Composites & nanocomposites

Hyperspectral imaging and chemometrics reveal wood acetylation on different spatial scales

Mikko Mäkelä[1,2,*] (ID), Michael Altgen[3], Tiina Belt[3,4], and Lauri Rautkari[3]

[1] *VTT Technical Research Centre of Finland, Ltd, PO Box 1000, 02044 VTT Espoo, Finland*

[2] *Department of Forest Biomaterials and Technology, Swedish University of Agricultural Sciences, Skogsmarksgränd, 90183 Umeå, Sweden*

[3] *School of Chemical Engineering, Department of Bioproducts and Biosystems, Aalto University, PO Box 16300, 00076 Aalto, Finland*

[4] *Natural Resources Institute Finland (Luke), Tietotie 2, 02150 Espoo, Finland*

ABSTRACT

Acetylation is a chemical treatment method commonly used to improve the hygroscopic properties of wood. Although acetylation has been industrially used for decades, its effects on the different hierarchical structures of wood are still poorly understood. In the laboratory, acetylation is generally measured gravimetrically. Weighing a sample before and after the modification procedure provides an indirect measure of the degree of acetylation within the entire sample but does not provide detailed information on the different structural regions of wood. Here, we determined acetylation of wood surfaces using hyperspectral near-infrared image regression. Our results show significant differences in the acetylation of earlywood and latewood, which suggests different durations for complete acetylation of earlywood and latewood cells. We have also illustrated our findings on the wood cell level based on the chemical differences in earlywood and latewood cell walls using cluster analysis of Raman images. These findings are an important step in understanding how chemical treatment affects the different hierarchical structures of wood on different spatial scales.

🙋 Springer

GRAPHICAL ABSTRACT
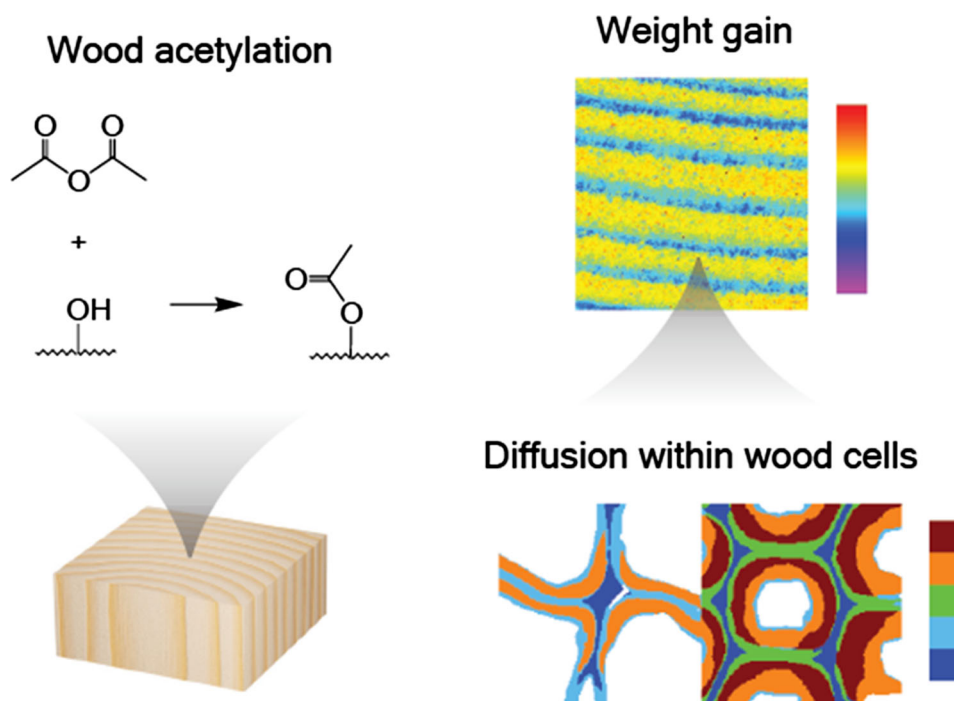
Introduction

Wood structure, which is made up of cellulose fibrils embedded in a matrix of hemicelluloses and lignin, provides a high strength to weight ratio but makes wood vulnerable to hygroscopic changes and dimensional instability. These weaknesses can be compensated by various chemical and thermal treatments routinely used to improve wood properties for different applications. One alternative is wood acetylation, where acetic anhydride is used for replacing accessible hydroxyl groups (–OH) on wood polymers with acetyl groups (–CH$_3$CO) generating acetic acid as a by-product [1]. Acetylation decreases the number of primary sorption sites for water molecules and limits the accessibility of unmodified hydroxyls due to the larger size of the acetyl groups [2, 3]. The degree of wood acetylation is generally measured gravimetrically by weighing a sample before and after the modification procedure. The obtained weight percent gain (WPG) provides an indirect measure of acetylation of the entire sample and is highly correlated with the improvement in wood properties [4, 5]. However, the WPG alone does not provide information on the potential variations in acetylation degree within the bulk sample. Although acetylation has been industrially used for decades, the distributions of acetyl groups within wood still remain poorly understood.

