a deterministic simulator. This means that the output is precisely the same, every time its multidimensional input is given the same value. Despite this, micromechanical parameters are physical quantities and as such come from some joint probability distribution, $f_X(\mathbf{x})$ effectively making them an uncertain multivariate random variable, $\mathbf{X}$. This uncertainty is propagated to the output through the simulator. Thus, the output $y = \eta(\mathbf{x})$ can be seen as a random variable $Y = \eta(\mathbf{X})$, with its own probability distribution. Very often the different inputs do not influence the output equally. Part of the computer-based investigation of the physical process is determining the influence of the uncertainty in the different inputs, or sets of inputs, on the variability of the output. This process is known as probabilistic *sensitivity analysis* (SA).

There are two main types of sensitivity analyses, namely local (LSA) and global (GSA). LSA is concerned with determining the effect of small, local perturbations in the parameter value around a given base point. A very common local sensitivity approach is the one based on derivatives, namely $\partial Y/\partial X_i$, where X_i is the i^{th} component of $\mathbf{X}$. The local method is not used in the current investigation, since it fails to capture the effect of the inputs when their values are arbitrarily chosen from the entire input domain. For more detail on LSA see [21].

In order to fully explore the input space, GSA relies on a number of points carefully spread according to some experimental design. There exists a variety of GSA techniques such as function decomposition in main and higher order effects, regression coefficients and variance-based methods, among others (see for example [21]).

3.1. Variance-based sensitivity analysis

Since the models of interest are deterministic, the variance of the output random variable will be entirely due to the uncertainty in the input values. This means that if one could learn the exact, true values of the inputs, the variance of Y would be reduced to 0. This leads to the notion that fixing one of the inputs X_i at a given value x_i and re-running the code will result in Y having a lower variance. Let $\mathbb{V}_{\sim i}[Y|X_i]$ be the conditional variance of Y, taken over all factors, but X_i (denoted $X_{\sim i}$) and given $X_i = x_i$. This conditional variance can be used as a measure of how influential the fixed parameter is. A severe drawback of this measure, however, is its dependence on the location of the point x_i. This problem could be resolved by taking the average of the conditional variance over all possible values of x_i, that is $\mathbb{E}_i[\mathbb{V}_{\sim i}[Y|X_i]]$. It is a well known fact in Probability theory that the variance of a random variable can be decomposed as:

$$\mathbb{V}[Y] = \mathbb{E}_i[\mathbb{V}_{\sim i}[Y|X_i]] + \mathbb{V}_i[\mathbb{E}_{\sim i}[Y|X_i]] \tag{1}$$

Eq. (1) gives another important diagnostic - $\mathbb{V}_i[\mathbb{E}_{\sim i}[Y|X_i]]$ - *the first order effect* of X_i on Y. The associated normalised sensitivity measure, also known as a *Sobol' index* [22] is:

$$S_i = \frac{\mathbb{V}_i[\mathbb{E}_{\sim i}[Y|X_i]]}{\mathbb{V}(Y)} \tag{2}$$

A high value of the Sobol' index for the given variable, means that it is important, i.e. if it is fixed, there will be a significant reduction in the variance of Y.

Many practical models are of the so-called *non-additive* class. That is, the effect that individual inputs have on the variance cannot be separated to account for 100% of the output variance. Instead, interactions between individual inputs or sets of inputs will play an important role. To capture such effects, the *higher-order Sobol' indices* can be constructed:

$$S_p = \frac{\mathbb{V}_p[\mathbb{E}_{\sim p}[Y|\mathbf{X}_p]]}{\mathbb{V}(Y)} \tag{3}$$

where $p \subset \{1, \ldots, d\}$ is a set of indices of all inputs under investigation. It can be shown that:

$$\sum_{i=1}^{d} S_i + \sum_{i<j}^{d} S_{i,j} + \sum_{i<j<k}^{d} S_{i,j,k} + \ldots + S_{1,2\ldots,d} = 1 \tag{4}$$

where the summation is carried out over all d dimensions, which means that summing over all Sobol' indices recovers the full variance. Full analysis of the main effects of the model inputs and their respective interactions will result in Eq. (4) having $2^d - 1$ terms. This means that with relatively low number of inputs, the summation components become too many to investigate individually. Homma and Saltelli [23] introduced the *Total Sobol' index*:

$$S_{T_i} = \left(1 - \frac{\mathbb{V}_{\mathbf{X}_{\sim i}}[\mathbb{E}_{X_i}[Y|\mathbf{X}_{\sim i}]]}{\mathbb{V}(Y)}\right) \tag{5}$$

This measure captures the effect of the $i-th$ input and all of its interactions, by fixing all other inputs. It is true that $S_i \leq S_{Ti}$, due to interactions between inputs. Equality can only arise in a perfectly additive model. An input is said to be truly non-influential if and only if $S_{Ti} = 0$. In [24] the authors have argued that a good, albeit non-exhaustive characterization of the input influences is given by the set of first order and total Sobol' indices. This is further discussed with the results presented in Section 4.

3.2. Gaussian process emulation

Simulators used to model complex scientific phenomena are usually very computationally expensive. This is to say that a single evaluation of the code's output at a given set of input values takes sufficiently long time, as to prohibit any type of analysis which requires a large number of model runs. The multiscale FE code used in this work is no exception. Since sensitivity analysis relies on Monte Carlo (MC) approximations of integrals, the estimators of the Sobol' indices will converge to their true value as the number of points used to estimate them approaches infinity. Clearly, the analysis cannot be carried out using the code directly. In such cases it is common to use a less expensive approximation of the code output. These approximations are widely known as *metamodels* or *emulators*. There is a number of existing metamodelling techniques, but for the purposes of this study, Gaussian process emulators (GPE) are used. Formally, the model structure is expressed as:

$$\eta(\mathbf{x}) = h(\mathbf{x})^T\boldsymbol{\beta} + Z(\mathbf{x}) \tag{6}$$

where $\eta(\mathbf{x})$ is the simulator output as a function of its inputs, $h(\mathbf{x})^T$ is a known function of the inputs, $\boldsymbol{\beta}$ is a vector of unknown coefficients and $Z(\mathbf{x})$ is a Gaussian process with zero mean, and covariance, $\sigma^2 c(\mathbf{x}, \mathbf{x}'; \psi)$. The function $h(\mathbf{x})$ should express any expert opinion about the form of the simulator output and together with the parameter $\boldsymbol{\beta}$ reflects its overall trend. In practice however, the trend is often taken to be constant as $h(\mathbf{x}) = 1$, charging the Gaussian process in Eq. (6) with the responsibility of capturing the behaviour of the underlying function. In the formulation above, σ^2 is a scale parameter, $c(\mathbf{x}, \mathbf{x}'; \psi)$ is a known correlation function and ψ is a parameter specifying the behaviour of the correlation function. The parameters of the Gaussian process are also commonly referred to as *hyperparameters* [25] to distinguish them from the model parameters.

