




DATA PAPER

Open Access



# Traceability and quality assessment of Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco) logs: the TreeTrace\_Douglas database

Fleur Longuetaud<sup>1\*</sup> , Guillaume Pot<sup>2</sup>, Frédéric Mothe<sup>1</sup>, Alexis Barthelemy<sup>1,3</sup>, Rémi Decelle<sup>1,3</sup>, Florian Delconte<sup>1,3</sup>, Xihe Ge<sup>2</sup>, Grégoire Guillaume<sup>2</sup>, Théo Mancini<sup>2</sup>, Tojo Ravoajanahary<sup>4</sup>, Jean-Claude Butaud<sup>2</sup>, Robert Collet<sup>2</sup>, Isabelle Debled-Rennesson<sup>3</sup>, Bertrand Marcon<sup>2</sup>, Phuc Ngo<sup>3</sup>, Benjamin Roux<sup>2</sup> and Joffrey Viguié<sup>2</sup>

## Key message

The *TreeTrace\_Douglas* database includes images and measurements at several stages of the processing of Douglas fir logs, from sawmill logyard to machine grading and destructive testing of boards, and is suitable for research on quality assessment and traceability. A total of 52 long logs, 156 short logs, 208 wood discs, and 346 boards were analyzed. The image data includes RGB images of log ends and board ends, RGB images and CT slices of strips, and a set of images of the boards (RGB, laser, and X-rays) obtained with an industrial board grading machine. The measurements include wood density, growth ring widths, pith and board location in the logs, heartwood and sapwood areas, mechanical properties of each board obtained by vibratory and static testing, and visual grading of the boards. Dataset is available at <https://doi.org/10.15454/YUNEGL> and associated metadata are available at <https://metadata-afs.nancy.inra.fr/geonetwork/srv/fre/catalog.search#/metadata/d9eef6e4-f195-41f4-b6c2-2ab46adc637e>.

**Keywords:** Image analysis, RGB images, Computed tomography, Wood density, Sapwood, Growth ring width, Mechanical properties, Lumber grading, Strength grading

## 1 Background

The *TreeTrace\_Douglas* database mainly provides RGB images, X-ray images, and wood quality measurements of logs, discs and boards, at several steps of the processing of Douglas fir logs, from the forest to the sawmill logyard, and from the sawmill logyard to the machine grading and destructive testing of boards. An effort was made to provide images of untreated log-ends (i.e., freshly sawn without any further preparation like sanding). The data

were collected in the framework of the TreeTrace project (ANR-17-CE10-0016).

Increasing the use of wood requires improving the organisation and competitiveness of the forest-wood chain and developing a circular bioeconomy (European Commission 2017). There is a challenge to better characterize and trace the wood all along the chain. With the increasing amount of imaging devices installed at sawmills, the opportunity of using these data for optimizing the use of wood material and for increasing revenues in the wood processing industries is growing.

The database was created for three purposes:

- Wood traceability along the forest-wood chain on the basis of intrinsic characteristics of logs or boards

Handling Editor: Véronique Lesage.

\*Correspondence: fleur.longuetaud@inrae.fr

<sup>1</sup> Université de Lorraine, AgroParisTech, INRAE, Silva, 54000 Nancy, France  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

(biometric methods like for fingerprint or iris recognition) by image processing of cross-sections only.

- Extraction of wood quality characteristics by processing images of cross-sections, applicable in the forest, on a log sorting platform or at the sawmill.
- Development of strength grading algorithms such as that of Viguiet et al. (2017), Olsson et al. (2018), and Rais et al. (2021) since the dataset includes the main measurements made by current industrial board grading machines: RGB images, fiber orientation data, X-ray local density maps, and vibratory testing. The location of the boards in the log can also be used in models like that of Lukacevic et al. (2019) and Hu et al. (2018).

The characterization and traceability of logs between the forest and industry are major issues today because they aim to guarantee the origin of wood, fight against illegal logging, contribute to maintaining a local and sustainable resource for the industries present, and optimize the flow of material and the transformation of this resource by taking better account of its characteristics.

In the literature, very few references propose algorithms for image processing of untreated log-ends. Some algorithms exist for images of sanded and polished cross-sections in the context of dendrochronology but these algorithms are not suitable to process images of rough cross-sections. A first pre-processing of untreated log-end images is to segment the wood cross-section in the entire image. The method proposed by Schraml and Uhl (2014) to segment wood cross-sections of spruce is a similarity-based region growing algorithm and requires pith location (Schraml and Uhl 2013). The final objective of the authors was then to trace logs from the forest to the sawmill (Schraml et al. 2015). More recently, neural networks were used to perform this task

(Wimmer et al. 2021). Another application in the field of wood quality is to discriminate heartwood and sapwood in cross-sections. Raatevaara et al. (2020) developed a method based on region growing techniques followed by a post-processing to segment heartwood on images of rough cross-sections of pine. For further processing it is often needed to locate the pith in cross-sections. For example, Norell and Borgefors (2008) estimated the pith location in order to further detect tree rings (Norell 2011). Kurdthongmee et al. (2018) also proposed an algorithm for the pith detection in images of untreated cross-sections. And more recently, deep neural networks were trained in order to address the problem of the pith detection in cross-section images of logs (Kurdthongmee 2020) and within boards from their longitudinal faces (Habite et al. 2021, 2022). The work already carried out with the *TreeTrace\_Douglas* database is listed in Section 6.

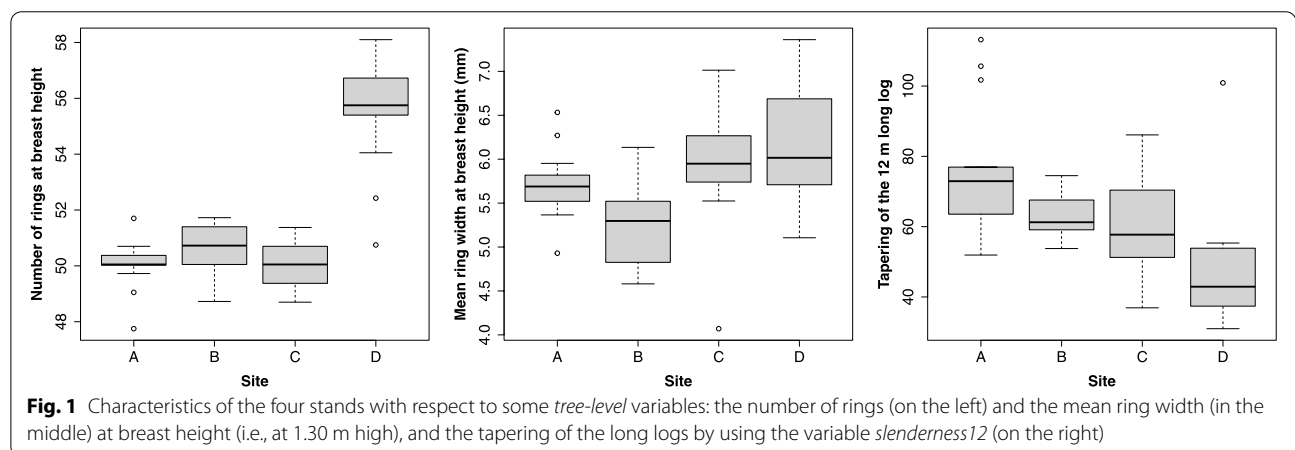
## 2 Methods used for the main sampling

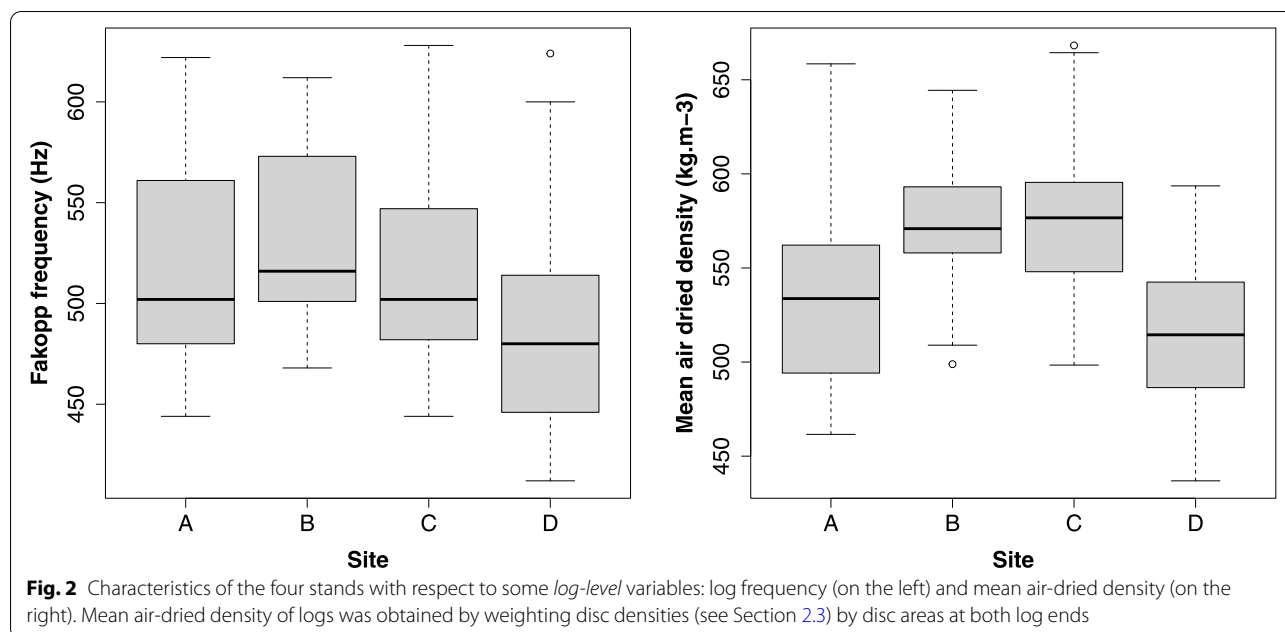
This sampling campaign was carried out in January 2020 in a sawmill of Bourgogne-Franche-Comté region.

### 2.1 Sampling and measurements on the logyard

Four harvesting sites were selected to provide the samples. In each stand, 13 trees were harvested and analyzed in details for wood quality. These trees were representative of what is usually processed at the sawmill. Since no details were available about the stands, they were named here A, B, C and D. Figures 1 and 2 illustrate the differences between the four sites due to stand management or site itself.

From these 52 trees, 52 long logs (about 12.5 m long) were transported to the logyard. An estimated breast height diameter (*D130Logyard\_cm*) was measured on the long logs. Measurements on wood discs sampled at 0 m and 4 m were used to





interpolate at breast height (1.30 m): the log diameter ( $D_{130Interpolated\_cm}$ ), the heartwood diameter ( $DHW_{130\_cm}$ ), the number of annual growth rings ( $nbRings_{130}$ ), the number of annual growth rings in the heartwood ( $nbRingsHW_{130}$ ), the mean ring width ( $ringWidth_{130\_mm}$ ), and the air-dry density ( $airDryDensity_{130\_kgpm3}$ ). A slenderness coefficient ( $slenderness_{12}$ ) was computed as:

$$slenderness_{12} = 12 / (D_0 - D_{12}) \tag{1}$$

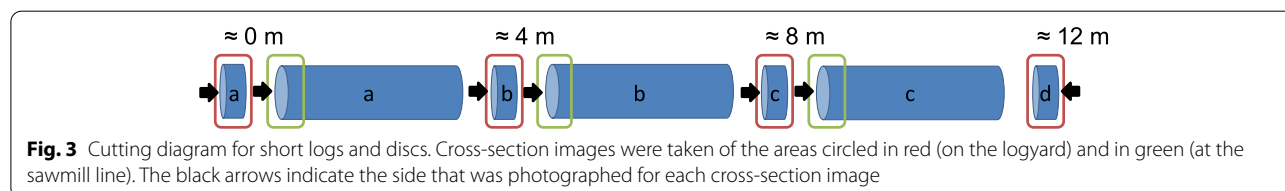
where  $D_0$  (m) is the log diameter at 0 m and  $D_{12}$  (m) is the log diameter at 12 m.

All the *tree-level data* are available in the file `TreeTrace_Douglas/tables/trees.txt`.

For each long log, three short logs were cut (Fig. 3). For example, B07c refers to the top log of tree #7 from site B. The length of each short log was measured ( $length\_cm$ ). The lengths of the bottom (logs #a), middle (logs #b), and top (logs #c) logs were in average 4.16 m, 3.66 m, and 4.18 m, respectively. For each short log, a frequency ( $frequency\_Hz$ ) was registered with the Fakopp Log Grader for Android (Divos 2014, <https://fakopp.com/en/product/rlg/>).