Chemical imaging provides tools to combine the spatial information of an image with the chemical information of a spectrum. Different spectroscopic techniques can be used for acquiring spectral information from a range of spectral variables either from a single pixel, a line of pixels or an entire image scene at a time. Hyperspectral near-infrared (NIR) imaging records diffuse reflectance or transmittance on a continuous range of several hundred different wavelengths in the NIR region of the electromagnetic spectrum [6]. The NIR region provides information on the overtones and combinations of the vibrational modes of polar molecular bonds making it extremely well-suited for analyzing lignocellulosic materials

and their derivates. Modern push broom instruments are fast and can be used with minimal sample preparation due to the lower absorptivity of the overtones and combinations of fundamental vibrational modes from the mid-infrared [7]. However, many of these peaks in NIR spectra overlap and multivariate image analysis based on chemometric techniques is generally required to identify the spatial variation in the chemical properties of a sample. Raman microspectroscopy, on the other hand, operates by mapping Raman spectra generated by the inelastic scattering of monochromatic light from a preselected region on the sample surface one pixel at a time. The theoretical maximum resolution is determined by wavelength of the excitation laser and the numerical aperture of the microscopy objective [8]. Although the spatial resolution achieved in practical applications is often below the theoretical maximum, resolutions in the range of micrometers can easily be achieved [9]. The majority of chemometric techniques for spectral image interpretation are based on dimensionality reduction using latent variables and enable visualizing the important variations in an image dataset in just a few interpretable dimensions. The use of latent variables can also enhance selectivity by taking into account information of a specific analyte from multiple different wavelengths and avoids problems of multicollinearity in calibration and regression if the number of spectral variables exceeds the number of available samples or objects.

Chemical imaging is a promising tool for determining the chemical properties of wood. Hyperspectral NIR images have previously been used for evaluating the chemical composition of wood [10, 11], identifying different wood species [12], predicting wood moisture content and physical properties [13, 14], and visualizing the level of wood extractives [15]. The use of NIR images has also been reported in predicting vapor-phase acetylation of wood [16]. The authors recorded NIR images within 1050–2550 nm and obtained coefficients of determination in the range 0.89–0.90 for predicting the concentration of acetyl groups on wood surfaces. However, although the authors were able to visualize the spatial variation in the concentration of acetyl groups independent of wood moisture content, the results were based on a multiple linear regression model on second derivative absorbance values at 1727 nm [16]. Raman microspectroscopy has also been extensively utilized in studying wood chemistry, with typical applications involving the visualization of cellulose and lignin distributions on a cellular and subcellular level in different types of natural and treated samples [9, 17]. Raman images have also been used to study the distributions of wood extractives [18] and various wood modification agents [19–21].

We have recently shown that hyperspectral NIR imaging and image regression can successfully determine the penetration of acetic anhydride into the wood structure after wood acetylation [22]. Our method was based on calibrating NIR images with external WPG values measured from dried samples after confined one-sided acetylation of solid wood. Now we extend this work by evaluating the spatial differences on wood surfaces after unconfined acetylation followed by wood conditioning under different relative humidities. This approach enables visualizing the acetylation-induced changes on wood surfaces practically independent of equilibrium moisture content. We also show how the most important wavelengths that correspond with wood
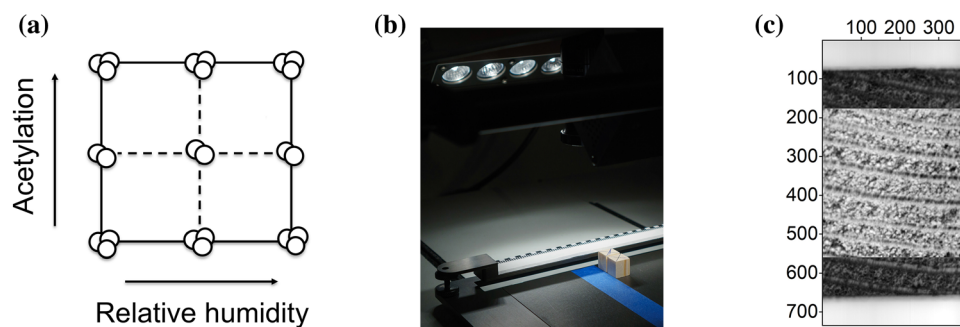
Figure 1 a An illustration of the experimental design, b an image of similar wood samples under illumination of the NIR camera, c and grayscale images of a sample (darker gray) and the selected region of interest, ROI (lighter gray). The values in c denote image pixel dimensions.

acetylation can be systematically determined by selecting subsets of spectral variables based on interval regression. The results show that considerably higher WPG values were found in the early-wood parts of wood surfaces. As the spatial resolution of most hyperspectral NIR cameras is limited to the macrostructural level, we have also illustrated the differences in earlywood and latewood acetylation on the cellular level using unsupervised classification of microscopic Raman images.

Experimental section

Sample preparation

Defect-free samples of dimensions $15 \times 15 \times 15$ mm^3 were cut from kiln-dried Scots pine (*Pinus sylvestris L.*) sapwood and prepared according to an experimental design, where acetylation and sample conditioning were varied on different levels. The initial central composite design included three levels for both variables, which resulted in WPG values in the range 0–18% and sample moisture contents of 0–29%, Fig. 1a. Additional calibration and validation samples were also prepared inside the design range. These samples were not evenly distributed within the design but provided additional variable levels for image calibration. All samples were prepared in duplicate or triplicate, which provided a total of 36 imaging samples distributed across 13 experimental locations.

The samples were first extracted in a Soxhlet apparatus with acetone for 6 h and then oven dried at 103 °C to determine their initial mass. The acetylation was performed using neat acetic anhydride. The samples were first impregnated with the anhydride under vacuum at room temperature and then transferred to a reaction flask, where they were treated with the acetic anhydride for 0, 20, 100, 180 or 360 min at 120 °C under reflux for adjusting acetylation degree. At the end of the treatment, the reaction flask was transferred onto ice to stop the reaction and the anhydride was replaced with acetone. The acetylated samples were again Soxhlet extracted with acetone for 6 h to remove leftover anhydride and reaction by-products and then oven dried at 103 °C to determine their modified mass and the WPG.