Using the GPE, a *posterior* probability distribution for the mean of the computer code's output can be constructed, conditional on a relatively small number of simulator runs with outputs $\mathbf{y}$ and the estimated parameter values, $\hat{\boldsymbol{\theta}} = \{\hat{\beta}, \hat{\sigma}^2, \hat{\psi}\}$. It can be shown [26] that at a new unobserved set of input values, $\mathbf{x}^*$, the posterior distribution has the form of a multivariate Gaussian distribution:

$$\eta(\mathbf{x}^*)|\mathbf{y}, \hat{\boldsymbol{\theta}}, \sim \mathcal{N}(m(\mathbf{x}^*), C(\mathbf{x}^*, \mathbf{x}'^*)) \tag{7}$$

with posterior predictive mean function:

$$m(\mathbf{x}^*) = \hat{\beta} + \mathbf{t}(\mathbf{x}^*)^T \boldsymbol{C}^{-1}(\mathbf{y} - \mathbf{1}\hat{\beta}) \tag{8}$$

and posterior predictive covariance function:

$$C(\mathbf{x}^*, \mathbf{x}'^*) = \hat{\sigma}^2(c(\mathbf{x}^*, \mathbf{x}^*) - \boldsymbol{t}(\mathbf{x}^*)^T \boldsymbol{C}^{-1} \boldsymbol{t}(\mathbf{x}'^*)) \tag{9}$$

In Eqs. (8) and (9) $\mathbf{C} \in \mathbb{R}^{n \times n}$ such that $C_{ij} = c(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{t}(\mathbf{x}^*) \in \mathbb{R}^n$ such that $\mathbf{t}(\mathbf{x}^*) = (c(\mathbf{x}^*, \mathbf{x}_1), \ldots, c(\mathbf{x}^*, \mathbf{x}_n))^T$ and $\mathbf{1} \in \mathbb{R}^n$ such that $\mathbf{1} = (1, \ldots, 1)^T$. The process of estimating $\boldsymbol{\theta}$ (i.e. constructing $\hat{\boldsymbol{\theta}}$) from observed data is referred to as *training* and is very well described in [10] from a classical prospective or in [26, 27] from a Bayesian standpoint. Once the emulator is trained, its posterior distribution can be sampled many times at an affordable cost to provide data for various analyses.

4. Micro-Macro analysis

4.1. Gaussian process emulator validation

The micro-macro analysis deals with the investigation of relations between the 13 microscopic properties and the 6 macromechanical parameters described in Section 2. The 6 macro parameters are analysed independently by fitting one Gaussian process per parameter. Therefore, the black-box function is of the form $M_j = \eta(m_1, \ldots, m_{13})$, where the M_j is the j^{th} macro parameter and $m_1, \ldots, m_{13}$ are the micro parameters.

[Table 1 about here.]

Since the multiscale model is expensive, it should only be run as many times as necessary. When performing computer experiments it is common to apply the $10d$ rule [22] for selecting the size of the training sample for the GPE.

Therefore, 130 uniformly distributed points were selected via a Latin hypercube sampling (LHS). LHS was chosen because it best represents each individual dimension. Another 60 LHS points were chosen as a validation set to check the quality of the GPE. The GPE was coded in MATLAB® and the model was run 190 times. The material properties of the model and the lower and upper bounds are retrieved from Saavedra Flores et al. [8, 11]. Table 1 gives the ranges of the each micromechanical parameter. The values were constrained in order to match physically possible values and the available experimental data. The properties are assumed to be stochastically distributed as uniform random variables because they are susceptible to considerable variations when measured experimentally [11]. A genetic algorithm was used to perform a direct search for the optimal hyperparameter values and the mean and variance were calculated via maximum likelihood estimation (MLE) [10]. There are a variety of validation techniques, which could be used for identifying problems with the emulator (see for example [28]). Here we have used individual prediction errors which are represented by the normalised difference between the real and predicted values of each test point:

$$D_i^I = \frac{y_i - \mathbb{E}[\eta(\mathbf{x}_i^*)|\mathbf{y}]}{\sqrt{\mathbb{V}[\eta(\mathbf{x}_i^*)|\mathbf{y}]}} \tag{10}$$

where the expected value of the posterior distribution, $\mathbb{E}[\eta(\mathbf{x}_i^*)|\mathbf{y}]$ and its variance $\mathbb{V}[\eta(\mathbf{x}_i^*)|\mathbf{y}]$ are given in functional form in Eq. (8) and Eq. (9), respectively. If the emulator can accurately represent the simulator, these errors should have a standard *Student-t* distribution. With a large number of degrees of freedom the *Student-t* approaches a standard normal distribution and thus, any errors with absolute value greater than 2 (i.e. outside of the

95 % credible interval) can be considered local conflicts between emulator and simulator. Patterns of errors lying outside the $[-2, 2]$ region could indicate more serious problems. A useful visual validation tool is the plot of predictions at the test points versus their true values. Figure 3 shows the validation results for all 6 macro parameters. It can be seen that there is a close correspondence between predictions and observations. Each point also displays the 95% credible interval, which is based on the posterior predictive variance. The individual prediction errors plotted against prediction values are shown in Figure 4. All but a few of the errors lie within the desired boundaries, which together with the plots in Figure 3 suggest that the emulator is a valid representation of the simulator. Once the GPE was validated the 60 points used for the process were added to the training sample and the surface was refit based on all 190 points.