In the same time, four wood discs (#a, #b, #c and #d) per long log were sampled (Fig. 3). For example, B07c refers to the wood disc sampled at the large end of the short log with the same name, B07c. And B07d refers to the wood disc sampled at the small end of the short log B07c.

For each of the three short logs, one image of the large end was taken on the logyard, with the log generally held by the harvester, before cutting the corresponding disc (Fig. 4 on the left). A chainsaw was used for some of the bottom logs #a when the diameter was too big for the harvester (Fig. 4 on the right). For the last short log (log #c), one image of the small end was also taken before cutting disc #d. Images on the logyard were taken with an iPhone X (f/1.8 aperture, focal length 4 mm). *Two hundred eight RGB images* (12 MPx, 4032 × 3024 px) are available as “4 sites × 13 trees × 4 discs” (+ 208 images with labels) on the data repository in `TreeTrace_Douglas/logs_and_discs/color_images/discs`. A 25-cm-long calibration scale was used to calibrate spatially each image. The pixel widths in millimeters for these images are given by the variable





**Fig. 4** Large ends of logs A12d, B05c, C04b, and D04a (from left to right) at the logyard. The image of log D04a was taken directly on the long log; the others were taken with the log held by the harvester. The 25-cm-long scale was used to calibrate each image

*pixelWidthLogyard\_mm*. In some cases, the photos were poorly framed and the wood cross-sections could be incomplete. A variable *incompleteAreaUBLogyard* indicates whether the section was significantly incomplete (= 1) or complete or nearly complete (= 0). For each image, a manual processing was done to locate the pith (*pithX\_pixel* and *pithY\_pixel*) and to delineate the outline of the wood cross-section under-bark and the outline of the heartwood (Fig. 5 on the top-right). The segmented images are available in *TreeTrace\_Douglas/logs\_and\_discs/segmented\_images/discs*. For each image, the under-bark area of the wood cross-section was computed (*areaUBLogyard\_cm2*) as well as the area of the heartwood (*areaHWLogyard\_cm2*).

In total, 156 short logs were cut and 208 wood discs were collected to be analyzed for wood quality (see section 2.3). For simplicity, the heights of the bottom of short logs (*bottomHeight\_m*) and of wood discs (*height\_m*) were estimated at 0 m, 4 m, 8 m, and 12 m, for the levels #a, #b, #c, and #d, respectively. All the *log-level data* are available in the file *TreeTrace\_Douglas/tables/logs.txt* and all the *disc-level data* are available in the file *TreeTrace\_Douglas/tables/discs.txt*.

### 2.2 Sampling at the sawmill line

The objective here was to automatically take images of large log-ends on the sawmill line. The camera was installed just after debarking, before the main saw unit, at



**Fig. 5** Images of log C04a: Painted small end and large end images taken on the logyard, and corresponding manually segmented image (first row); large end image taken on the sawmill line and corresponding manually segmented image (second row). On the segmented images: The white area is the sapwood located between heartwood and bark; the red area is the heartwood



the top of the wood log loader. For the triggering, a laser telemeter was used.

The selected camera was a BASLER ACE acA5472-5gm/gc (20 MPx, 5472 × 3648 px). Its sensor width was 13.1328 mm and its focal length was 16 mm. A custom software to manage the camera, the telemeter, and the data was developed by a software application engineer from AMVALOR.

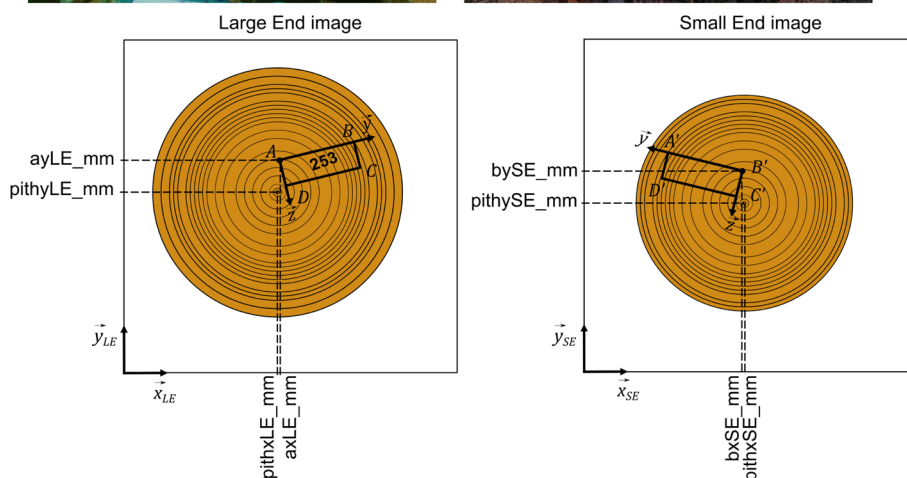
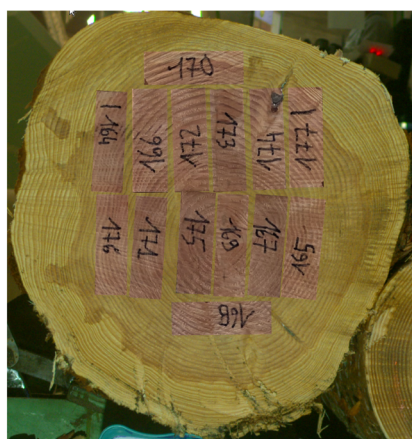
The distance between the camera and the log-end ( $logDistance\_m$ ) arising from the laser telemeter used to trigger the log end capture was used to compute the pixel size ( $pixelWidthSawmill\_mm$ ) in the images as follows:

$$\begin{aligned}
 imageWidth\_mm &= 13.1328/16 \times logDistance\_m \times 1000 \\
 pixelWidthSawmill\_mm &= imageWidth\_mm/5472
 \end{aligned}
 \tag{2}$$

where  $imageWidth\_mm$  is the width in millimeters of an object placed at  $logDistance\_m$  from the camera that would occupy the entire width of the image.

From the 156 short logs, 4 were not sawn and kept to be processed by rotary peeling (A03b, B10b, C13b, and

D12b) and to manufacture laminated veneer lumber as described in Duriot et al. (2021a, b). A variable *not-Sawn* was added to the corresponding data file to identify these four logs. The 152 remaining logs were sawn. For a sub-sample of 32 short logs (4 sites × 4 trees × 2 logs per tree, logs #a and #c), boards were collected to be analyzed for quality at Arts et Metiers Institute of Technology (LaBoMaP). The sawing pattern was designed to produce a single board size of section 55 mm × 160 mm with one or two stacks of boards depending on the log diameter and optional edge boards. Figure 6 illustrates a sawing pattern with two stacks of six boards each and two additional edge boards. The variable *nbBoards* indicates the number of boards collected for each log. A total of 346 boards was obtained from the 32 short logs (some edge boards were not collected). The small end of these 32 logs was painted so that they could be recognised during the sawing and that the location of the sawn boards could be reconstructed (Fig. 5 on the top-left and Fig. 6).



**Fig. 6** Manual positioning of the board end images on the end images of a log: example of log C04a at the top of the figure (large end on the left, small end on the right), and illustration of how the position of a board is analyzed in these images at the bottom of the figure

The logs external envelopes measured by a laser scanner (Optilog from Sovilor) were used to estimate log volumes (*volumeUBLaserScanning\_m3*). A log volume named *volumeUB\_m3* was also estimated from the log length (*length\_cm*) and the under-bark areas measured on the images of both log ends (in the sawmill or on the logyard depending on the case).

RGB images were acquired by the camera on the line. One hundred fifty-two RGB images are available as “4 sites  $\times$  13 trees  $\times$  3 logs – 4 rotary cut logs” on the data repository in `TreeTrace_Douglas/logs_and_discs/color_images/logs/large_ends`. The same manual processing of images as the one described in Section 2.1 was done and segmented images (Fig. 5 on the bottom-right) are available in `TreeTrace_Douglas/logs_and_discs/segmented_images/logs`. The log-level data file contains for these images the pith coordinates (*pithX\_pixel* and *pithY\_pixel*) and the under-bark area of the wood cross-section (*areaUBSawmill\_cm2*) as well as the area of the heartwood (*areaHWSawmill\_cm2*).

### 2.3 Measurements on wood discs

The wood discs were sawn in the freshly cut state to get four radial strips (about 8 cm thick in the tangential and longitudinal directions). The strips were numbered 1 to 4, always starting from the longest radius. A total of 810 samples was obtained, some of the discs being too damaged to provide all four strips. After several weeks of air-drying, the radial-tangential surface of each strip was sanded and scanned in 400 dpi (i.e., pixel width about 0.06 mm). These RGB images are available on the data repository in `TreeTrace_Douglas/strips/color_images`. The samples were then air-dried several months and processed for density measurements by a medical X-ray CT scanner (General Electric Brightspeed Excel). The scanner settings were as follows: 80 kVp, 50 mA, “DETAIL” convolution filter, and 1.25 mm of slice thickness. One CT slice only was collected for each wood disc, with all strips stacked together, with a pixel size of 0.98 mm  $\times$  0.98 mm. The images were calibrated using the protocol described by Freyburger et al. (2009) to compute wood density for each pixel. These original images in DICOM format and the calibrated images in TIF format are available on the data repository in `TreeTrace_Douglas/strips/xray_images`.

The ImageJ software (Schneider et al. 2012) was used to record on each strip a radial profile of density (`TreeTrace_Douglas/tables/strips_density_profiles.txt`). The Plot profile function of ImageJ was used to draw a 10 pixel wide line from pith to wood/bark boundary and to record the density values at each pixel

position along the line (averaging 10 pixels in the perpendicular direction).

Most of the *ring-level data for the radial strips* (`TreeTrace_Douglas/tables/strips_rings.txt`) were measured on the RGB images using the Windendro software (WinDendro 2021). Cambial age (*cambialAge*), growth year (*growthYear*), distance to pith (*externalRadius\_mm*), and ring width (*ringWidth\_mm*) were recorded for each ring of each radial strip. The radial profile of ring widths was coupled with the radial profile of density to assess the mean density of each ring (*airDryDensity\_kgpm3*). The coupling algorithm consisted first in establishing from the density profile of each sample a linear interpolation function of the air-dry density in any relative radial position. Then, for each annual ring, 100 relative radial positions from the beginning to the end of the ring were considered to generate 100 interpolated density values. The mean of these 100 values was assigned to the ring density.

The ring-level data from the orthogonal radial strips were then averaged to obtain the *ring-level data for the discs* (`TreeTrace_Douglas/tables/rings.txt`). The annual ring widths of a disc (*ringWidth\_mm*) were obtained by subtraction of successive quadratic means of the measured radii (i.e., root mean square of external radii of a given ring minus root mean square of internal radii of the same ring). The mean density of each ring (*airDryDensity\_kgpm3*) was obtained by weighting the ring density by the area represented by the ring on each radial strip.

The ring-level data on the radial strips were also used to compute *strip-level data* (`TreeTrace_Douglas/tables/strips.txt`). The number of rings (*nbRings*), the average ring width (*ringWidth\_mm*), the number of heartwood rings (*nbRingsHW*), and the distances from pith to sapwood (*lengthHW\_mm*) and from pith to bark (*length\_mm*) were calculated. The mean density of each strip (*weightedDensity\_kgpm3*) was computed by averaging the pixel values weighted by their theoretical representative area, i.e., by the area of a one pixel wide ring of radius equal to the distance to the pith. The unweighted mean density of the pixel values (*rawDensity\_kgpm3*) was also computed.

Then, the average ring width and the average wood density of the discs were computed as follows and added to the *disc-level data* (`TreeTrace_Douglas/tables/discs.txt`). The average ring width (*ringWidth\_mm*) of each disc was computed as the arithmetic mean of all annual ring widths. The average density (*airDryDensity\_kgpm3*) of a disc was computed as the average density of the strips (*weightedDensity\_kgpm3*) weighted by the disc area represented by each strip using the strip length (i.e.,  $\pi \times \text{length\_mm}^2$ ).

### 2.4 Measurements on boards

All *board-level data* are available in the file `TreeTrace_Douglas/tables/boards.txt`.

In this section, for a better reading of the equations, easy-to-read symbols have been associated to some of the variable names recorded in the data table. They are described in Table 1.