The acetylated samples were then conditioned in different relative humidities for determining the effects of acetylation on sample moisture content. Silica gel, salt solutions of calcium chloride, potassium iodide, potassium chloride or pure water were used as reagents for generating relative humidities 0, 33, 70, 85 and > 97%, respectively [23–25]. The acetylated samples were placed on sample holders above the reagents in plastic containers, which were sealed and stored in room temperature 20 °C for approximately 4 weeks.

Hyperspectral NIR imaging

The hyperspectral NIR images were taken from the sawn, transverse surfaces directly after the wood samples had been acetylated and conditioned in different RHs. The imaging setup has been previously been described in ref. [26] and consisted of a Specim SWIR 3 camera (Specim, Spectral Imaging Ltd.) equipped with a 105 mm OLES macro lens provided by Specim. As illustrated in Fig. 1b, two opposite rows of quartz halogen lamps generated polychromatic light and the reflected wavelengths were separated by a grating-prism monochromator followed by a HgCdTe detector array. The images were taken in line-scanning mode where a line of 384 pixels was continuously recorded on different wavelengths. A spectral range of 1000–2500 nm was used with a sampling interval of 5.6 nm, which provided 267 spectral variables. The lens field of view was 10 mm which resulted in a nominal pixel size of 26×26 μm^2 [26]. The images were taken in reflectance mode and converted to absorbance units after correction with Spectralon white reference and dark current intensities. After imaging, the samples were placed back into the sample containers and weighed after all the samples had been imaged.

Raman microspectroscopy

Raman images were obtained from earlywood and latewood sections in a sample with 8.9% WPG and 9.2% moisture content located in the center of the experimental design (Fig. 1a). A cross section with a thickness of approx. 20 μm was cut from a water-saturated block on a sledge microtome (Lab-Microtome, Swiss Federal Research Institute WSL). The cross section was then placed on a microscope slide with a drop of deionized water, covered with a glass coverslip and edge-sealed with nail polish. The Raman images were acquired using a WITec alpha

300 RA Raman microscope equipped with a 532 nm frequency-doubled Nd:YAG laser operated at 30 mW, a 100 × immersion oil objective with a numerical aperture 1.25, and a DU970-BV EMCCD camera behind a 600 lines mm^{-1} grating. The image size was 50×50 μm^2 with 180 lines per image and 180 points per line. The integration time was set to 0.3 s for collecting a Raman spectrum in each image pixel. Separate images were acquired from earlywood and latewood cells within the same cross section.

Image and data analysis

The obtained NIR images were interpreted using exploratory image analysis and hyperspectral image regression. As illustrated in Fig. 1c, square regions of interest (ROIs) of 384 × 384 pixels were selected from the middle of each sample image and median filtered to remove the effects of dead pixels in the camera detector. One sample ROI from each experimental design location was selected and combined into a mosaic of 9 sample ROIs. The mosaic was decomposed using a principal component analysis (PCA) [27, 28] model after the pixel spectra had been preprocessed with standard normal variate (SNV) [29] correction and mean centering and unfolded into a two-way array, Eq. (1):

$$\mathbf{X} = \mathbf{t_1}\mathbf{p_1^T} + \ldots + \mathbf{t_n}\mathbf{p_n^T} + \mathbf{E_n} \tag{1}$$

where $\mathbf{X}$ denoted an unfolded matrix of preprocessed and mean-centered pixel spectra, $\mathbf{t}$ and $\mathbf{p}$ the orthogonal score and orthonormal loading vectors, respectively, and $\mathbf{E_n}$ a residual matrix after n principal components. The principal component score vectors $\mathbf{t}$ were then refolded back to the mosaic dimension to illustrate the results through score images.

Image calibration was performed using partial least squares (PLS) regression [30, 31] on measured WPG values. The model vector $\mathbf{b}$ in the general regression equation, Eq. (2):

$$\mathbf{y} = \mathbf{Zb} + \mathbf{e} \tag{2}$$

where $\mathbf{y}$ denoted a vector of measured and mean-centered WPG values, $\mathbf{Z}$ a matrix of SNV transformed and mean-centered calibration spectra, and $\mathbf{e}$ a vector of model residuals, was determined using the SIMPLS algorithm [32]. The image ROIs were divided into separate calibration and test sets that represented the entire experimental design range and six average

spectra were obtained from each sample ROI for calibration and validation. Root-mean-squared errors (RMSE) of calibration (RMSEC) and prediction of the test set (RMSEP) were calculated based on Eq. (3):

$$\text{RMSE} = \sqrt{\frac{\sum_j^n \left(y_j - \hat{y}_j\right)^2}{n}} \tag{3}$$

where y_j and $\hat{y}_j$ denoted the measured and predicted values, respectively, and n the number of calibration or test objects. In addition, a separate RMSEP parameter was determined based on the pixel populations of the test set ROIs to evaluate image over-fitting [33], Eq. (4):

$$\text{RMSEP}_{im} = \sqrt{\frac{\sum_i^k \sum_p^m \left(y_{pi} - \hat{y}_{pi}\right)^2}{km}} \tag{4}$$

where y_{pi} and $\hat{y}_{pi}$ denoted the measured sample and predicted pixel values, respectively, and m the number of image pixels in k test set images.