[Figure 3 about here.]

[Figure 4 about here.]

4.2. Calculation of Sobol' indices

As mentioned in Section 3.1, the calculation of Sobol' indices requires the evaluation of both conditional and unconditional expectations and variances. These operations are associated with the calculation of a number of integrals. In order to evaluate them, the integrals can be approximated by Monte Carlo (MC) simulation. Since this is the case, a relatively large sample size is required to achieve reasonably accurate estimation results. This is often a problem because, coupled with the computational complexity of most scientific and engineering codes, extensive sampling results in a very costly sensitivity analyses. Using GPEs as inexpensive approximations to the output of the code, together with the use of parallel computers, enables MC based analyses to be performed within reasonable time periods. The unconditional variance of the simulator output can be written as:

$$\mathbb{V}[Y] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \tag{11}$$

When using the emulator, the simulator output Y in Eq. (11) is substituted with the posterior mean of the emulator $\mathbb{E}[\eta(\mathbf{X})|\mathbf{y}]$ (Eq. (8)). Then, the Monte Carlo approximations of the terms in Eq. (11) are given by:

$$\hat{\mathbb{E}}[Y] = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[\eta(\mathbf{x}^{(n)})|\mathbf{y}] \tag{12}$$

$$\hat{\mathbb{V}}[Y] = \frac{1}{N-1}\sum_{n=1}^{N}\mathbb{E}[\eta(\mathbf{x}^{(n)})|\mathbf{y}]^2 - \hat{\mathbb{E}}[Y]^2 \tag{13}$$

Here we only give the estimator for the first-order and the *total* conditional variances, since any higher-order variances could be calculated from their definitions in Eq. (2) and in Eq. (5).

$$\hat{\mathbb{V}}[\mathbb{E}[Y|X_i]] = \frac{1}{N-1}\sum_{n=1}^{N}\mathbb{E}[\eta(x_i^{(n)}, \mathbf{x}_{\sim\mathbf{i}}{}^{(n)})|\mathbf{y}]\mathbb{E}[\eta(x_i^{(n)}, \mathbf{x}'_{\sim\mathbf{i}}{}^{(n)})|\mathbf{y}] - \hat{\mathbb{E}}[Y]^2 \tag{14}$$

$$\hat{\mathbb{V}}[\mathbb{E}[Y|\mathbf{X}_{\sim i}]] = \frac{1}{N-1}\sum_{n=1}^{N}\mathbb{E}[\eta(x_i^{(n)}, \mathbf{x}'_{\sim\mathbf{i}}{}^{(n)})|\mathbf{y}]\mathbb{E}[\eta(x_i'^{(n)}, \mathbf{x}'_{\sim\mathbf{i}}{}^{(n)})|\mathbf{y}] - \hat{\mathbb{E}}[Y]^2 \tag{15}$$

where $\mathbf{x}$ and $\mathbf{x}'$ come from two distinct sets of values for $\mathbf{X}$ each of size $N \times d$. The full algorithm for the calculation of the first order and total Sobol' indices is given in [29]. In some instances, analytical expressions from the GPE are available for all quantities of interest, but these rely on some modelling assumptions and hence we resort to sampling the posterior of the GPE directly. A straightforward convergence study was carried out to determine a suitable sample size. The procedure was based on obtaining 100 Sobol' index estimates based on 20 different sample sizes between $N = 1000$ and $N = 20\,000$. In Figure 5 we plot the mean and one standard deviation of the first order indices for density and select the appropriate sample size as the one after which there is no appreciable change in the index' standard deviation. In this case the size was selected to be $N = 10\,000$ points per variable. Inspecting Figure 6, which shows the same study based on the total indices, confirms the correctness of the choice. A sample of $N = 10\,000$ points per variable results in a total of $M = N \times (d + 2) = 150\,000$ points for estimating first and total Sobol' indices for all 13 variables. It is immediately obvious that such a sample could have not come directly from the model at a reasonable computational cost. Figure 7 shows a set of bar graphs representing the first order and total Sobol' indices of the 13 micromechanical parameters for density, longitudinal and transverse Young's modulus, in rows (a) - (c), respectively. Figure 8 follows the same logic and depicts the Sobol' index ranking for bending, compression and shear stiffness in rows (a) - (c), respectively.

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

The indices represent a ratio of variances so they can not (in theory) be negative. However some estimated values that are close to 0 are negative, due to the fact that all integrals are estimated using sums (see Eq. (14)). It is useful therefore to have a measure of confidence in the estimations.

Traditionally the use of *bootstrap* [30] has been employed when the data generating process is expensive and limits the size of available observations; see for instance [31]. When using the emulator however, this is not the case and predictions for any given input combination are readily available. The fact that the GPE is only an approximation to the output of the real code can be accounted for by sampling the whole posterior distribution, instead of just its mean. The error bars on Figures 7 and 8 depict ± 2 sample standard deviations obtained from sampling the emulator 1000 times. These measures give a 95% confidence interval for the indices and reflect the validity in the predictions from the GPE shown in Figures 3 and 4

5. Discussion

Probabilistic sensitivity analysis used in conjunction with GPE provides an affordable way of constructing Sobol' indices. Using this framework any number of indices or combinations thereof can be easily computed. For practical (visualisation) reasons we only compute the first and total Sobol' indices. We remind the reader that these two indicators measure the influence of the micromechanical properties on the uncertain macroscopic response. Here, a micromechanical parameter is considered to be non-influential (or with little impact) on the macroscopic response if both Sobol' indices are zero (or nearly zero). In general, the same trends in the first-order and total Sobol' indices are observed in both Figures 7 and 8, which represents weak interaction among parameters. Interaction among input variables is indicated as the relative increase in the total Sobol' indices as compared to the first order terms. It is noted that this increase quantifies that part of the response variability which cannot be written off as a simple superposition of input effects. Weak interactions is not to say that their relative magnitude with respect to the corresponding first order effect is small, but rather that there are no major changes in the ordering of the inputs by importance. The insets in each figure show a magnification of those indices which can change order in the overall importance ranking due to their quantified uncertainty. Most of the affected parameters have relatively low Sobol' indices and are thus simply a demonstrator of the fact that sensitivity ranking is a probabilistic measure and should not be taken to have a fixed numerical value. For example the inset in Figure 8(b) shows that the importance of the thickness of late wood fibres can dominate that early wood fibres for compression stiffness. In general, parameters that were identified as important kept their positions after the inclusion of uncertainty. For the sake of clarity only two first order Sobol' indices are show in the inset. On the other hand all total indices whose error bars could not be clearly distinguished are shown in the insets in the left column of Figures 7 and 8. It is worth mentioning that the results presented in Section 4 depend on the assumed parametric distribution (see Table 1). This however, does not hold for the general methodology, which is independent of the modelling assumptions and can be applied to a wide variety of problems.