The 55 mm × 160 mm boards sawn from the logs were dried and then planed to a nominal cross-section of 45 mm × 145 mm. A reference end was chosen for each board, which was the end corresponding to the large end of the log from which it came. A “Top” face was arbitrarily chosen for each board, and the top left corner, when looking at the side of the reference end, was named A, and defined as the origin of the board coordinate system  $(A, \vec{x}, \vec{y}, \vec{z})$  (Fig. 7). As shown in the Fig. 7, “Left,” “Right,” and “Bot” (abbreviated from bottom) faces were defined relatively to the Top face when looking at the reference end side. The board number was handwritten on the reference end, always readable

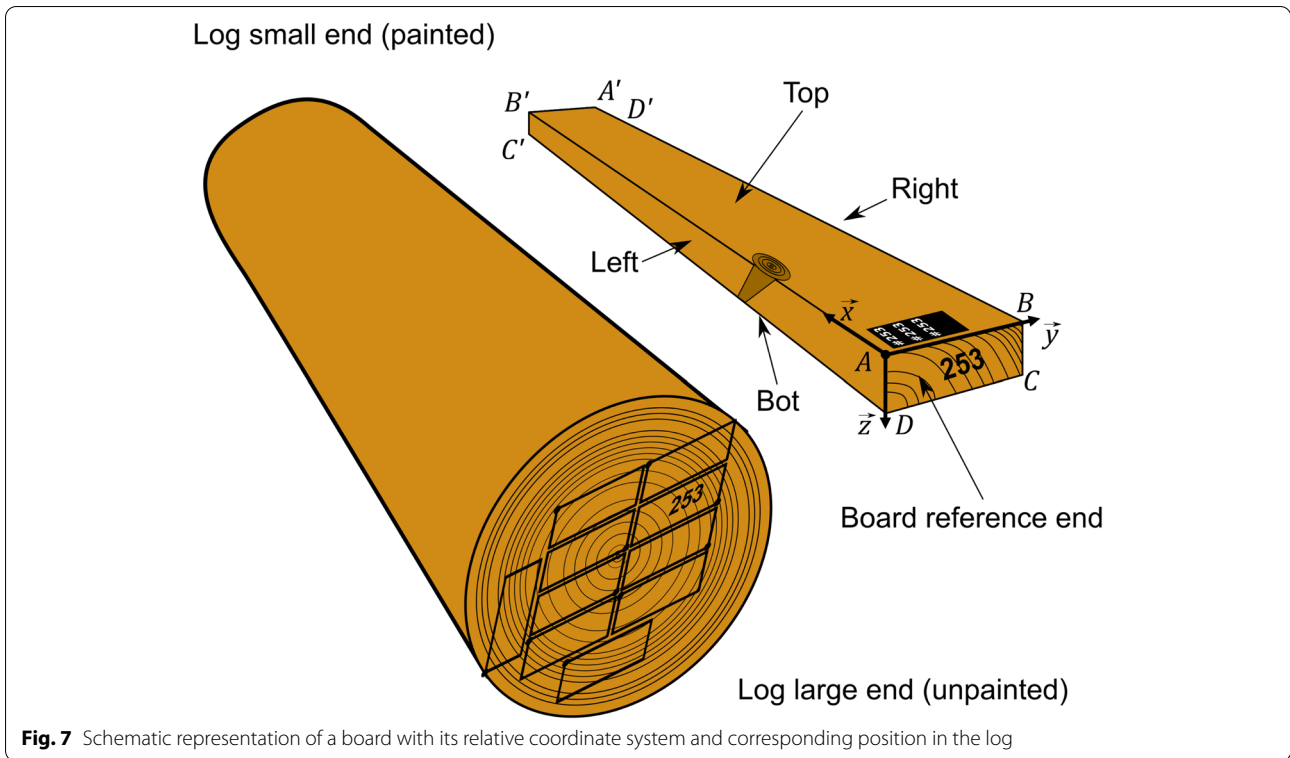
from left to right when Top face is upside. In addition, the board number was printed in white on black on adhesive paper, and affixed on the Top face on the side of the reference end, always readable from left to right when left face is upside (i.e. board defined top left corner “A” placed in the top left corner when looking at the Top face). The board number was printed several times on this adhesive paper in the following manner: “#253”, as seen in the example of Fig. 7. This is the *boardID* recorded in the *board-level data* table.

#### 2.4.1 Recording the position of boards within logs

The log ends were photographed on the logyard for the small ends (painted faces) (Fig. 5 on the top-left), on the sawmill feeder for the large ends (Fig. 5 on the bottom-left). After processing, the 346 boards were stacked into five piles that were also photographed in the fresh state (thus prior to planing) on both sides, with the board numbers handwritten on the unpainted side in the order of arrangement of the boards (Fig. 8). After several

**Table 1** List of main symbols and corresponding variable names in data Tables

Variable name (in data tables)	Variable symbol (in equations)	Variable description	Unit
<i>length_m</i>	$L$	Board length measured manually with a tape	m
	$\vec{x}$	Lengthwise direction of the board	
<i>width_m</i>	$w$	Board width measured manually with a caliper	m
	$\vec{y}$	Crosswise direction parallel to the width of the board	
<i>thickness_m</i>	$t$	Board thickness measured manually with a caliper	m
	$\vec{z}$	Crosswise direction parallel to the thickness of the board	
<i>massManual_kg</i>	$m_{board}$	Board mass measured manually with a scale	kg
<i>boardDensityManual_kgpm3</i>	$\rho_M$	Board density measured from weight measured manually	kg m <sup>-3</sup>
<i>boardDensityXRays_kgpm3</i>	$\rho_{XR}$	Board density computed from X-ray data	kg m <sup>-3</sup>
<i>cwDensity12pc_kgpm3</i>	$\rho_{CW}$	Clear wood density of the small clear wood sample at 12 % moisture content	kg m <sup>-3</sup>
<i>mcCWBendingTest_pc</i>	$u_{CW}^{test}$	Moisture content of the small clear wood sample at the moment of the destructive bending test	%
<i>firstAxialFreq_Hz</i>	$f_{a,1}$	Board first axial resonance frequency	Hz
<i>dynMoEAxial_Pa</i>	$E_{a,dyn}$	Board dynamic modulus of elasticity obtained from the first axial resonance frequency	Pa
<i>dynEWBendFreqN_Hz</i>	$f_{b,N}$	Board N <sup>th</sup> edgewise transversal frequency	Hz
<i>dynEWBendMoEBernoulliN_Pa</i>	$E_{b,dyn,N}$	Board dynamic modulus of elasticity obtained from N <sup>th</sup> edgewise transversal frequency	Pa
<i>dynEWBendMoETimoshenko_Pa</i>	$E_{b,dyn}$	Board dynamic modulus of elasticity obtained from Timoshenko bending theory	Pa
<i>dynEWBendShearMoE_Pa</i>	$G_{dyn}$	Board dynamic shear modulus obtained from Timoshenko bending theory	Pa
	$l_{load}$	Distance between loading heads in the 4-point bending test	m
<i>supportSpan_m</i>	$l$	Span of the 4-point bending test	m
<i>startloadMoE_N</i>	$F_1$	Load to which board static modulus of elasticity starts to be calculated	N
<i>endLoadMoE_N</i>	$F_2$	Load to which board static modulus of elasticity ends to be calculated	N
<i>localStatEWBendMoE_Pa</i>	$E_{m,l}$	Board static local modulus of elasticity as defined by EN 408:2012	Pa
<i>globalStatEWBendMoE_Pa</i>	$E_{m,g}$	Board static global modulus of elasticity as defined by EN 408:2012	Pa
<i>maxLoadStatEWBend_N</i>	$F_{max}$	Ultimate load in 4-point bending	N
<i>strenghtEWBend_Pa</i>	$f_m$	4-point bending strength of the board	Pa



**Fig. 7** Schematic representation of a board with its relative coordinate system and corresponding position in the log

months of air-drying, the boards were planed and new images of the piles of boards were taken. The images are available on the data repository in *TreeTrace\_Douglas/boards/piles*.

The images of the board piles taken in the fresh state were used by an operator to manually segment both ends of each board on the pile images (assuming that the section was rectangular and parallel to the image edges). Then, by scaling, rotating, and translating these board images, their positions on the log end images were found (Fig. 6). This work was done with the

open-source software LibreOffice Draw (<https://www.libreoffice.org/>) which allows to display the coordinates, dimensions, and rotation angle of each object in the page. The 2D coordinates were transformed to be relative to the bottom left corner of each image. The dimensions in millimeters of the scaled images of board ends at fresh state on the large end side of the log (abbreviated “LE”), *freshWidthLE\_mm* and *freshThicknessLE\_mm*, and the rotation angle of the board in the log large end image (*rotationAngleLE\_deg*) were recorded in the *board-level data* table. The scale of the corresponding



**Fig. 8** A pile of boards at the sawmill: Small ends covered with paint on the left, large ends with board numbers on the right. The 25-cm-long scale was used to calibrate the images in size



log end image  $scaleLE\_pixelspmm$  was also given. The coordinates of the top left reference corner, named A as in Figs. 6 and 7, were given ( $axLE\_mm$  and  $ayLE\_mm$ ). In addition, the coordinates within the log end image of the three other corners of the board end image (named B, C, D clockwise, see Fig. 6 and 7) were computed and provided in the board-level data table. The same variables were recorded for the small end side (abbreviated “SE”), but for letters A’, B’, C’, and D’ as shown in Figs. 6 and 7, with coordinates named  $axSE\_mm$ ,  $aySE\_mm$ ,  $bxSE\_mm$ , etc.

### 2.4.2 Dimensions

Each board length (symbol  $L$  in the equations, name  $length\_m$  in the data table) was measured with a measuring tape (length 5 m, precision 0.5 mm). Since boards were not cross cut in length after sawing, the tolerance of the boards length can be up to 1 cm.

Thickness ( $t$ ,  $thickness\_m$ ) and width ( $w$ ,  $width\_m$ ) dimensions were measured with a digital caliper (resolution 0.01 mm, precision 0.01 mm) in three positions near the center and at the two ends of the boards. Their average values were immediately calculated and recorded. It should be noted that despite the planing, some boards

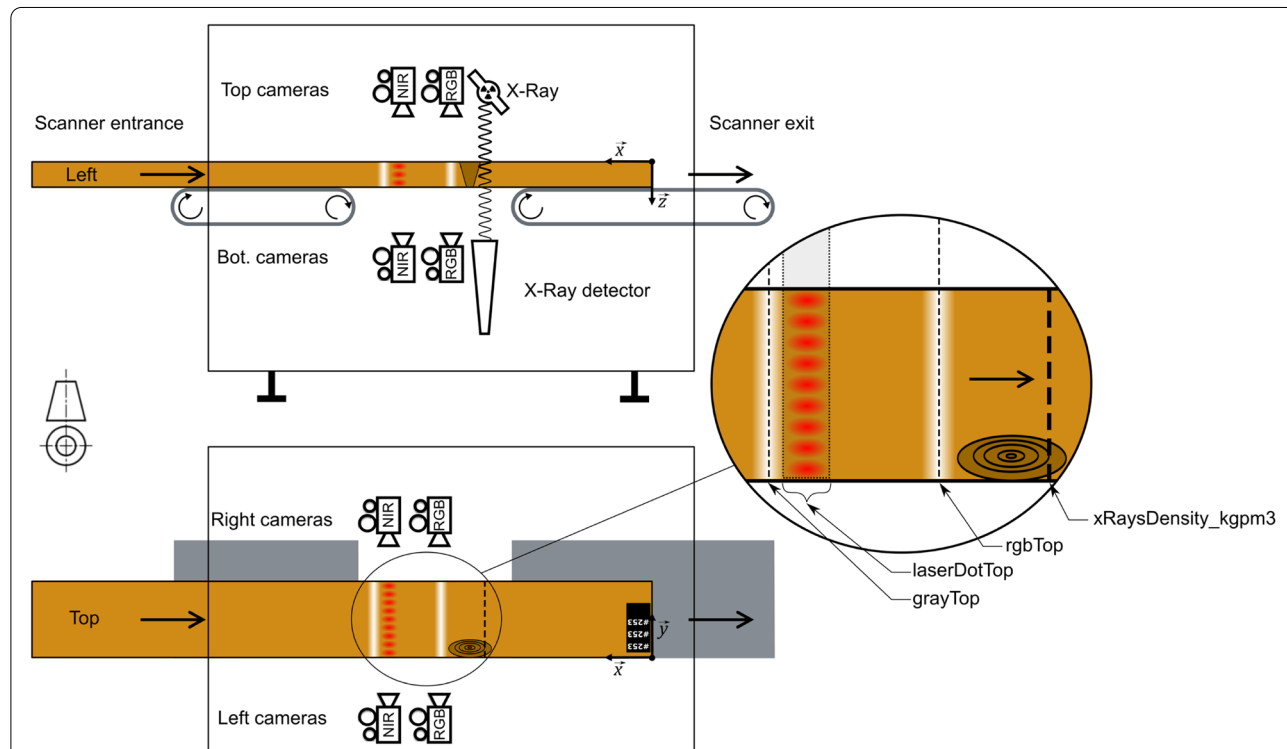
could exhibit 2 mm of width variation along the length, which has been reported in the column *observations*.