The measured Raman images were interpreted using PCA and PCA-based pixel cluster analysis after the earlywood and latewood images were combined into an image mosaic. Wavelengths outside 500–3500 cm^{-1} were excluded to remove unnecessary noise, and the pixel spectra were preprocessed by baseline removal using a second-order polynomial followed by normalization to unit length to correct for intensity differences between the two images. Random spikes were removed with a median filter applied in the image dimension. Principal component scores and loadings were determined as in Eq. (1) after preprocessing, mean-centering and mosaic unfolding. The obtained PCA scores were used for pixel clustering based on a k-means algorithm that determined similarities between individual pixels and the cluster centroids based on correlation (c_{ij}) in the principal component score space, Eq. (5):

$$c_{ij} = \mathbf{t_i} \cdot \mathbf{t_j} = |\mathbf{t_i}||\mathbf{t_j}|cos\theta \tag{5}$$

where $\mathbf{t_i}$ and $\mathbf{t_j}$ denote the score vectors of pixels i and j normalized to unit length after n retained principal components (Eq. 1). As illustrated in Eq. (5), for vectors which are unit length, correlation described the angle and not the distance between the determined pixel scores in principal component space. A given number of cluster centroids was initially chosen based on largest distances from the center of the score space and updated upon the assignment of each
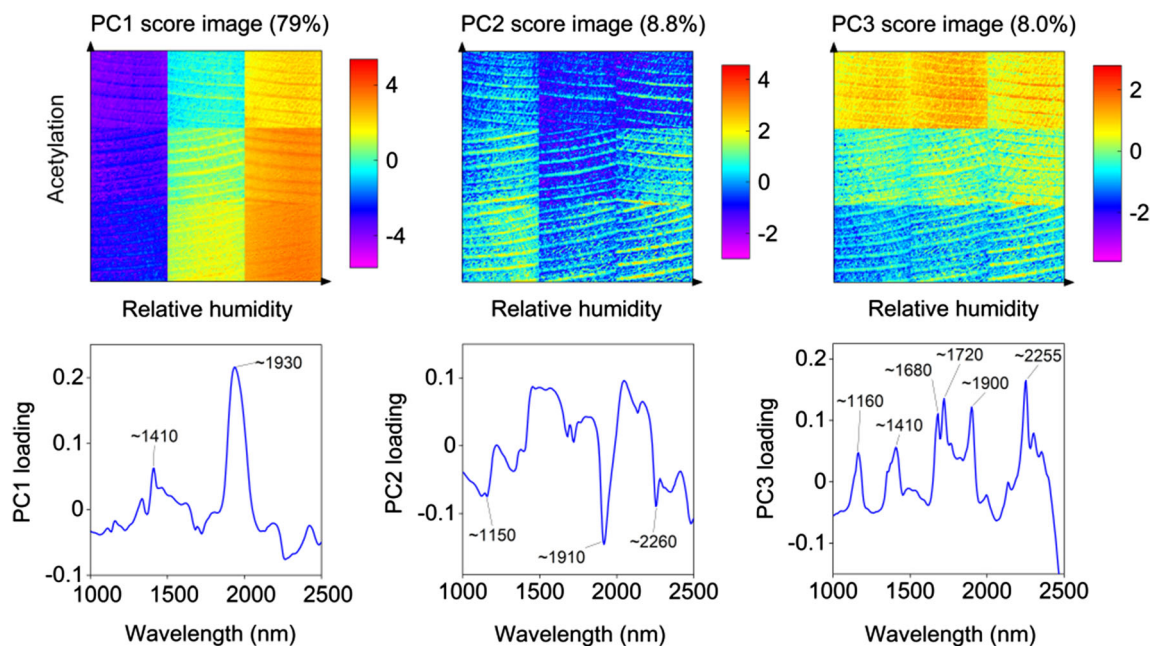
Figure 2 Score images and the respective loading vectors of the first three principal components based on the NIR image mosaic. The color scales denote pixel score values.

pixel. All data analyses were performed with in-house Matlab® (Version 2018b, The MathWorks, Inc.) routines that utilized commercial functions from the PLS (Version 8.7) and MIA toolboxes (Version 3.0.7) from Eigenvector Research, Inc. The initial NIR image segmentation was done in Breeze (Version 2019.02, Prediktera AB), and the final results were plotted in OriginPro (Version 9.6.0.172, OriginLab Corp.).

Results

We acetylated and then conditioned wood samples in different relative humidities based on an experimental to also determine the effects of acetylation and WPG on moisture uptake based on hyperspectral NIR images. The determined PCA model showed that the effects of acetylation and sample conditioning were clearly visible in the NIR image mosaic. The first three principal components explained 96% of the variation in the preprocessed and mean-centered spectra and provided a meaningful chemical interpretation for the differences within the mosaic. The first principal component explained 79% of data variation and mainly separated the samples based on relative humidity during conditioning. As illustrated in Fig. 2, samples conditioned under higher relative humidity showed an increase in the pixel score

values. In general, positive scores in PCA are associated with increased absorbance on positive loadings, which indicated that sample conditioning increased absorbance at approximately 1410 and 1930 nm based on the loading vector. These are the main NIR peaks for water-related vibrational modes and are generated by the first overtone of the O–H stretching vibration generally detected at 1414 nm and the O–H stretching vibration and O–H deformation vibration from H_2O itself within 1916–1942 nm [34]. The first score image also suggested that the acetylated samples had lower final moisture contents as the score values within a specific relative humidity decreased towards increasing acetylation degree. It should be noted that small deviations from NIR peaks reported in the literature are normal in hyperspectral image analysis due to the comparatively lower spectral resolution of imaging spectrographs.