A great influence of the cellulose content on the CLT stiffness parameters K_{bend}, K_{comp} and K_{sh} is observed in Figures 8. The greatest influence is produced by the cellulose volume fraction f_c on the in-plane shear CLT stiffness K_{sh}, with a First order Sobol' index S_i^1 close to 0.36, and a total Sobol' index S_i^T around 0.38. Furthermore, its influence on the CLT stiffness is lower for the bending and compression deformation mechanisms (i.e., K_{bend} and K_{comp}, respectively), with S_i^1 and S_i^T just about 0.3 in both cases. This can be attributed to the fact that during the in-plane shear deformation process, the three CLT layers contribute greatly to the overall shear stiffness of the CLT plate. Nevertheless, for the bending and compression deformation modes, only the two external CLT layers (whose wood fibres are aligned with the loading direction) contribute significantly to the overall stiffness. The central CLT layer provides little stiffness because the wood fibres are perpendicular to the loading direction.

We note that the above strong influence of the cellulose content on the overall stiffness was expected. Nevertheless, neither the difference between the influence of the cellulose content on the shear deformation mode and on the bending and compression behaviour, nor its numerical quantification, was evident. This represents the main justification of carrying out the present sensitivity analysis.

Other influencing parameters on the CLT stiffness components are the *late-wood* and *early-wood* cell-wall thicknesses, t_{cv} and t_{cp}, respectively, with first-order and total Sobol' indices around 0.2 for the compression and shear stiffness, and t_{cp} over 0.25 for bending.

The influence of the microfibril angle, *MFA*, on the CLT stiffness is lower than that produced by the cell-wall thickness parameters. Here, the total Sobol' indices reach a maximum value of 0.07 for shear, 0.14 for compression and 0.12 for bending, and first order indices of 0.03, 0.12 and 0.11 , respectively. The remaining micromechanical parameters influence very little on the CLT stiffness components.

As expected, the density ρ is strongly affected by the *late-wood* and *early-wood* cell-wall thicknesses, and by the angle θ. Their corresponding total indices exceed 0.37, 0.31 and 0.24, respectively. In particular, the strong influence of *late-wood* is due to their thicker cell-walls when compared with *early-wood* fibres. The influence of the remaining micromechanical parameters on wood density can be neglected.

The influence on the longitudinal Young's modulus E_0 is mainly governed by the degree of cellulose crystallinity f_{cc}, the cellulose volume fraction f_c, and by the angle *MFA*. Their corresponding total Sobol' indices are 0.29, 0.27 and 0.23, respectively. The *late-wood* and *early-wood* thicknesses, and the angle θ also influence the longitudinal Young's modulus, but their indices are lower. The first two have a total index of 0.11 with 0.08 for the third.

Contrary to E_0, the transverse Young's modulus E_{90} is greatly influenced by the *late-wood* cell-wall thickness, with first-order and total Sobol' indices nearly 0.6. This behaviour can be attributed to the fact that E_{90} is mainly governed by the cell-wall matrix's response. The little influence of the *MFA* on E_{90} also contrasts with the great influence of *MFA* on E_0. Here, both Sobol' indices approach zero. Nevertheless, a greater influence on E_{90} could eventually be found if the values of *MFA* were greater. However, the emulator was trained with small values of $MFA \in [0°\text{-}22°]$.

The angle θ is another influencing parameter on E_{90}. Its corresponding Sobol' indices are close to 0.2. This behaviour is explained by the fact that the angle θ determines the transverse shape of wood fibres. Therefore, it also affects the behaviour of E_{90}.

In general, the tangential and radial dimensions of wood fibres, T_p, T_v, R_p and R_v, have virtually no impact on the macroscopic response. Similarly, the volume fraction of hemicellulose f_h, the length of the crystalline cellulose fraction L_{cc} and the volume fraction of *early-wood* fibres with respect to the total volume of growth ring P_{ew}, have also very little impact on the macroscale. The relevance of identifying these non-influencing parameters is that they can be removed from the modelling process in order to develop simpler and much more efficient models.

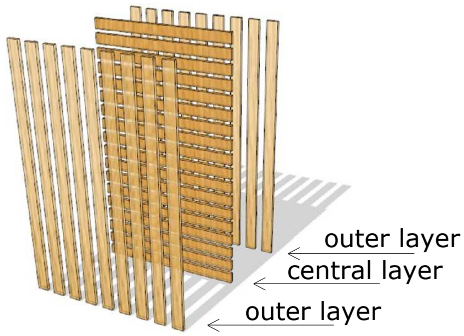
6. Conclusion

The influence of micromechanical properties of wood on its uncertain macroscopic response was investigated by means of a probabilistic sensitivity analysis. A homogenisation-based multiscale approach was adopted to capture the micro-macro relations existing in wood. Due to the relevance within the engineering community, the influence on the structural response of sawn wood and CLT plates was studied. The most influential microscopic parameter on the CLT stiffness components was found to be the cellulose content f_c. The degree of cellulose crystallinity, and the early and late wood thicknesses also played an important role on the CLT stiffness. The hemicellulose volume fraction, the tangential and radial dimensions of the wood fibre and the length of the crystalline cellulose showed very little influence on the macroscopic response. The volume fraction of early wood fibres with respect to the total volume of growth rings also showed little effect on the macroscopic stiffness. A practical insight into the definition of the micromechanical parameters allows to have an idea of the relevance of each parameters in the determination of the macroscopic response after the homogenisation procedure. Thanks to the sensitivity results presented in this work, the relevance of the parameters can not only be verified, but also quantitatively measured. These results are of practical interest, as they provide a simple criterion to weight the micromechanical parameters for future optimization of the macroscopic responses.