The global mass of each board ( $m_{board}$ ,  $massManual\_kg$ ) was measured with an OHAUS numerical balance (resolution 0.01 kg) at the moment of the non-destructive tests. As this measurement was done manually, it is subject to errors, and thus boards #75 and #102 had wrong mass informations which were recorded as NA.

### 2.4.3 Scanning

An industrial scanner was used to analyze the boards. By conveying the boards longitudinally, this scanner provides different images of the four long faces of the boards, as well as local density and local orientation of the fibers. The principle of board scanning is presented in Fig. 9. All the boards were fed in the scanner with the board reference end first, printed number on the top. All the boards were successfully scanned; however, it should be noted that for board #206, a small foreign object appears in the top images and in the X-ray data.

In the scanner, there are successively between two conveyors, for each of the four faces of a scanned board: a white LED lighting, a row of near infrared (830



**Fig. 9** Illustration of board scanning with side view on the top of the figure, and top view on the bottom of the figure, corresponding to left and top faces of the board, respectively. In the detail view, dashed lines represent a line of pixels which is recorded by the cameras while the board is conveyed longitudinally, and the grayed area with dotted line contour represents the region of interest used in the camera to treat ellipses information

nm) laser dots, and another white LED lighting. There is a first row of four monochrome matrix cameras (one for each board face), and then a second row of four linear RGB cameras, all of them with the top of the field of view facing toward the entrance of the scanner. The monochrome cameras of the first row are sensible in near infrared region. Finally, there is a X-ray source and a X-ray detector to measure the local density in the  $(\vec{x}, \vec{y})$  plane of the boards.

Although the first row of cameras were matrix cameras, they were mainly used as linear cameras, with several regions of interest (ROI) consisting in single lines of pixels. These ROI are shown in the detail view of Fig. 9. The images were obtained by the longitudinal ( $\vec{x}$  direction) displacement of the board during the scanning at the speed of  $62 \text{ m min}^{-1}$ . As a result, the reconstructed images of the four sides of the boards appeared as shown in Fig. 10, with the board coordinate system shown on it.

The longitudinal resolution depends on a roller encoder (not represented in Fig. 9), so it is the same for all the data:  $1.00 \text{ pixel mm}^{-1}$ . A roller presser ensures that there

is no slipping. The resolution in the transverse directions ( $\vec{y}, \vec{z}$ ) has been determined for each variable and is given in the following.

Since the positions of the cameras relative to the different faces were not exactly the same, adjustments have been made to the reconstructed image data in order that the longitudinal position of the board for each image is the same in all the data. As a result, for each board, all the images obtained from the cameras have the same number of pixels in the longitudinal direction of the board, and the longitudinal position of a pixel is in the same position physically in the board coordinate system whatever the data considered. Conversely, in the transverse directions of the boards, firstly the transverse resolution change depending on the position of the ROI in the field of view of the cameras and their angle relative to board surface, and secondly, boards that were not perfectly straight, boards that were not fed perfectly in line in the scanner, and vibrations of the boards induce many regions of black pixels (with 0 value) in the images (see Fig. 10). As a result, the number of pixels varies from one source of

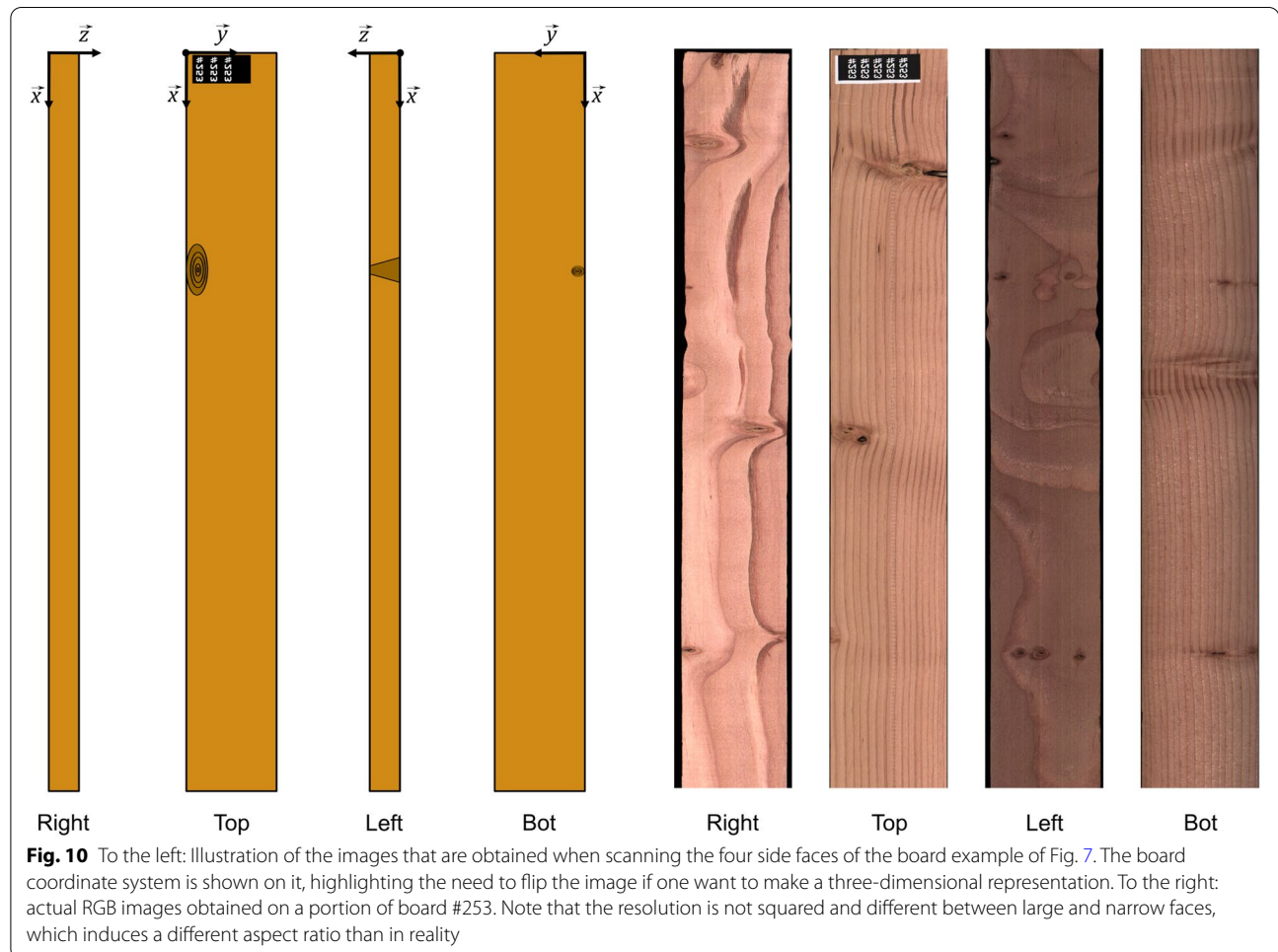


image to another. The images have been cropped so that there is the minimum of unnecessary pixels; however, this is a problem that remains to be solved if one want to make a perfect three-dimensional representation.

The images and data obtained from the scanner for each board are stored in the data repository TreeTrace\_Douglas/boards/faces in several images and text files named as “boardboardID\_variableName.png” or “boardboardID\_variableName.txt,” each of the *variableName* corresponding to the variables detailed in the following and represented in the detail view of Fig. 9.

- **Gray level images** *grayTop*, *grayBot*, *grayLeft*, and *grayRight* correspond to the reconstructed images of the line of pixels located in the middle of the white LED lighting in the NIR cameras for the top, bottom, left, and right faces, respectively. The measured corresponding resolutions are 4.30, 4.29, 8.61, and 8.57 pixels mm<sup>-1</sup> for the top, bottom, left, and right faces, respectively.
- **laserDot data** are tabulated text data resulting from laser dots detection and processing. For each board face, a 830-nm laser dot is diffracted in several laser dots thanks to a diffractive optical element, resulting in lines of dots spaced of 4 mm. Due to anisotropic light diffusion properties of wood, the observed light from a laser dot on the surface of wood resembles the shape of an ellipse. In the literature, this phenomenon is called the “tracheid effect,” because it was first studied on softwoods (Nystrom 2003; Simonaho et al. 2004). The major axis of the observed ellipse is oriented in the same direction as the fiber orientation, or more precisely in the same direction as the projection of the fiber direction on the surface of the board (thus, there is no information about the diving angle). In the near infrared cameras, the region of interest including the line of dots shown in the detail view of Fig. 9 was binarized with the same threshold for all four faces. The contour of the binarized ellipses were directly recorded during scanning. Then, these data were post-processed and the contours were fitted to an ellipse equation in the least square sense. The reader can refer to Besseau et al. (2020) for more details about this subject. The post-processing step provides for each ellipse, in each *laserDotTop*, *laserDotBot*, *laserDotLeft*, and *laserDotRight* table:

- In first column: The transverse position of the center of the fitted ellipse in pixels, with a transverse resolution of 4.24, 4.23, 8.58, and 8.58 pixels mm<sup>-1</sup> for the top, bottom, left, and right faces, respectively;
- In second column: The longitudinal position of the center of the fitted ellipse in pixels, in the

coordinate system of the other images and thus same resolution;

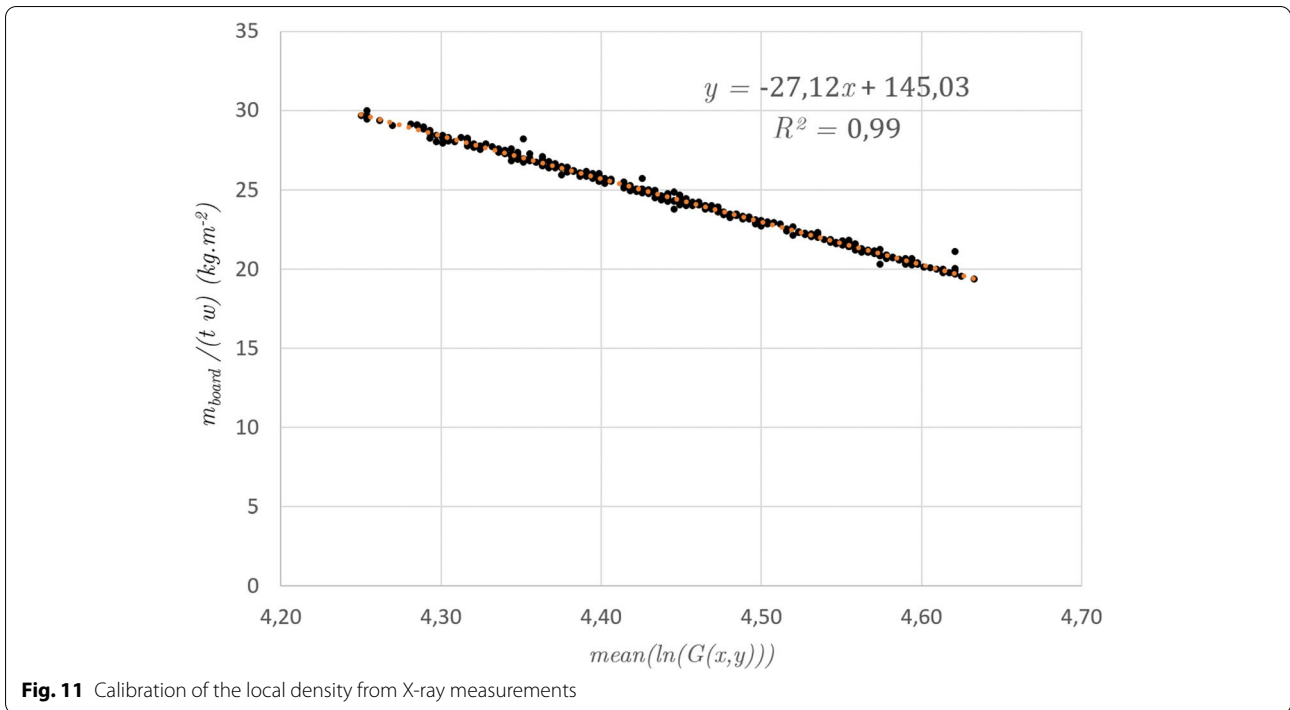
- In third column: The angle in degrees of the major axis of the fitted ellipse relatively to the image coordinate system as shown in Fig. 10, thus the angle should be multiplied by -1 on Left and Bot faces if ones want to be in the same board coordinate system for all four faces;
  - In fourth column: The length of the fitted ellipse major axis, in pixels;
  - In fifth column: The length of the fitted ellipse minor axis, in pixels.
- **rgb images** *rgbTop*, *rgbBot*, *rgbLeft* and *rgbRight* correspond to the reconstructed images of the line of pixels located in the middle of the white LED lighting in the RGB cameras for the top, bottom, left, and right faces, respectively. The corresponding resolutions are 3.79, 3.75, 10.84, and 10.65 pixels mm<sup>-1</sup> for the top, bottom, left, and right faces, respectively.
  - **Local density** was determined by means of a source emitting X-rays across the large face of the boards and a line sensor recording the photons having gone through the thickness of the board on the opposite face while each board was conveyed longitudinally. This scanning resulted in a two-dimensional map of 8-bit (0-255) gray levels in the ( $\bar{x}, \bar{y}$ ) plane,  $G$  (following board coordinate system as defined in Fig. 9). This raw data has been converted into local density maps, named *xRaysDensity\_kgpm3*, by using Beer-Lambert law with a calibration made on the boards for which the mass was available (all but two): each board mass per unit of surface can be written as a linear function of the average of the logarithm of the pixel values:

$$\frac{m_{board}}{w \times L} = a \overline{\ln(G)} + b \tag{3}$$

where  $a$  and  $b$  are linear calibration coefficients obtained after fitting a linear regression curve.  $\overline{\ln(G)}$  is the average computed only for pixel values included between 10 and 240; this thresholding allows to take into account only the area where there is wood material without voids. The calibration curve is presented in Fig. 11 and gives  $a = -27.12 \text{ kg m}^{-2}$  and  $b = 145.03 \text{ kg m}^{-2}$ . The local density map *xRaysDensity\_kgpm3* was computed for each thresholded pixel of each board as:

$$xRaysDensity\_kgpm3 = \frac{-27.12 \ln(G) + 145.03}{t} \tag{4}$$

The measured transverse resolution is 2.93 pixels mm<sup>-1</sup>. The pixels removed from the computation



**Fig. 11** Calibration of the local density from X-ray measurements

because of the threshold are given a NaN value. The density data maps have been saved in 16-bit images with the gray level value (theoretically between 0 and 65,535) being directly the rounded density values in  $\text{kg m}^{-3}$ .

**2.4.4 Boards global density**

The global mass measured for each board and its dimensions measured manually allowed to compute the global air dried density of each board, *boardDensityManual\_kgpm3* (symbol:  $\rho_M$ ) as:

$$boardDensityManual\_kgpm3 = \rho_M = \frac{m_{board}}{Ltw} \quad (5)$$

Because of missing values in mass and incertitude in the board lengths, this density value was not used and the density obtained from boards X-Rays scanning is preferred. For this purpose, for each board, the average of non-zero local density contained in *xRaysDensity\_kgpm3* maps has been computed to obtain the air dried average density of the whole board, *boardDensityXRays\_kgpm3* (symbol:  $\rho_{XR}$ ), which is recorded in the board-level data table.

**2.4.5 Vibratory measurements**

Every board has been subjected to an axial vibratory test and a transversal (bending) vibratory tests in edgewise

direction, i.e.,  $(\vec{x}, \vec{y})$  plane. The boards were put on two elastic supports to simulate free-free conditions (distance between the supports 207 cm). Vibrations were generated by a hammer impact on the reference end side of each board, and a microphone recorded the vibrations at the other end. BING<sup>®</sup> software (Paradi et al. 2017) was used to record the full signals and to detect the resonance frequencies. The full temporal signals are available for each board and type of vibratory test in separate text files in the data repository *TreeTrace\_Douglas/boards/grading/vibratory\_tests*. The text file *TreeTrace\_Douglas/boards/grading/vibratory\_tests/readme.txt* give some details about those data. The bending, axial, and shear moduli of elasticity (MoE) were computed as explained in the following and recorded in the board-level data table.

The dynamic MoE for axial vibrations *dynMoEAxial\_Pa* (symbol:  $E_{a,dyn}$ ) of a board was determined by using the first compression longitudinal resonance frequency ( $f_{a,1}, firstAxialFreq\_Hz$ ) as:

$$dynMoEAxial\_Pa = E_{a,dyn} = 4L^2 \rho_{XR} f_{a,1}^2 \quad (6)$$

This equation comes from classical board theory, supposing the board to be homogeneous and on free-free boundary conditions. Both the frequency and the MoE computed from this equation are recorded in the board-level data table.



For edgewise bending vibrations, the four first bending frequencies were recorded (name *dynEWBendFreqN\_Hz*, symbol  $f_{b,N}$  with N replaced by 1 to 4). The Bernoulli model was first applied to compute the corresponding dynamic Bernoulli MoEs in edgewise bending, named *dynEWBendMoEBernoulliN\_Pa* (symbol:  $E_{b,dyn,N}$ ) following the equation:

$$dynEWBendMoEBernoulliN\_Pa = E_{b,dyn,N} = 4\pi^2 L^4 \rho_{XR} \cdot \frac{A f_{b,N}^2}{I_{Gz} P_N} \tag{7}$$

with :  $A = tw$

$$I_{Gz} = \frac{tw^3}{12}$$

By supposing the board to be on free-free boundary conditions,  $P_N$  is a parameter obtained when solving the equation  $\cos \sqrt[4]{P_N} \sin \sqrt[4]{P_N} = 1$ , giving  $\sqrt[4]{P_1} = 4.73004$ ,  $\sqrt[4]{P_2} = 7.85320$ ,  $P_N = (2N + 1)\pi/2$ . Each frequencies and Bernoulli MoEs were computed from this equation and recorded in the board-level data table. A NA value is provided when the frequency was not observable (fourth frequency of board #12). Bernoulli theory assumes pure bending; thus, there is no effect of shearing, which is a very approximative assumption, especially for higher order modes.

Timoshenko bending theory, which accounts for shear effects, can be used to determine the dynamic Timoshenko MoE in edgewise bending (*E<sub>b,dyn</sub>, dynEWBendMoETimoshenko\_Pa*) and shear modulus in edgewise bending (*G<sub>dyn</sub>, dynEWBendShearMoE\_Pa*) by using several vibration modes. Brancheriau and Bailleres (2002) solution of the vibrating board's equation of motion at the first order can be applied as:

$$y_N = \frac{E_{b,dyn}}{\rho_{XR}} - \frac{E_{b,dyn}}{KG_{dyn}} \cdot x_N \tag{8}$$

where  $K$  is the shear factor ( $K = \frac{5}{6}$  for a rectangular cross-section) and  $x_N$  and  $y_N$  are parameters that depend on the  $N^{th}$  vibrational mode frequency (see Brancheriau and Bailleres (2002) for details). Therefore,  $E_{b,dyn}$  and  $G_{dyn}$  can be calculated by using linear regression, knowing the different points of vibrational mode frequency. The coefficient of determination *r2dynEWBendMoETimoshenkoCalculation* is provided in the board-level data table along with these moduli. According to Brancheriau and Bailleres (2002), the deviation of the equation is less than 1% if the sample's length/depth ratio is between 10 and 20. Here  $L/t$  is about 29, thus inducing a higher error. Most importantly, the most heterogeneous boards do not fulfill the homogeneous assumption thus implying low

*r2dynEWBendMoETimoshenkoCalculation* value and/or incoherent  $G_{dyn}$  value (some of them are negatives).

### 2.4.6 Visual grading

Visual grading is a straightforward method to grade timber allowed by European Standards. Each board has been graded according to NF B52-001-1 (2018) visual grading standard. There are 19 criteria based in this standard, although growth ring width and knot size are usually the most restrictive and essential. These two parameters have been measured and recorded, along with some other characteristics of the boards which are described below.

- *Board end grain pattern* varies according to the position of the boards in the logs and the way the log sawing is carried out. Although it is not a parameter to take into account following NF B52-001-1 (2018), it is related to the orientation of knots into the board and may induce differences in mechanical behavior. Four different sawing patterns were recorded after visual observation: the C letter in *boardEndGrainPattern* refers to boards for which the pith of the tree was visible wherever possible along the board ("coeur" in French); D was for flat sawn type end grain pattern (dosse in French), when both ends of the board exhibited growth rings almost parallel to the wide faces of the board; Q was for quarter sawn type end grain pattern, when both ends of the board exhibited growth rings almost parallel to the narrow faces of the board; FQ was for false quarter or rift sawn type end grain pattern, when at least one end of the board exhibited growth ring not parallel to the faces or the sides of the board.
- *Growth ring width* was measured according to NF B52-001-1 (2018) at a board end by measuring the longest perpendicular line to the growth rings (shift 25 mm from the pith if it is included in the board) and dividing it by the number of crossed growth rings. The mean growth ring width *growthRingWidth\_mm* was then calculated for each board as the average growth ring width between the two board ends.
- *Knot size* was measured using a digital caliper (resolution 0.01 mm, precision 0.01 mm) according to NF B52-001-1 (2018), i.e., perpendicularly to the board longitudinal axis. The size of the largest knots on wide and narrow faces according to this definition was recorded. Considering the fact that the largest knot could be located close to the end of the board, it may happen that this knot cannot technically be loaded in the bending test. As a result, for wide faces on one hand and narrow faces on the other hand, two

types of largest knots were recorded: the largest knots in the whole board (of size *knotSizeWideFaceWhole\_mm* and *knotSizeNarrowFaceWhole\_mm*), and the largest knots in the loadable area (of size *knotSizeWideFaceTested\_mm* and *knotSizeNarrowFaceTested\_mm*), defined as the central part of the boards excluding 101 cm to the ends (chosen as seven times the width of the board as regard to the geometry of the bending test).

- *Knot type* is defined as the way the knot looks. When knots are grouped, NF B52-001-1 (2018) allow to consider them as a single weak point and it provides rules to measure the corresponding knot size. The knot type “single” or “group” was recorded as “S” or “G”, respectively. Knots can also be observed as (0) healthy or sound, (1) non-adhesive or unsound, (2) leaving a hole, (3) encased, (4) stopped by pruning, and (5) decayed. As an example, a “S0” knot is a single sound knot and “G3” refers to a group of encased knot. For each *knotSize* variable corresponds a *knotType* variable.
- *The weakest loadable zone* was determined depending on the size of the largest knots on the wide and narrow faces, in the loadable area of the boards. NF B52-001-1 (2018) provides different rules for wide and narrow faces knots; thus, the most restrictive case was used to determine the weakest zone in the loadable area of the board. The location of this weakest loadable zone was recorded as the distance from the reference end of the board, named *weakPointx\_m* (see Fig. 12).
- *Other* visual observations have been recorded in the variable *observations*, for example cracks, large resin pockets, wane, bark inclusion, and global deformation.

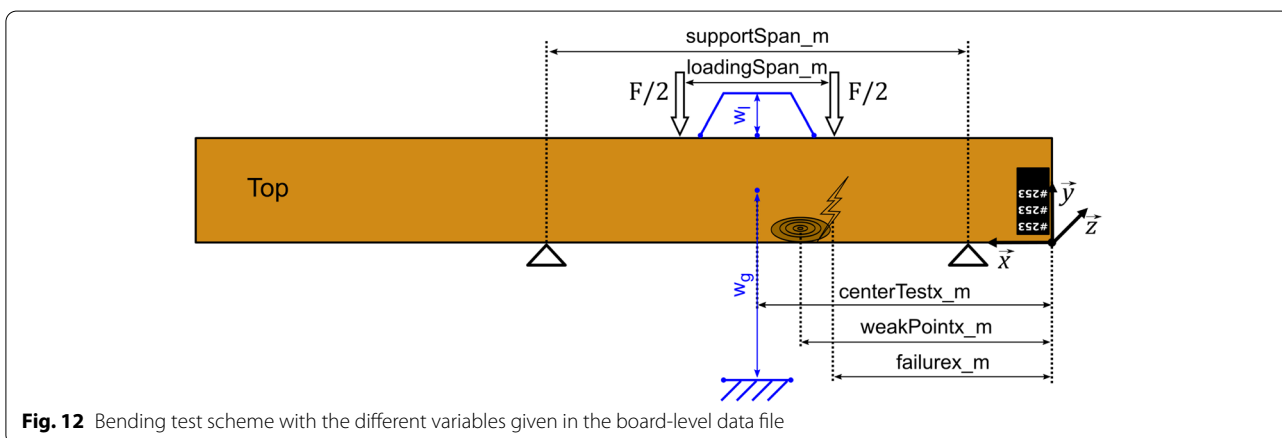
According to these visual measurements and the rules of NF B52-001-1 (2018), each board was graded in

ST-I, ST-II, or ST-III visual grades, or rejected. The corresponding value in the dataset are 1, 2, 3, and 4 for *gradeNFB52001Whole* and *gradeNFB52001Tested* variables, which imply the same distinction as before between the grade of the part of the board that can be loaded in the bending test and the whole board. ST-I corresponds to C30 class, ST-II to C24 class, and ST-III to C18 class of EN 338 (2016) standard.