The second and third principal components each explained 8–9% of the remaining data variation and illustrated the effects of wood acetylation within the mosaic. Especially, the earlywood parts of the sample surfaces showed significant changes with acetylation (Fig. 2). The second principal component represented a combination of acetylation and sample conditioning based on the score image. The negative loading peak identified at approximately 1150 nm most likely

Figure 3 **a** Calibration and prediction errors as a function of PLS latent variables, **b** minimum average test set prediction errors on five latent variables based on iPLS, **c** calibration and prediction errors with 9 intervals based on iPLS, **d** measured and predicted WPG values of the calibration and test sets with three latent variables and 9 intervals based on iPLS, and **e** the PLS model vector with three latent variables and 9 intervals based on iPLS.
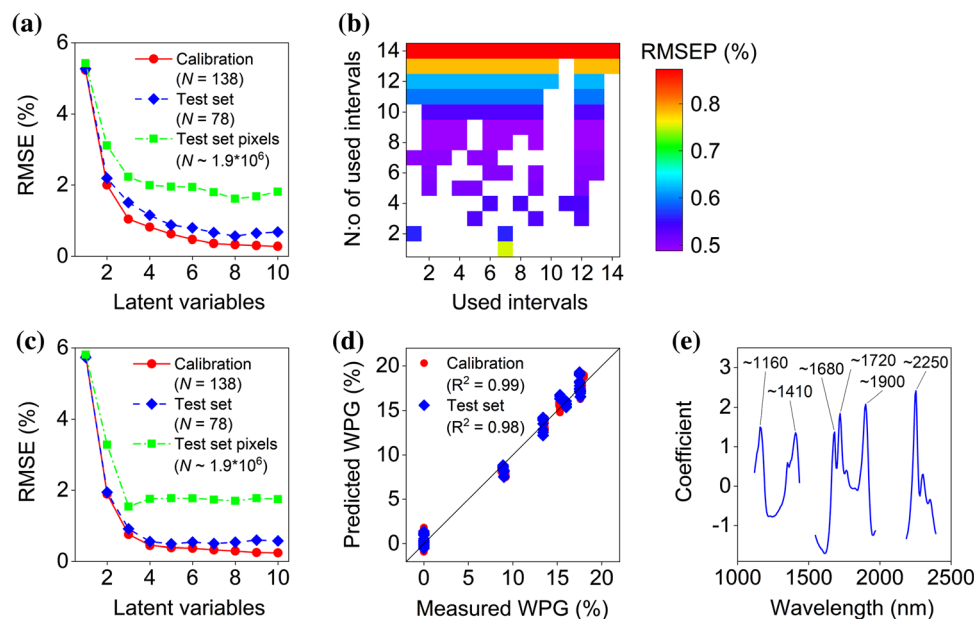

corresponded with the second overtone of the C-H stretching vibration of methyl groups and aromatic moieties in lignin or methyl groups of acetyl esters in hemicelluloses generally seen at 1143 and 1157 nm, respectively [34]. The second overtone of the C=O stretching vibration in hemicellulose is generally seen at 1910 nm [34], which suggested an increase in acetyl groups with negative pixel scores. The differences across the sample scores also highly resembled the first score image, which further illustrated the correlation between acetylation and sample moisture content. The second acetylation peak at circa 2260 nm was likely caused by O–H and C-O stretching vibrations of acetyl groups generally seen in acetylated wood at 2255 nm [34].

The third principal component provided the clearest distinction in acetylation based on the mosaic. A clear trend in the pixel scores was seen towards the direction of increasing sample acetylation. The positive loading peaks at approximately 1160, 1680, 1720, 1900 and 2255 nm were consistent with an increase in the methyl groups of acetyl esters in hemicelluloses (reported at 1157 and 1175 nm), acetyl groups in acetylated wood based on the first overtone of C-H stretching vibration (1681 and 1721 nm), carbonyl groups in hemicelluloses based on the second overtone of the C=O stretching vibration (1907 and 1910 nm), and acetyl groups in acetylated wood based on O–H and C–H stretching vibrations (2255 nm) [34]. The remaining broader loading peak at circa 1410 nm was likely a

combination of two separate peaks. The first overtone of O–H stretching vibrations from phenolic hydroxyl groups or wood lignin is generally seen at 1410 nm [34], of which the latter was in line with the general increase in lignin in earlywood compared with latewood [35]. The peak also showed a clear shoulder on the left at approximately 1350 nm, which suggested an increase in the first overtone of C–H stretching vibrations and C–H deformation vibrations from acetylated wood components (1350 nm) [34].

After PCA, we divided the 36 sample images into separate calibration and prediction sets with 23 and 13 objects, respectively. Six average spectra were then extracted from each sample image, and calibration models were determined for gravimetrically measured WPG values based on SNV transformed and mean-centered calibration spectra. The effects of spectral preprocessing on the extracted calibration spectra are illustrated in Fig. S.1 (Supplementary material). Average prediction errors based on model calibration (RMSEC), prediction of the test set (RMSEP), and prediction of the test set image pixels (RMSEP$_{im}$) were then used for determining the correct number of latent variables for the PLS calibration model. As illustrated in Fig. 3a, no clear shoulders were observed in the determined diagnostics as RMSEC and RMSEP monotonically decreased as more model components were used. Even RMSEP$_{im}$ did not clearly suggest a correct number of latent variables through increasing prediction errors based

Figure 4 Predicted weight percent gains (WPG, %) of the test images (above) and the respective pixel histograms (below) for **a** a non-acetylated sample, **b** an acetylated sample from the center of the design, and **c** a heavily acetylated sample. The red vertical lines in the histograms denote gravimetrically measured WPG values and the blue lines the respective model predictions expressed as an average of the pixel populations.
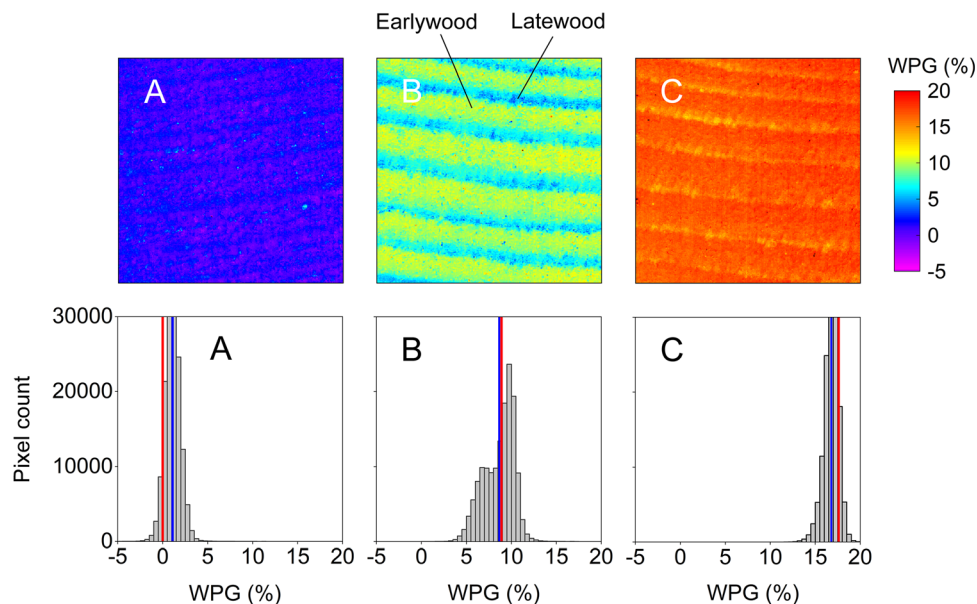
on the test set image pixels that would have indicated potential model overfitting [33].