Acknowledgments

E.I. Saavedra Flores acknowledges the financial support from the Chilean National Commission for Scientific and Technological Research (CONICYT), FONDECYT REGULAR research project No 1160691.
P.O. Hristov acknowledges the funding provided by Parker-Hannifin Manufacturing (UK) Ltd.

List of Figures



(a) Stacked layers.

(b) Glued layers.

Figure 1: Schematic representation of a CLT panel [32].

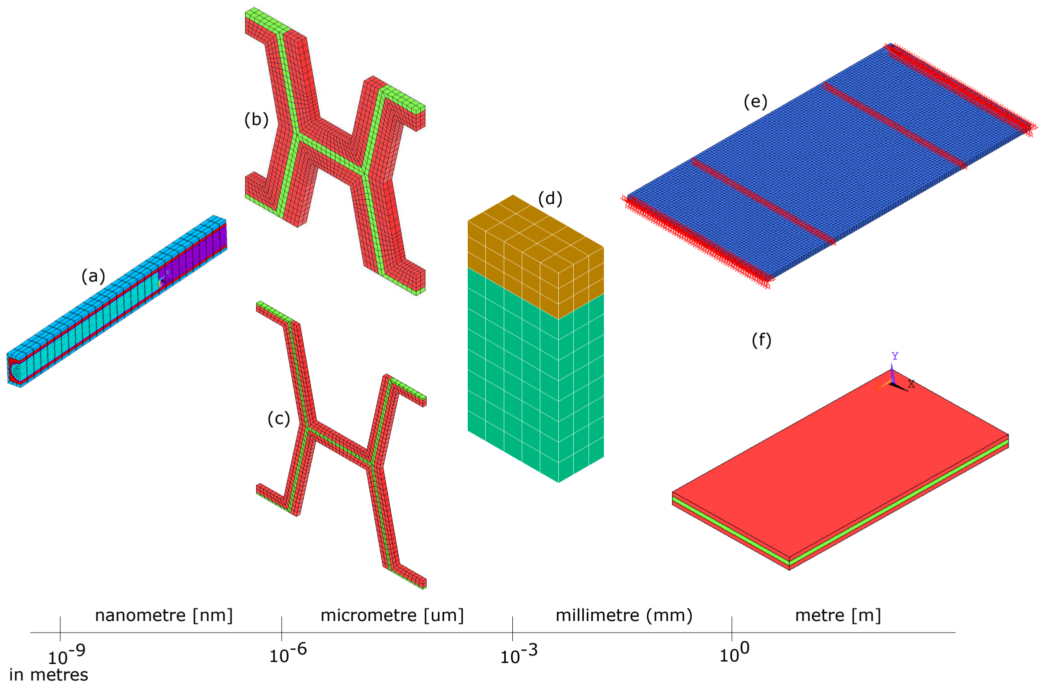


Figure 2: Finite element meshes of the RVEs and structures analysed in this paper along with their corresponding length scales. (a) RVE associated with the *microfibril* scale (for the sake of clarity, only one half of the RVE is shown here); (b) RVE associated with the modelling of *late-wood* fibres; (c) RVE associated with *early-wood* fibres; (d) growth ring RVE; (e) 4-cm-thick layer subject to four-point bending; (f) 12-cm-thick CLT panel (consisting of three 4-cm-thick layers). For the sake of clarity, the FE mesh has been hidden. Figure adapted from [32].

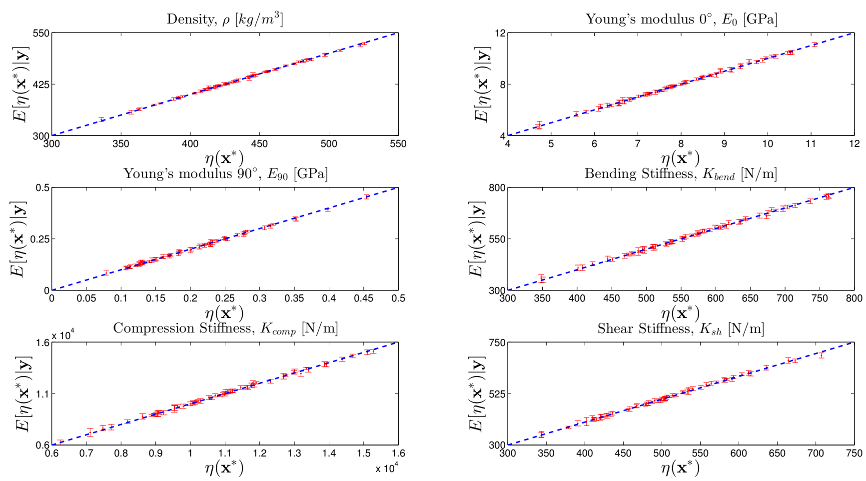


Figure 3: Simulated versus emulated values. The prediction is given by the posterior mean and the 95% credible interval (error bars) is given by the posterior variance of the emulator.

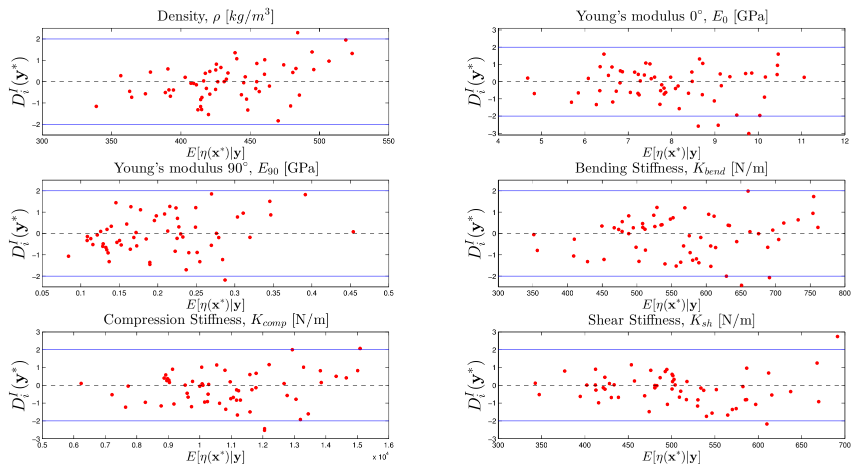


Figure 4: Individual prediction errors for all six macro parameters. The emulator value has a *Student-t* posterior distribution and the errors should lie within the interval $[-2, 2]$ with 95% confidence.