### 2.4.7 Bending tests

All boards were subjected to four-point edgewise bending tests according to EN 408 (2012). In accordance with this standard, the maximal descent speed of the loading heads was set to 0.003 times the board’s height in the test (that is  $w$ ) per second, which gives  $26.1 \text{ mm min}^{-1}$ . The bending test geometry is defined in the board coordinate system as shown in Fig. 12. The support span *supportSpan\_m* was set to 18 times the nominal width of the board of 0.145 m, which gives 2.610 m, while the distance between the two loading heads *loadingSpan\_m* was set to 6 times the nominal width of the board, which gives 0.87 m. Four Teflon pads were installed on the contact surfaces of the loading heads and bottom supports in order to minimize the static friction with the board permitting the pure bending stress to be applied in between the loading heads.

A 200 kN load cell was used to measure the force  $F$  applied to the loading heads. The maximum load in N is recorded as *maxLoadStatEWBend\_N* (symbol:  $F_{max}$ ). A rope line displacement sensor was positioned at center of the span to measure what is called the global deflection  $w_g$ . This deflection was measured thanks to a hook that was attached at the board’s neutral axis (see Fig. 12). A linear displacement sensor was used to measure the local deflection between the loading heads  $w_l$  on a distance of 5 times the nominal width of the board, which is 0.725 m (see Fig. 12).  $F$ ,  $w_l$  and  $w_g$  were recorded with time in a text file



**Fig. 12** Bending test scheme with the different variables given in the board-level data file

stored in the data repository `TreeTrace_Douglas/boards/grading/static_tests`. The text file `TreeTrace_Douglas/boards/grading/static_tests/readme.txt` give some details about those data.

Following EN 384 (2016) standard, the expected weakest point was positioned between the loading heads, where the board's bending moment was constant and maximum so that the board's bending strength was actually measured. However, the boards were not always positioned exactly such as the distance from the board reference end to the center of the test, denoted `centerTestx_m`, was equal to `weakPointx_m`. Indeed, in some cases, for example if the weakest point was close to the end of the board or if there were two very similar weak points, it was chosen to position the board in a manner so that all the weak points were subjected to the maximum bending. Even by doing so, the failure did not always occurred between the loading heads. A variable named `bendingTestSuccess` indicates if the bending test succeeded, i.e., if the failure occurred between the loading heads and if all the data was successfully recorded (sensor recording failed a couple of times). The distance to the reference end of the observable failure, `failurex_m`, was also recorded. Right after the test, a picture of the failure was taken for each board. These pictures are stored in the data repository `TreeTrace_Douglas/boards/grading/static_tests`. Note that pictures are missing for boards # 92, 247 and 266.

The local MoE `localStatEWBendMoE_Pa` (symbol:  $E_{m,l}$ ) and the global MoE `globalStatEWBendMoE_Pa` (symbol:  $E_{m,g}$ ) were computed for each board following the formulas of EN 408 (2012) standard, i.e., as:

$$localStatEWBendMoE_Pa = E_{m,l} = \frac{a \times 0.725^2}{16I_{G_z}} \cdot \frac{F_2 - F_1}{w_l(F_2) - w_l(F_1)} \quad (9)$$

and

$$globalStatEWBendMoE_Pa = E_{m,g} = \frac{3al^2 - 4a^3}{4tw^3 \cdot \frac{w_g(F_2) - w_g(F_1)}{F_2 - F_1}} \quad (10)$$

with :

$$l = supportSpan\_m$$

$$a = \frac{supportSpan\_m - loadingSpan\_m}{2}$$

and where `startLoadMoE_N` (symbol:  $F_2$ ) and `endLoadMoE_N` (symbol:  $F_1$ ) are respectively the start and end load values in between the MoE were computing, corresponding most of the time to  $0.1 \times F_{max}$  and  $0.4 \times F_{max}$  respectively, unless if the coefficient of determination between the load curve and the linear fit was greater than 0.99; in these very rare cases, the load interval has been reduced.

Each board strength in four-point edgewise bending ( $f_m$ , `strenghtEWBend_Pa`) was computed as:

$$strenghtEWBend\_Pa = f_m = \frac{3aF_{max}}{tw^2} \quad (11)$$

#### 2.4.8 Moisture content and clear wood density

Moisture content was measured for each board by different means and at different times of the experiment.

At first, the moisture content was measured non-destructively by a Testo 606-2 pin-type moisture meter (resolution 0.1 %, precision 1 %) at the time of the visual grading and vibratory tests. It was measured at the location of the weakest zone (`mcPinNDTestWeakPoint_pc`) and at the reference end of the board (`mcPinNDTestBoardEnd_pc`).

Then, it was measured with the same device at the time of the destructive bending test (these tests lasted several weeks), but only at the location of the weakest zone which is recorded as the variable `mcPinBendingTestWeakPoint_pc`.

To determine precisely the moisture content at the time of the bending test, the EN 13183-1 (2002) standard was applied. A small clear wood sample of the full width of the boards (whenever possible) was cut in the nearest zone possible to the failure zone, and its mass was measured right after the bending test with a KERN EMS 300-3 electronic scale (resolution 0.001 g, precision 0.002 g), and recorded as `massCWBendingTest_g` (symbol:  $m_{CW}^{test}$ ). Then, these samples were oven dried for at least 90 h at 103 °C and weighed again with the same scale to obtain what is supposed to be the mass of the sample at 0% moisture content, which was recorded as `massCWOvenDried_g` (symbol:  $m_{CW}^{dry}$ ) (no data available for boards #1, 150, 202, 296, and 305 because of lost/broken samples during the experiment). The moisture content of clear wood in the failure zone at the time of the bending test, `mcCWBendingTest_pc` (symbol:  $u_{CW}^{test}$ ), was thus computed as:

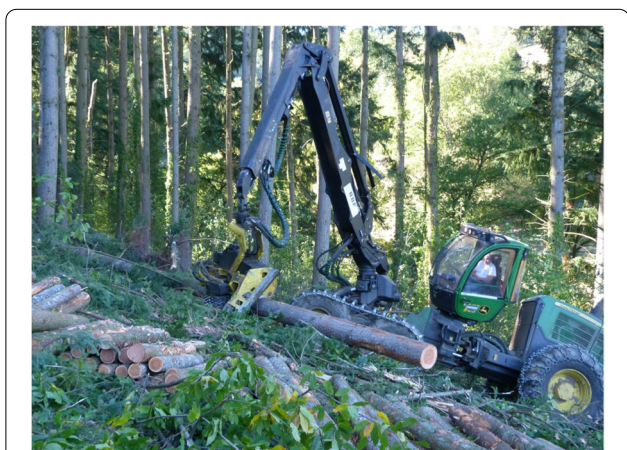
$$mcCWBendingTest\_pc = u_{CW}^{test} = \frac{(m_{CW}^{test} - m_{CW}^{dry})}{m_{CW}^{dry}} \times 100 \quad (12)$$

The same clear wood sample was then used to determine clear wood density ( $\rho_{CW}$ , `cwDensity12pc_kgpm3`), apart for boards #1, 150, 202, 296, and 305 for which a second sample has been used. Since the sample shapes were not perfectly squared, their volume was determined by drowning the samples into water and measuring the resulting buoyant force. First, the mass of each sample at the time of this measurement, `massCWDensity_g`

(symbol:  $m_{CW}^{\rho}$ ), was measured; then, it was drowned into a container of demineralized water with a device maintaining it underwater. The container being positioned on a scale set to zero, the scale gives a mass corresponding to the buoyant force;  $m_b$ . It is used to compute the clear wood density at 12% moisture content in the failure zone as:

$$cwDensity12pc\_kgpm3 = \rho_{CW} = \frac{m_{CW}^{\rho}}{m_b/\rho_w} \times (1 - 0.005 \times (mcCWDensity\_pc - 12)) \tag{13}$$

where  $\rho_w$  is the density of the demineralized water taken equal to  $1000 \text{ kg m}^{-3}$  and  $mcCWDensity\_pc$  is the moisture content of the small clear wood sample at the moment of density measurement. The correction applied to take account of this moisture content comes from EN 384 (2016) standard.



**Fig. 13** Douglas fir trees harvested in the Besle forest

### 3 Methods used for the preliminary sampling

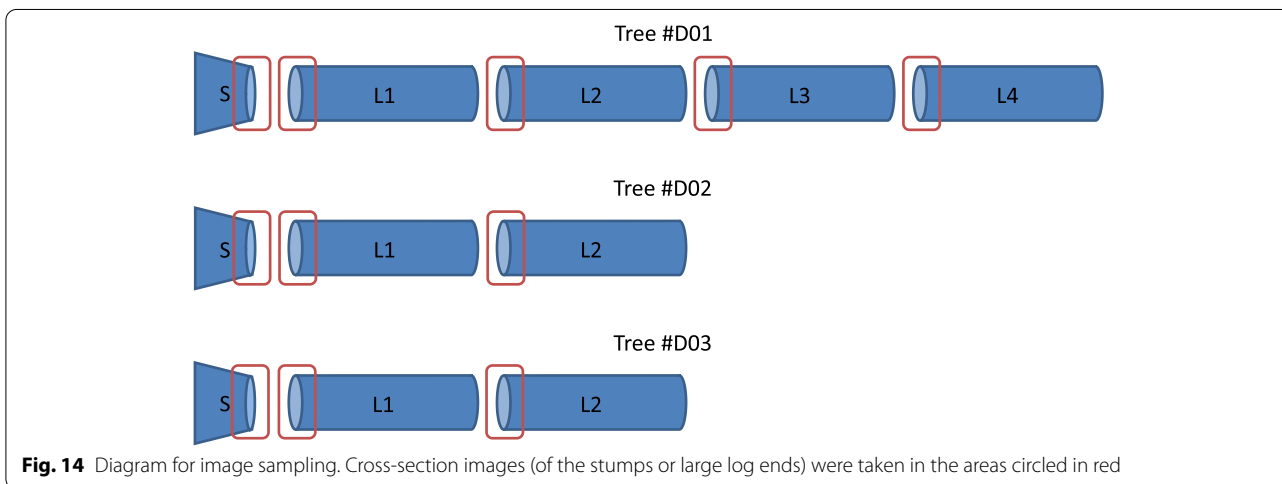
This dataset was collected at the beginning of the project to develop the protocol for the main sampling. Three trees were harvested in the Besle forest (Fig. 13) in October 2018. Several logs were cut in each tree by the harvester (four logs for tree #D01; two logs for tree #D02; two logs for tree #D03; Fig. 14). Images were acquired in

the field with a Panasonic FZ45 Lumix camera. For each tree, several images of the stump were taken. Then, several images of the large log ends were done. For example, *D01-L4-3.jpg* refers to the third image of log #4 (large end) from tree #1. *D01-S-2.jpg* refers to the second image of the stump of tree #1. And *D01-L4-label.jpg* refers to the image of the large end of log #4 after stapling the label. In total, 57 RGB images are available in this “forest” sample (+ eight images with labels).

The logs were then transported to the sawmill several days later, where other images of the same eight log ends were taken with a Huawei smartphone by an employee of the sawmill after removing the labels stapled in the field. In total, 32 RGB images are available in this “logyard” sample (+ eight images with labels).

One aim of this preliminary sampling was to test the feasibility of the traceability of logs from the forest to the sawmill. Figure 15 presents the tree stump and corresponding log-end images taken in the forest and at the sawmill.

The pith coordinates were manually recorded on all available images. Some images were then manually segmented with ImageJ software to identify wood (i.e., the limit between wood and bark) and heartwood areas



**Fig. 14** Diagram for image sampling. Cross-section images (of the stumps or large log ends) were taken in the areas circled in red





**Fig. 15** Example of images of log #1 from tree #2 with: one stump image on the left; one corresponding image of the large log-end taken in the forest; one image of the same log-end taken on the logyard at the sawmill several weeks later. The 25-cm-long scale was used to calibrate each image

(Fig. 16): 11 images for the “forest” sample (one for each of the eight log ends + 3 stumps) and 8 images for the “logyard” sample.