However, we initially used the entire wavelength range within 1000–2500 nm and not all wavelengths are necessarily equally important for a spectral calibration model. Thus, the best combination of spectral variables for predicting WPG and minimizing RMSEP was determined based on interval PLS (iPLS) [36]. One variable (1005 nm) was first omitted and the wavelength range was divided into 14 equal intervals, which each consisted of 19 spectral variables. All possible combinations of using 1–14 wavelength intervals were then applied for model calibration and determining RMSEP based on five latent variables. Five latent variables were initially chosen based on $RMSEP_{im}$, which indicated that adding more model components did not provide significant improvement in model performance (Fig. 3a). A total of 16,383 calibration models were determined, and the combinations with the lowest average test set prediction errors values are illustrated in Fig. 3b.

Based on the iPLS results, using 9 out of the original 14 intervals provided the lowest average prediction errors on five latent variables and decreased RMSEP from 0.87 to 0.49% WPG (Fig. 3b). As shown in Figs. 3c, using these 9 intervals also provided a clearer indication of the correct number of latent variables based on $RMSEP_{im}$ which suggested that using four or more model components increased prediction errors based on the pixel populations of

the test images. Three latent variables provided the lowest $RMSEP_{im}$ and led to satisfactory prediction of the calibration and test sets with average prediction errors of 0.75% and 0.91% WPG based on extracted mean spectra, respectively (Fig. 3d). The final model vector illustrated in Fig. 3e also included the main peaks affected by sample acetylation as was earlier indicated by the third principal component during the interpretation of the NIR sample mosaic.

The calibration model on three latent variables was then used for predicting the WPG values of the test images and examples are illustrated in Fig. 4. Clear differences were visible in the predicted WPG values of the earlywood and latewood regions of the wood surfaces especially with the samples that were situated in the center of the experimental design (Fig. 4b). The results indicated that the earlywood regions had significantly higher WPG values and had thus been more severely acetylated than the latewood regions of the wood surface. These earlywood and latewood differences were also clearly visible in the pixel histogram, which suggested that the WPG values of the individual pixels could be described with two different distributions (Fig. 4b). The differences were less pronounced but still visible in the heavily acetylated samples (Fig. 4c).

The spatial resolution in hyperspectral NIR imaging is commonly limited to the macrostructure level, which enables separating the earlywood and latewood parts of the wood surface but does not allow determining chemical differences on the cell level.

Figure 5 Score images and respective loading vectors of the first four principal components based on the Raman mosaic showing earlywood (left) and latewood (right) cells. The color scales denote pixel score values.
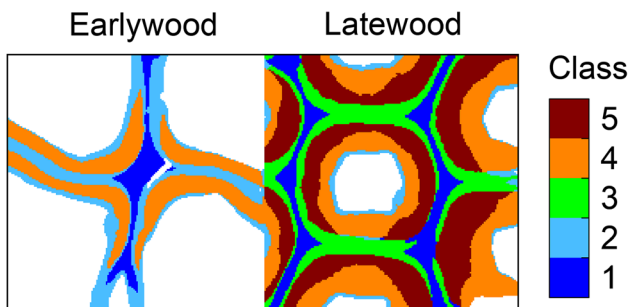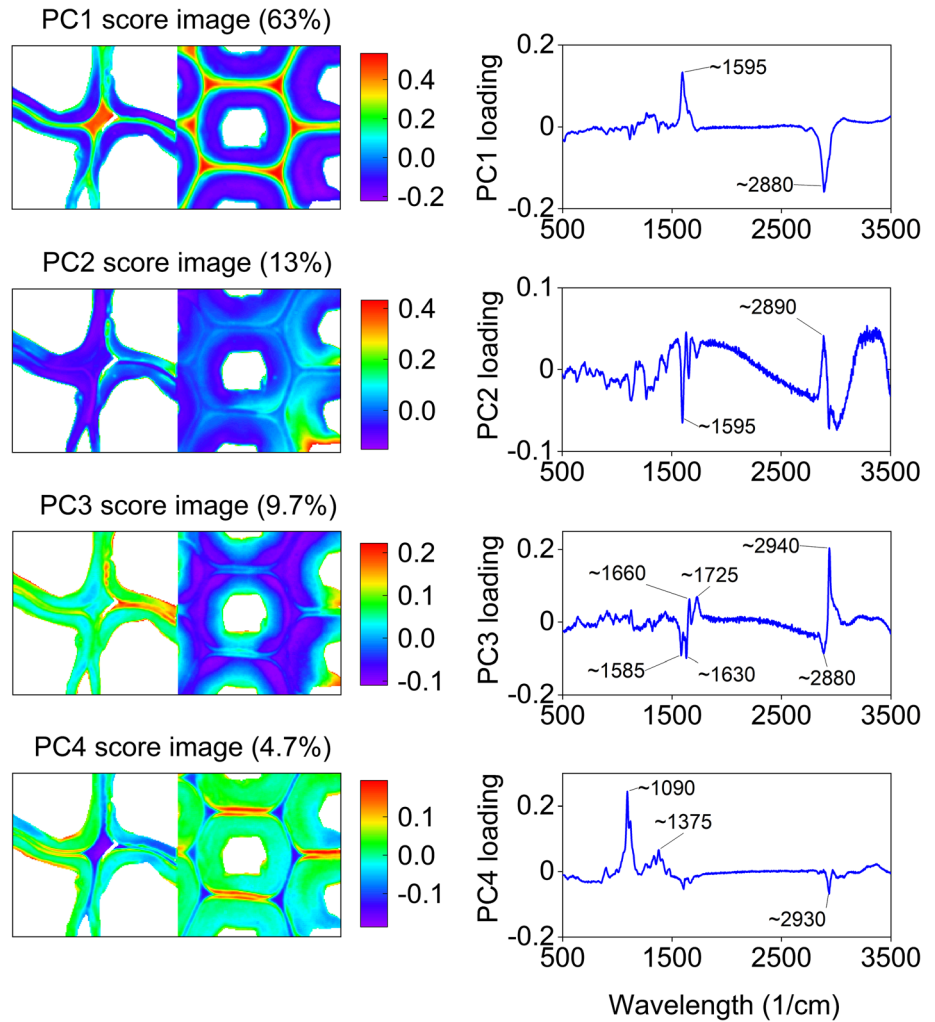
Figure 6 Class image of earlywood and latewood cells based on k-means clustering of principal component scores determined based on the Raman image mosaic.