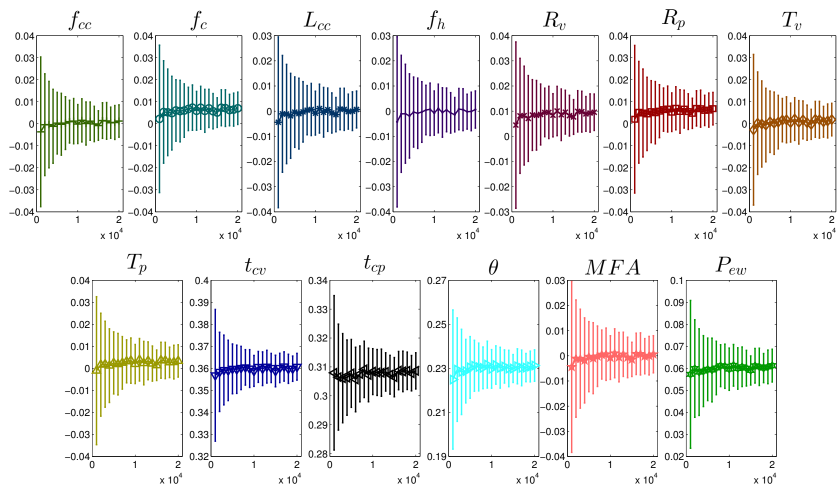


Figure 5: Convergence study for the first order parameter effects on density. Line and error bars are mean and ±1 standard deviation of the indices distribution estimated by sampling the emulator mean.

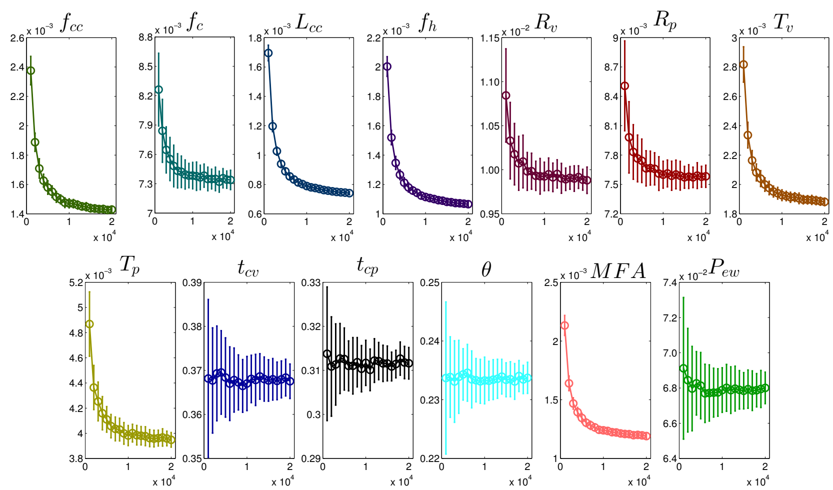
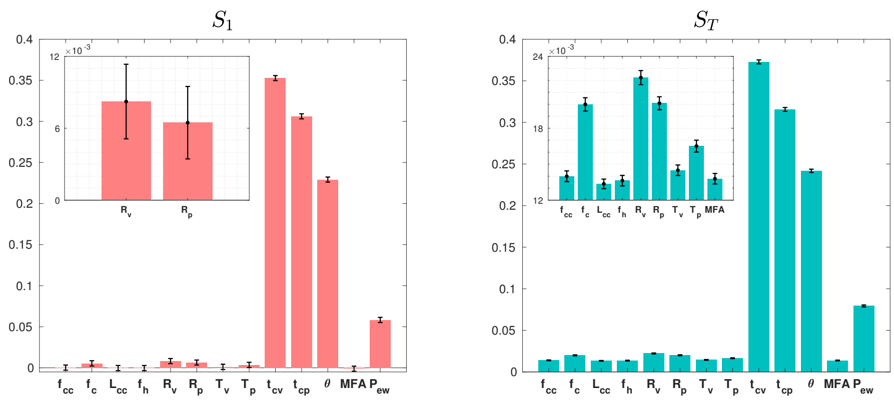
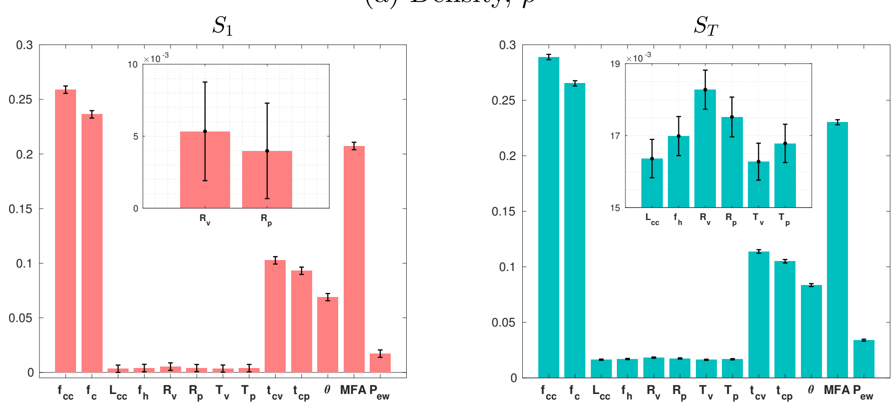