All images (taken in the forest, on the logyard and segmented) are available on the data repository in *TreeTrace\_Douglas/logs\_besle*. All the *Besle data*, among others, pith coordinates (*pithX\_pixel* and *pithY\_pixel*), wood cross-section (*areaLIBForest\_cm2* and *areaLIBLogyard\_cm2*), and heartwood (*areaHWForest\_cm2* and *areaHWLogyard\_cm2*) areas, are available on the data repository in the directory *TreeTrace\_Douglas/tables*.

#### 4 Access to the data and metadata description

The images and the database (Longuetaud et al. 2022) are available at Recherche Data Gouv repository: <https://doi.org/10.15454/YUNEGL>. Associated metadata access is at <https://metadata-afs.nancy.inra.fr/geonetwork/srv/fre/catalog.search#/metadata/d9eef6e4-f195-41f4-b6c2-2ab46adc637e>.

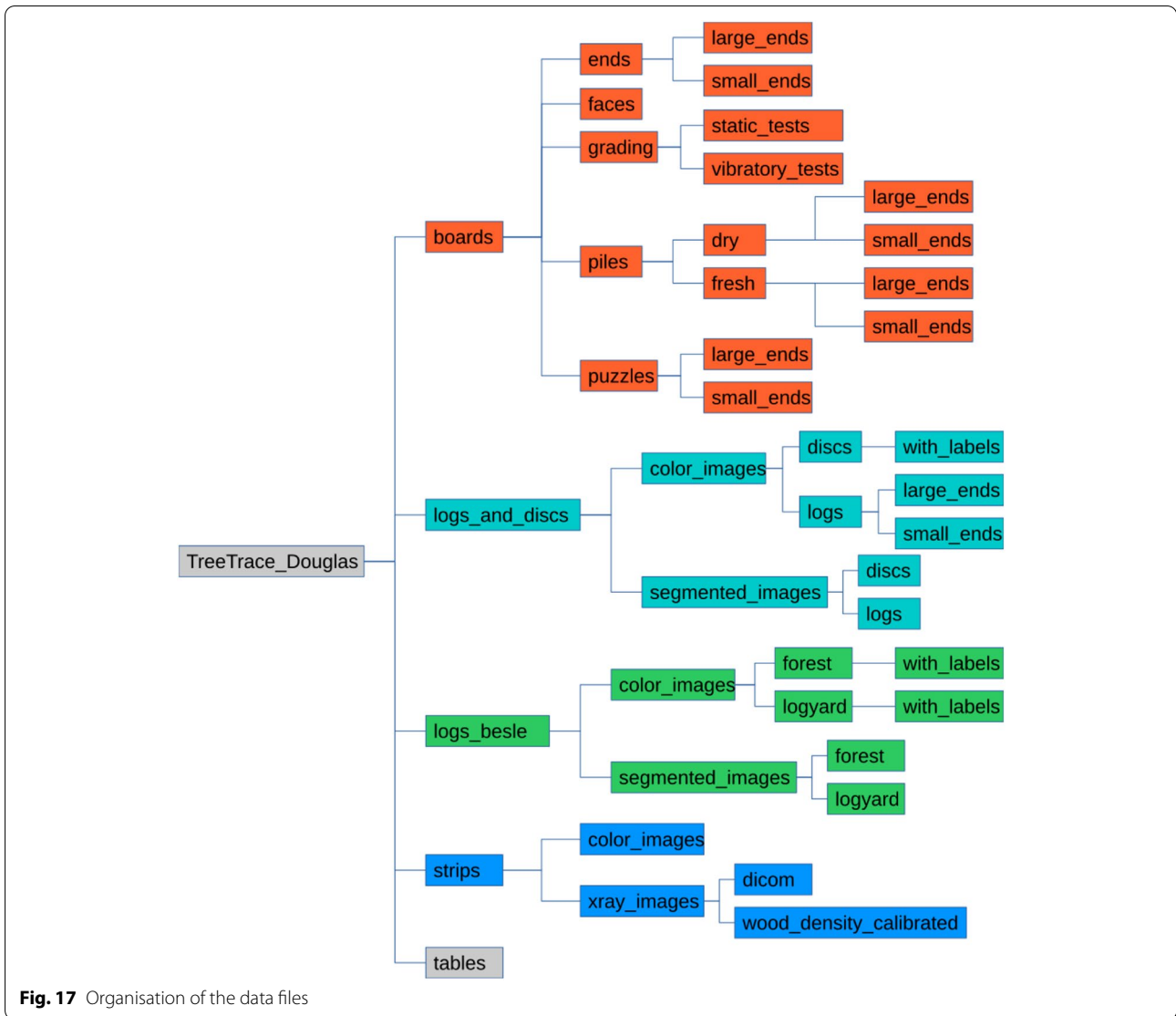
The arborescence of the *TreeTrace\_Douglas* directory was detailed in the previous section and is summarized in Fig. 17. In practice, due to limitations of the downloading software, the data files are available as compressed files, one for each top-level directory: *boards.zip*, *logs\_and\_discs.zip*, *logs\_besles.zip*, *strips.zip* and *tables.zip*.

The main numeric data are in the *TreeTrace\_Douglas/tables* directory. There are 10 tables, all in the form of an array with variables in columns and a first line indicating the variable names. Each data table is accompanied with a variable table (with the same name followed by *\_variable*) describing each variable. For the main sampling, the data tables are the following with the different levels of measurements:

- The trees table contains 52 lines, one for each long log, and 10 columns. Most variables were interpolated at breast height from disc measurements to provide a description of the trees.
- The logs table contains 156 lines, one for each short log, and 17 columns. Some of the measurements were performed on the logyard. Others were performed on images of the log ends taken on the sawmill line. And some other variables were computed from measurements made on images or discs from both ends of the logs.
- The discs table contains 208 lines, one for each disc, and 14 columns. The measurements were performed directly on the discs at the lab or on images of the log ends taken on the logyard.
- The strips table contains 832 lines, one for each radial strip, and 13 columns. The variables were computed



**Fig. 16** Raw RGB image of log #1 from tree #1 taken in the forest (on the left) and the corresponding manually segmented image (on the right). On the segmented image: The white area is the sapwood located between heartwood and bark; the red area is the heartwood



from ring widths measured on RGB images and density profiles issued from CT images.

- The `strips_density_profiles` table gives the raw air-dried density values with one line of data per pixel along the profiles.
- The `strips_rings` table contains 37,856 lines, one for each annual growth ring of each strip, and 8 columns. The variables were computed from ring widths measured on RGB images and density profiles issued from CT images.
- The `rings` table contains 9464 lines, one for each annual growth ring, and 8 columns. They are almost the same variables as in `strips_rings` but averaged over the values of the four orthogonal strips.

- The `boards` table contains 346 lines, one for each board collected, and 83 columns. The measurements include the mechanical assessment of the boards and the location of board end images in log end images.

Data tables are also available for the preliminary sampling of Besle:

- The `logs_besle_forest` table contains variables measured manually on log end images taken in the forest just after harvesting.
- The `logs_besle_logyard` table contains the same variables but measured on images of the same log ends taken on the logyard after transportation to the sawmill.

## 5 Technical validation

All measurements were carefully checked by graphical analyses with R software. The ring measurements were controlled by using a binocular magnifying glass when in doubt, and by checking for consistency between ring profiles from the four strips per disc and the four discs per tree. The X-ray CT scanner used to measure wood density is controlled and calibrated by the manufacturer three times a year. A specific calibration procedure (Freyburger et al. 2009) was used to convert Hounsfield Units to wood density. Scales, load cells, and displacement sensors used for board testing have been controlled and calibrated by the manufacturer no more than 3 years before the tests. Board scanner resolutions were measured by using a ruler glued on the faces of a given board and counting the pixels.

## 6 Reuse potential and limits

All data tables are in plain text format with tabulations as separator. They can be imported easily with R or with any spreadsheet software. The RGB images are in standard JPEG or PNG formats. A scale (black and white checkerboard) included in each image allows to perform spatial calibration except for the images taken on the sawmill line and board face images where another system was used. In all cases, pixel size ( $\text{mm pixel}^{-1}$ ) or image resolution ( $\text{pixel mm}^{-1}$ ) is given for each image.

The CT images (original images in DICOM format and calibrated images in 16-bits TIFF format) can be read with ImageJ. Sometimes information is missing for some variables, which is translated in “NA” value in the tables.

The *TreeTrace\_Douglas* database is suitable for the development of image processing algorithms for untreated log cross-sections with two main purposes: traceability of logs from the forest to the sawmill and during their processing up to the boards; wood quality assessment directly in the forest, on the logyard or on the sawmill line before sawing. Both were objectives of the TreeTrace project.

We give below some examples of past uses of the database.

Pith detection algorithms were developed (Decelle et al. 2019, 2021a, c) because they often serve as a basis for the detection of other characteristics and because pith position provides information on the severity of the eccentricity potentially related to the presence of compression wood for softwood species.

A neural network algorithm to automatically segment the wood cross-section in the entire image was also developed (Decelle and Jalilian 2020). This work has also led to the development of original algorithms in the field of mathematical morphology (Decelle et al. 2021b).

Several Master students worked in the project: for automatic reading of the black and white calibration scale (Delconte 2019), for automatic repositioning of individual board faces in the log cross-section (Barthélémy 2020), for board machine grading (Ge 2020), and board destructive testing (Mancini 2021; Guillaume 2021).

Moreover, some of the studied logs were peeled into veneers that were used to study and manufacture laminated veneer lumbers (Duriot et al. 2021a, b); thus, a comparison with the mechanical properties of the boards of the present dataset may be performed.

In addition to the past uses, further works are in progress for segmenting heartwood area and measuring ring width. Such algorithms can be developed and validated with this database providing raw images and ground truth for each characteristics (pith location, cross-section area of wood, heartwood area, annual ring widths, board coordinates).

The high resolution RGB images of strips (four orthogonal strips per disc) could be used to develop new image processing algorithms for applications in dendrochronology. These images are provided with the corresponding annual growth ring readings.

The data can also be used to analyze some wood quality characteristics of Douglas fir and to estimate the amount of information that can be recovered from the analysis of log ends. For instance, a study is underway to analyze the relationship between tree growth and wood density at the ring level from the database. Indeed, an automatic detection of ring widths associated with a ring width – ring density model would provide relevant quality information.

It would also be possible to link the properties of the logs (ring widths, wood density, juvenile wood area, heartwood and sapwood amounts, pith eccentricity) to the mechanical performances of the boards obtained from these logs.

Another possible application is the development of strength grading algorithms such as that of Viguier et al. (2017), Olsson et al. (2018), and Rais et al. (2021) since the dataset includes the main measurements made by current industrial board grading machines. The location of the boards in the logs can also be used in the models like that of Lukacevic et al. (2019) and Hu et al. (2018).

The main difficulty in such studies performed in an industrial context is that very few information is available on the sites and trees sampled since the logs were selected on the logyard.

### Acknowledgements

Many thanks to the BBF sawmill for helping us with the sampling on the logyard and on the sawmill line, and with the sawing process to collect the boards. We also thank Fakopp for helping us with the frequency measurements with the Log Grader. We also would like to thank all the people involved in the measurements and especially Daniel Rittié (ring width

measurements); Charline Freyburger and the SilvaTech team (CT scanning); Anaëlle Rajaoarivo, Lou Bissieux, and Noah Farré (manual segmentations, board puzzles, etc.); and Fabrice Cottin and Roger Letourneau (board tests). Thanks a lot to Damien Maurice and Alain Benard for helping us upload the data to the data repository "Data INRAE." Thanks to our Austrian colleagues, Rudolf Schraml, Karl Entacher, Alexander Petutschnigg, and Andreas Uhl, without whom this project would not have been possible. An equivalent article on the common database developed during the TreeTrace project, this time on Norway spruce, is in preparation.

#### Authors' contributions

FL initiated the funding request; managed the project for the French part; participated in the data collection, data analysis, and writing of the manuscript; and supervised several Master students. GP participated in the data collection, data analysis, and writing of the manuscript and supervised several Master students. FM initiated the funding request; participated in the data collection, data analysis, and writing of the manuscript; and supervised several Master students. AB worked on these data for the location of board ends in log ends during his Master thesis. RD did his PhD thesis on the basis of these data and participated in the data collection and data analysis. FD worked on these data for the automatic detection of the black and white scale during his Master thesis. XG participated in the data collection and data analysis. GG participated in the data collection and data analysis. TM participated in the data collection and data analysis. TR worked on these data to attribute a wood density to each individual growth ring by developing a new approach during his Master thesis. JCB participated in the data collection and data analysis. RC participated in the data collection, especially for the collaboration with the sawmill; data analysis; and writing the manuscript. IDR supervised the PhD student and several Master students and participated to the data analysis. BM participated in the data collection, data analysis, and writing the manuscript. PN supervised the PhD student and several Master students and participated in the data analysis. BR participated in the data collection, especially with the installation of the camera at the sawmill, and data analysis. JV participated in the data collection, data analysis, and writing the manuscript and supervised several Master students. All authors read and approved the final manuscript.