For determining cell level differences between earlywood and latewood acetylation, we recorded Raman images of wood cells separated from different wood sections from the sample illustrated in Fig. 4b. The Raman images were combined into a mosaic and preprocessed using baseline removal, normalization and median filtering. Pixels corresponding to cell lumens were then removed, and the remaining mosaic was then decomposed using PCA. The first four principal components shown in Fig. 5 explained 91% of data variation and the first, third and fourth components provided a meaningful interpretation of the chemical differences within the earlywood and latewood cell walls. The second principal component mainly represented measurement uncertainties and a loss of focus due to differences in sample height on the focal plane. After four principal components, the noise in the loadings also increased significantly.

The first component explained 61% of data variation within the Raman mosaic and separated the image pixels into areas of high lignin and high carbohydrate contents. As illustrated in Fig. 5, these areas showed increasing intensities from the aromatic ring stretch band at 1600 cm^{-1} and the carbohydrate C–H stretch at 2880 cm^{-1} [37], respectively. The

effects of wood acetylation appeared primarily in the third principal component, which explained 9% of data variation. The score images showed positive score values in the cell walls and the middle lamellae in earlywood, whereas in latewood positive scores were only seen in the inner cell wall layers. The loading vector showed corresponding positive bands at 2940 cm^{-1} and 1660 cm^{-1} due to the C–H stretch of O–CH$_3$ groups and the C=O of acetyl groups previously observed in the Raman spectra of acetylated lignin [37]. The loading also includes a positive band at 1725 cm^{-1}, which was likely due to the carbonyl stretching of acetyl groups previously reported for acetylated lignin and wood [37, 38]. The fourth component explained 5% of the remaining data variation and mainly illustrated the outer layers of the cell walls based on a strong contribution from the orientation sensitive cellulose band at 1090 cm^{-1} [39].

The score values of the first, third and fourth principal components were then used for unsupervised classification of the mosaic pixels. The classification was based on a k-means clustering algorithm which evaluated similarity based on correlation in principal component space. A given number of cluster centroids was initially chosen, and the final number of clusters were determined based on average correlation between the pixels and the cluster centroids. We finally used five clusters with an average correlation of 0.91 as using six or more provided only minute improvements in average correlation (Fig. S.2) and did not enhance interpretation of the results. The obtained class image is illustrated in Fig. 6. Average mean-centered spectra of the corresponding classes are given in Fig. S.3 for brevity.

As illustrated by the results, the effects of acetylation were associated with classes 2 and 4 which represented earlywood cell walls and middle lamellae (Fig. 6). The average spectrum of class 2 showed a negative band of the carbohydrate C-H stretch region at 2890 cm^{-1} (Fig. S.3). The class 4 spectrum featured a positive band at 2890 cm^{-1} and negative bands at the aromatic ring stretch region at approximately 1600 cm^{-1}. The average spectra of both classes also showed positive acetylation marker bands at approximately 2940 and 1720 cm^{-1} which indicated that classes 2 and 4 represented acetylated low and high carbohydrate regions, respectively. In earlywood, these regions encompassed the entire cell wall and most of the middle lamellae, while the cell corner region was represented by class 1. The average

spectrum of class 1, however, contained no acetylation marker bands, which suggested that their intensities were close the mean intensity of the entire data set.

In latewood, the cell corner regions were also represented by class 1, but the acetylated regions represented by classes 2 and 4 extended only halfway through the cell wall (Fig. 6). The remaining central cell wall layer was represented by class 5, while the outer cell wall layers and parts of the middle lamellae were represented by class 3. The mean-centered average spectra of these classes showed either no or featured negative acetylation marker bands, which indicated that the outer cell wall regions were less acetylated in latewood cells.