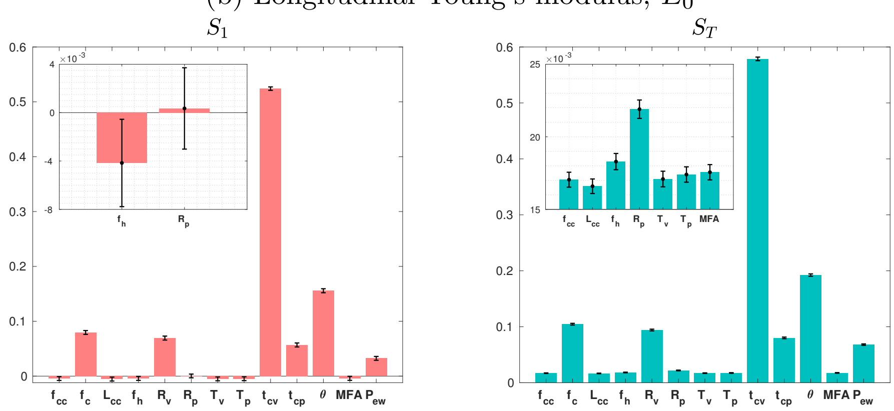
Figure 6: Convergence study for the total parameter effects on density. Line and error bars are mean and ±1 standard deviation of the indices distribution estimated by sampling the emulator mean. All non-influential values are overestimated at small sample sizes.



(a) Density, ρ

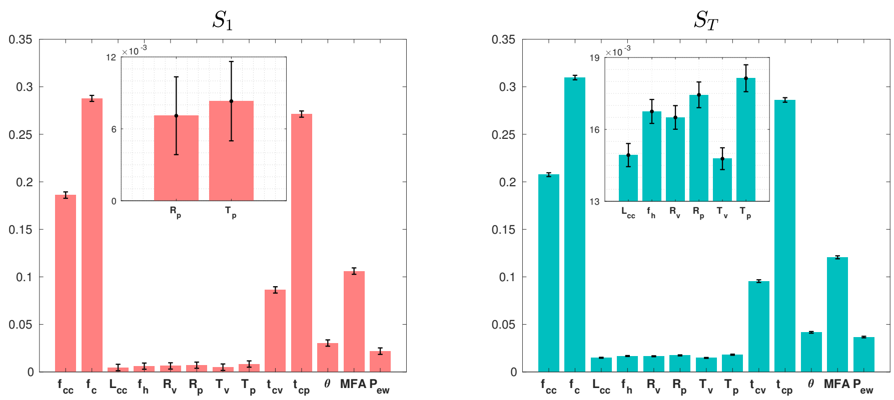


(b) Longitudinal Young's modulus, E_0

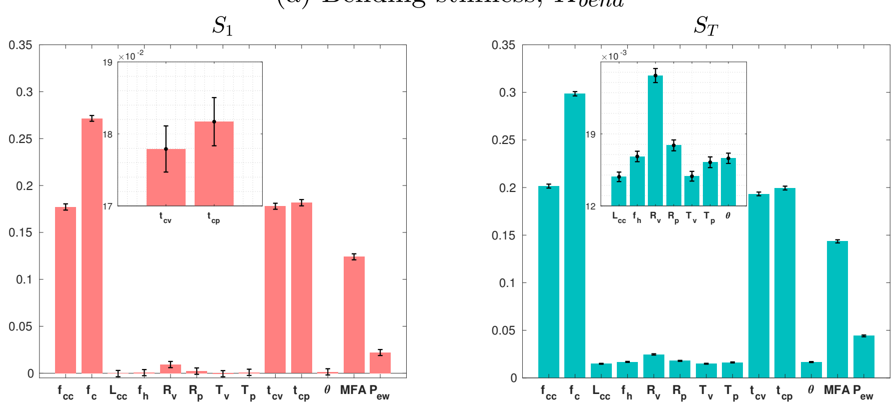


(c) Transverse Young's modulus, E_{90}

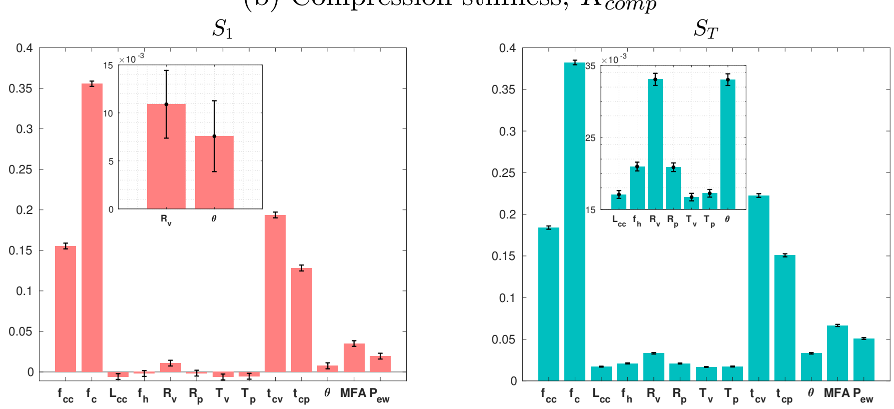
Figure 7: First and total Sobol' indices for density, longitudinal and transverse Young's moduli, (a) to (c), respectively. Error bars show ± 2 standard deviations of the indices obtained from the Gaussian process posterior. The insets show magnification of some sets of indices which could change importance due to errors. Refer to the text for more detail.



(a) Bending stiffness, K_{bend}



(b) Compression stiffness, K_{comp}



(c) Shear stiffness, K_{sh}

Figure 8: First and total Sobol' indices for bending, compression and shear stiffness, (a) to (c), respectively. Error bars show ± 2 standard deviations of the indices obtained from the Gaussian process posterior. The insets show magnification of some sets of indices which could change importance due to errors. Refer to the text for more detail.