#### Funding

SilvaTech facility is supported by the French National Research Agency through the Laboratory of Excellence ARBRE (ANR-11-LABX-0002-01). This research was made possible thanks to the financial support of the French National Research Agency (ANR) in the framework of the TreeTrace project, ANR-17-CE10-0016.

#### Availability of data and materials

The images and the database are available at Research Data Gouv repository: <https://doi.org/10.15454/YUNEG>. Associated metadata are available at <https://metadata-afs.nancy.inra.fr/geonetwork/srv/fre/catalog.search#/metadata/d9eef6e4-f195-41f4-b6c2-2ab46adc637e>.

#### Code availability

Not applicable.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Université de Lorraine, AgroParisTech, INRAE, Silva, 54000 Nancy, France. <sup>2</sup>Arts et Métiers Institute of Technology, LaBoMaP, Université Bourgogne Franche-Comté, HESAM Université, F-71250 Cluny, France. <sup>3</sup>Université de Lorraine, CNRS, LORIA, F-54000 Nancy, France. <sup>4</sup>Université de Lorraine, AgroParisTech, INRAE, Silva, 54000 Nancy, France.

Received: 4 May 2022 Accepted: 6 October 2022

Published online: 28 November 2022

#### References

- Barthélémy A (2020) Reconstruction de billon. Master's thesis, M2 informatique, Parcours type Apprentissage, Vision, Robotique (AVR). Université de Lorraine, Nancy. <https://hal.univ-lorraine.fr/hal-03606199>
- Besseau B, Pot G, Collet R, Viguier J (2020) Influence of wood anatomy on fiber orientation measurement obtained by laser scanning on five European species. *J Wood Sci* 66(1):74. <https://doi.org/10.1186/s10086-020-01922-y>
- Brancheriau L, Bailleres H (2002) Natural vibration analysis of clear wooden beams: a theoretical review. *Wood Sci Technol* 36(4):347–365. <https://doi.org/10.1007/s00226-002-0143-7>
- Decelle R, Jalilian E (2020) Neural networks for cross-section segmentation in raw images of log ends. In: 4th International Conference on Image Processing, Applications and Systems (IPAS). IEEE, p 131–137. <https://doi.org/10.1109/IPAS50080.2020.9334960>
- Decelle R, Ngo P, Debled-Rennesson I, Mothe F, Longuetaud F (2019) A new algorithm to automatically detect the pith on rough log-end images. In: 21st International Nondestructive Testing and Evaluation (NDTE) of Wood Symposium. Freiburg. <https://hal.inria.fr/hal-02275651>
- Decelle R, Ngo P, Debled-Rennesson I, Mothe F, Longuetaud F (2021a) Ant colony optimization for estimating pith position on images of tree log ends. *Image Process Line*. <https://hal.archives-ouvertes.fr/hal-03442289>
- Decelle R, Ngo P, Debled-Rennesson I, Mothe F, Longuetaud F (2021b) Digital straight segment filter for geometric description. In: International Conference on Discrete Geometry and Mathematical Morphology. Springer, p 255–268. [https://doi.org/10.1007/978-3-030-76657-3\\_18](https://doi.org/10.1007/978-3-030-76657-3_18)
- Decelle R, Ngo P, Debled-Rennesson I, Mothe F, Longuetaud F (2021c) Pith estimation on tree log end images. In: International Workshop on Reproducible Research in Pattern Recognition. Milan, p 101–120. [https://doi.org/10.1007/978-3-030-76423-4\\_7](https://doi.org/10.1007/978-3-030-76423-4_7)
- Delconte F (2019) Estimation de la qualité de grumes de bois. Master's thesis, M2 informatique, Parcours type Apprentissage, Vision, Robotique (AVR). Université de Lorraine, Nancy. <https://hal.univ-lorraine.fr/hal-03603108>
- Divos G (2014) RESONANCE LOG GRADER - for log grading - Android software - User's guide
- Duriot R, Pot G, Girardon S, Denaud L (2021a) New perspectives for LVL manufacturing from wood of heterogeneous quality-Part 2: Modeling and manufacturing of variable stiffness beams. *Forests* 12. <https://doi.org/10.3390/f12091275>
- Duriot R, Pot G, Girardon S, Roux B, Marcon B, Viguier J, Denaud L (2021b) New perspectives for LVL manufacturing from wood of heterogeneous quality-Part 1: Veneer mechanical grading based on online local wood fiber orientation measurement. *Forests* 12. <https://doi.org/10.3390/f12091264>
- EN 13183-1 (2002) Moisture content of a piece of sawn timber – Part 1: Determination by oven dry method. <https://www.boutique.afnor.org>
- EN 338 (2016) Structural timber – strength classes. <https://www.boutique.afnor.org>
- EN 384 (2016) Structural timber – determination of characteristic values of mechanical properties and density. <https://www.boutique.afnor.org>
- EN 408 (2012) Timber structures – structural timber and glued laminated timber – determination of some physical and mechanical properties. <https://www.boutique.afnor.org>
- European Commission (2017) Strategy for forest-based industries
- Freyburger C, Longuetaud F, Mothe F, Constant T, Leban JM (2009) Measuring wood density by means of X-ray computer tomography. *Ann For Sci* 66(8):804. <https://doi.org/10.1051/forest/2009071>
- Ge X (2020) Etude de la qualité (au sens des propriétés mécaniques) de sciages en relation avec des informations d'arbre. Master's thesis, M1 Ecole Nationale Supérieure Paris-Scalpy
- Guillaume G (2021) Contribution à l'évaluation des propriétés mécaniques longitudinales, transversales et au classement du Douglas français (ANR TreeTrace). Master's thesis, M2 Génie Mécanique et Matériaux, Université Bretagne Sud
- Habite T, Abdeljaber O, Olsson A (2021). Automatic detection of annual rings and pith location along norway spruce timber boards using conditional adversarial networks. <https://doi.org/10.1007/s00226-021-01266-w>



- Habite T, Abdeljaber O, Olsson A (2022) Determination of pith location along norway spruce timber boards using one dimensional convolutional neural networks trained on virtual timber boards. 329(127):129. <https://doi.org/10.1016/j.conbuildmat.2022.127129>
- Hu M, Olsson A, Johansson M, Oscarsson J (2018) Modelling local bending stiffness based on fibre orientation in sawn timber. 76(6):1605–1621. <https://doi.org/10.1007/s00107-018-1348-2>
- Kurdthongmee W (2020) A comparative study of the effectiveness of using popular DNN object detection algorithms for pith detection in cross-sectional images of parawood. *Heliyon* 6(2). <https://doi.org/10.1016/j.heliyon.2020.e03480>
- Kurdthongmee W, Suwannarat K, Panyuen P, Sae-Ma N (2018) A fast algorithm to approximate the pith location of rubberwood timber from a normal camera image. In: 15th International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE, p 1–6. <https://doi.org/10.1109/JCSSE.2018.8457375>
- Longuetaud F, Pot G, Mothe F, Barthelemy A, Decelle R, Delconte F, Ge X, Guillaume G, Mancini T, Ravoajanahary T, Butaud JC, Collet R, Debled-Rennesson I, Marcon B, Ngo P, Roux B, Viguier J (2022) "TreeTrace\_Douglas". [dataset], V1. Recherche Data Gouv. <https://doi.org/10.15454/YUNEG>
- Lukacevic M, Kandler G, Hu M, Olsson A, Füssl J (2019) A 3d model for knots and related fiber deviations in sawn timber for prediction of mechanical properties of boards. 166:107617. <https://doi.org/10.1016/j.matdes.2019.107617>
- Mancini T (2021) Effets de la traçabilité des arbres pour le classement mécanique des bois dans la scierie du futur (ANR TreeTrace). Master's thesis, Arts et Métiers Sciences et Technologies
- NF B52-001-1 (2018) Règles d'utilisation du bois dans la construction - Classement visuel pour l'emploi en structures des bois sciés résineux et feuillus - Partie 1 : bois massif
- Norell K (2011) Automatic counting of annual rings on *Pinus sylvestris* end faces in sawmill industry. *Comput Electron Agric* 75(2):231–237. <https://doi.org/10.1016/j.compag.2010.11.005>
- Norell K, Borgefors G (2008) Estimation of pith position in untreated log ends in sawmill environments. *Comput Electron Agric* 63(2):155–167. <https://doi.org/j.compag.2008.02.006>
- Nyström J (2003) Automatic measurement of fiber orientation in softwoods by using the tracheid effect. *Comput Electron Agric* 41(1–3):91–99. [https://doi.org/10.1016/S0168-1699\(03\)00045-0](https://doi.org/10.1016/S0168-1699(03)00045-0)
- Olsson A, Pot G, Viguier J, Faydi Y, Oscarsson J (2018) Performance of strength grading methods based on fibre orientation and axial resonance frequency applied to norway spruce (*Pinus abies* L.), douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco) and european oak (*Quercus petraea* (Matt.) Liebl./*Quercus robur* L.) 75(4). <https://doi.org/10.1007/s13595-018-0781-z>
- Paradis S, Brancheriau L, Bailleres H (2017) Bing: Beam identification by non destructive grading. <https://doi.org/10.18167/62696e67>
- Raatevaara A, Korpunen H, Tiitta M, Tomppo L, Kulju S, Antikainen J, Uusitalo J (2020) Electrical impedance and image analysis methods in detecting and measuring scots pine heartwood from a log end during tree harvesting. *Comput Electron Agric* 177:105690. <https://doi.org/10.1016/j.compag.2020.105690>
- Rais A, Bacher M, Khaloian-Sarnaghi A, Zeilhofer M, Kovryga A, Fontanini F, Hilmers T, Westermayr M, Jacobs M, Pretzsch H, van de Kuilen JW (2021) Local 3d fibre orientation for tensile strength prediction of european beech timber. 279:122527. <https://doi.org/10.1016/j.conbuildmat.2021.122527>
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675. <https://doi.org/10.1038/nmeth.2089>
- Schraml R, Hofbauer H, Petutschnigg A, Uhl A (2015) Tree log identification based on digital cross-section images of log ends using fingerprint and iris recognition methods. In: International Conference on Computer Analysis of Images and Patterns. Springer, p 752–765. [https://doi.org/10.1007/978-3-319-23192-1\\_63](https://doi.org/10.1007/978-3-319-23192-1_63)
- Schraml R, Uhl A (2013) Pith estimation on rough log end images using local fourier spectrum analysis. In: Proceedings of the 14th Conference on Computer Graphics and Imaging (CGIM'13), Innsbruck, AUT. <https://doi.org/10.2316/P.2013.797-012>
- Schraml R, Uhl A (2014) Similarity based cross-section segmentation in rough log end images. In: IFIP International Conference on Artificial Intelligence Applications and Innovations. Springer, p 614–623. [https://doi.org/10.1007/978-3-662-44654-6\\_61](https://doi.org/10.1007/978-3-662-44654-6_61)
- Simonaho SP, Palviainen J, Tolonen Y, Silvennoinen R (2004) Determination of wood grain direction from laser light scattering pattern. *Opt Lasers Eng* 41(1):95–103. [https://doi.org/10.1016/S0143-8166\(02\)00144-6](https://doi.org/10.1016/S0143-8166(02)00144-6)
- Viguier J, Bourreau D, Bocquet JF, Pot G, Bléron L, Lanvin JD (2017) Modelling mechanical properties of spruce and douglas fir timber by means of x-ray and grain angle measurements for strength grading purpose. 75(4):527–541. <https://doi.org/10.1007/s00107-016-1149-4>
- Wimmer G, Schraml R, Hofbauer H, Petutschnigg A, Uhl A (2021) Two-stage cnn-based wood log recognition. *CoRR abs/2101.04450*. [https://doi.org/10.1007/978-3-030-87007-2\\_9](https://doi.org/10.1007/978-3-030-87007-2_9)
- WinDendro (2021) WinDENDRO: An image analysis system for annual tree-rings analysis. [https://regentstruments.com/assets/windendro\\_about.html](https://regentstruments.com/assets/windendro_about.html)

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