List of Tables

Table 1: Input parameter distribution - $U(a, b)$.

	f_{cc}	f_c	L_{cc}	f_h	R_v	R_p	T_v	T_p	t_{cv}	t_{cp}	θ	MFA	P_{ew}
	%	%	nm	%	µm	µm	µm	µm	µm	µm	deg	deg	%
a	0.45	0.30	26.50	0.25	31.00	37.00	25.00	28.00	4.30	3.10	10.00	0.00	0.67
b	0.60	0.50	36.40	0.29	37.00	40.00	27.00	30.00	8.00	4.30	27.50	22.00	0.80

[1] W. Munoz, M. Mohammad, S. Gagnon, Lateral and withdrawal resistance of typical CLT connections, in: A. Ceccotti (Ed.), Proceedings of the 11th World Conference on Timber Engineering (WCTE), Trentino, Italy, 2010.

[2] E. I. Saavedra Flores, K. Saavedra, J. Hinojosa, Y. Chandra, R. Das, Multi-scale modelling of rolling shear failure in cross-laminated timber structures by homogenisation and cohesive zone models, International Journal of Solids and Structures 81 (2016) 219 – 232.

[3] J. C. Michel, H. Moulinec, P. Suquet, Effective properties of composite materials with periodic microstructure: a computational approach, Computer Methods in Applied Mechanics and Engineering 172 (1999) 109–143.

[4] M. Oudjene, E.-M. Meghlat, H. Ait-Aider, J.-L. Batoz, Non-linear finite element modelling of the structural behaviour of screwed timber-to-concrete composite connections, Composite Structures 102 (2013) 20 – 28.

[5] V.-D. Tran, M. Oudjene, P.-J. Mausoone, Experimental and numerical analyses of the structural response of adhesively reconstituted beech timber beams, Composite Structures 119 (2015) 206 – 217.

[6] M. Osei-Antwi, J. de Castro, A. P. Vassilopoulos, T. Keller, Modeling of axial and shear stresses in multilayer sandwich beams with stiff core layers, Composite Structures 116 (2014) 453 – 460.

[7] C. OLoinsigh, M. Oudjene, E. Shotton, A. Pizzi, P. Fanning, Mechanical behaviour and 3d stress analysis of multi-layered wooden beams made with welded-through wood dowels, Composite Structures 94 (2) (2012) 313 – 321.

[8] E. I. Saavedra Flores, F. A. DiazDelaO, M. I. Friswell, R. M. Ajaj, Investigation on the extensibility of the wood cell-wall composite by an approach based on homogenisation and uncertainty analysis, Composite Structures 108 (2014) 212–222.

[9] E. I. Saavedra Flores, F. A. DiazDelaO, R. M. Ajaj, M. I. Friswell, G. F. Fernando, Mathematical modelling of the stochastic mechanical properties of wood and its extensibility at small scales, Applied Mathematical Modelling 38 (15–16) (2014) 3958–3967.

[10] A. Forrester, A. Sobester, A. Keane, Engineering design via surrogate modelling: A practical guide, Wiley, 2008.

[11] E. I. Saavedra Flores, I. Dayyani, R. Ajaj, R. Castro-Triguero, F. DiazDelaO, R. Das, P. González Soto, Analysis of cross-laminated timber by computational homogenisation and experimental validation, Composite Structures 121 (2015) 386 – 394.

[12] P. Fratzl, I. Burgert, J. Keckes, Mechanical model for the deformation of the wood cell wall, Zeitschrift für Metallkunde / Materials Research and Advanced 7 (95) (2004) 579–584.

[13] J. Dinwoodie, Timber – Its nature and behavior, Von Nostrand Reinhold, New York, 1981.

[14] J. Bodig, B. Jayne, Mechanics of wood and wood composites, Von Nostrand Reinhold, New York, 1982.

[15] INN – Instituto Nacional de Normalización, NCh 801. Of. 2003. elementos de contrucción -Paneles- Ensayo de carga compresión, Chile.

[16] INN – Instituto Nacional de Normalización, NCh 802. EOf. 1971. Arquitectura y construcción - Paneles prefabricados - Ensayo de carga horizontal, Chile.

[17] INN – Instituto Nacional de Normalización, NCh 803. Of. 2003. elementos de contrucción - Paneles - Ensayo de carga flexión, Chile.

[18] C. Miehe, J. Shotte, M. Lambrecht, Homogenization of inelastic solid materials at finite strains based on incremental minimization principles. application to the texture analysis of polycrystals, Journal of the Mechanics and Physics of Solids 50 (10) (2002) 2123–2167.

[19] R. Castro-Triguero, E. Garcia-Macias, E. I. Saavedra Flores, M. I. Friswell, R. Gallego, Multi-scale model updating of a timber footbridge using experimental vibration data, Engineering Computations 34 (3) (2017) 754–780.

[20] ANSYS, Parametric Design Language Guide. Release 16.1, ANSYS, Inc., Canonsburg, PA 15317, http://www.ansys.com, 2016.

[21] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global sensitivity analysis. The primer, John Wiley & Sons, 2008.

[22] I. M. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation 55 (1-3) (2001) 271–280.

[23] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering & System Safety 52 (1) (1996) 1–17.

[24] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, Sensitivity analysis in practice: A guide to assessing scientific models, Wiley, 2004.

[25] J. Bernardo, A. Smith, Bayesian theory, Wiley Series in Probability & Statistics, Wiley, 1994.

[26] J. Oakley, Bayesian uncertainty analysis for complex computer codes, Ph.D. thesis, University of Sheffield (1999).

[27] W. Becker, Uncertainty propagation through large nonlinear models, Ph.D. thesis, University of Sheffield (2011).

[28] L. S. Bastos, A. OHagan, Diagnostics for Gaussian process emulators, Technometrics 51 (4) (2009) 425–438.

[29] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Computer Physics Communications 145 (2) (2002) 280–297.

[30] B. Efron, R. J. Tibshirani, An introduction to the bootstrap, CRC press, 1994.

[31] G. E. B. Archer, A. Saltelli, I. M. Sobol, Sensitivity measures, ANOVA-like techniques and the use of bootstrap, Journal of Statistical Computation and Simulation 58 (2) (1997) 99–120.

[32] E. I. Saavedra Flores, R. M. Ajaj, I. Dayyani, Y. Chandra, R. Das, Multi-scale model updating for the mechanical properties of cross-laminated timber, Computers and Structures 177 (2016) 83 – 90.