Discussion

Our objective was to determine the spatial differences in wood acetylation based on hyperspectral NIR image regression. Already the exploration of the NIR image mosaic based on PCA suggested that there were significant differences in the acetylation and resulting moisture content of earlywood and latewood and proved useful for visualizing the chemical differences within the experimental domain. NIR image calibration based on extracted mean spectra provided a reliable calibration model for predicting acetylation through sample WPG. The variable selection procedure enabled choosing the most important wavelength combinations and improved estimating the pseudorank of the final model, which resulted in average prediction errors of 0.91% WPG based on the test set mean spectra and three latent variables. Such a prediction error within a range of 17.6% WPG makes our spectral calibration model suitable for quality control applications based on the general rules of thumb in the NIR field [7]. Our capability to predict WPG within a broad humidity range further demonstrated the suitability of the calibration model.

Once the reliability of the calibration model had been established, we predicted the WPG values of the individual test set image pixels which showed clear differences in the acetylation of earlywood and latewood. These differences earlywood and latewood acetylation, however, decreased in the heavily acetylated samples, which suggests that they were caused by the different rates of acetylation in

earlywood and latewood. Since the rate of acetylation is generally controlled by anhydride diffusion in the wood cell walls [41], the thin-walled earlywood can be expected to acetylate faster than the thick-walled latewood. A significantly higher reactivity of acetic anhydride has previously been found for lignin compared with hemicelluloses and cellulose based on isolated cell wall compounds [40, 42]. As earlywood generally has higher lignin contents than latewood [35, 43], this difference in chemical composition may also have contributed to a higher acetylation rate in earlywood.

We have recently observed similar differences in the acetylation rate of earlywood and latewood when the acetylation was initiated after a confined, one-sided flow of acetic anhydride into the wood [22]. However, earlier studies in the wood field, in which wood samples were completely soaked in acetic anhydride prior to acetylation, have found little to no difference in the WPGs of earlywood and latewood. The results from these earlier studies stand in contrast to our present study, which can likely be explained by differences in experimental methodologies. As an example, Moon et al. [44] and Sadeghifar et al. [45] both performed bulk measurements on mechanically separated and heavily acetylated earlywood and latewood with WPGs in the range 15–22%. It is possible that the authors have been unable to detect the narrow low-WPG regions that persisted in heavily acetylated latewood based on our results. With increasing acetylation time and the resulting high WPG values, the differences between earlywood and latewood acetylation are, however, likely to disappear completely.

We also determined the effects of acetylation on the wood cell level using additional Raman measurements. Unsupervised classification of the Raman images enabled choosing the chemically meaningful principal components for pixel classification after PCA decomposition of the Raman mosaic. The classification results supported the NIR results and revealed that the degree of acetylation decreased from the cell lumens towards the cell corner regions in the thick-walled latewood cells. This explains the observed difference in acetylation degree between earlywood and latewood and supports our hypothesis that the acetylation of the latewood cell walls was limited by the diffusion of acetic anhydride within the wood cells.

Conclusions

We have illustrated how hyperspectral imaging coupled with chemometrics revealed wood acetylation on the wood surface and cellular levels. These results extend our recent findings where we predicted the WPG values of image pixels based on cross-sectional hyperspectral NIR images after confined one-sided acetylation of wood samples. Our current results showed that this approach is not limited to evaluating the penetration depth of acetic anhydride and can be successfully extended to take into account changes in wood equilibrium moisture contents. Unsupervised classification of Raman images obtained from acetylated earlywood and latewood cells supported our hypothesis that the observed differences in earlywood and latewood acetylation on the wood surfaces were caused by differences in the diffusion of acetic anhydride within the earlywood and latewood cells. Overall, our findings pave way for developing novel machine vision methods for the wood field and are an important step for understanding how chemical treatment affects the different hierarchical structures of wood and their hygroscopic properties on different spatial scales.

Author contributions

Conceptualization (MM, MA, LR), Methodology (MM, MA), Experiments (MM, MA, TB), Software (MM.), Data analysis (MM, MA, TB), Visualization (MM), Writing (original draft, review & editing; MM, MA, TB, LR).

Data availability

The datasets used in this study are available from the corresponding author upon reasonable request.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Electronic supplementary material: The online version of this article (https://doi.org/10.1007/s108 53-020-05597-0) contains supplementary material, which is available to authorized users.

References

[1] Fuchs W (1928) Genuine lignin. I Acetylation of pine wood Berichte 61B:948–951

[2] Beck G, Strohbusch S, Larnøy E, Militz H, Hill C (2017) Accessibility of hydroxyl groups in anhydride modified wood as measured by deuterium exchange and saponification. Holzforschung 72:17–23. https://doi.org/10.1515/hf-20 17-0059

[3] Popescu C-M, Hill CAS, Curling S, Ormondroyd G, Xie Y (2014) The water vapour sorption behaviour of acetylated birch wood: how acetylation affects the sorption isotherm and accessible hydroxyl content. J Mater Sci 49:2362–2371. https://doi.org/10.1007/s10853-013-7937-x

[4] Hill CAS (2008) The reduction in the fibre saturation point of woo due to chemical modification using anhydride

reagents: a reappraisal. Holzforschung 62:423–428. https://d oi.org/10.1515/HF.2008.078

[5] Hunt CG, Zelinka SL, Frihart CR, Lorenx L, Yelle D, Gleber S-C, Vogt S, Jakes JE (2018) Acetylation increases relative humidity threshold for ion transport in wood cell walls - a means to understanding decay resistance. Int Biodeterior Biodegrad 133:230–237. https://doi.org/10.1016/j.ibiod.201 8.06.014

[6] Boldrini B, Kessler W, Rebner K, Kessler RW (2012) Hyperspectral imaging: a review of best practice, performance and pitfalls for in-line and on-line applications. J Near Infrared Spectrosc 20:483–508. https://doi.org/10.1255/1003

[7] Manley M (2014) Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chem Soc Rev 43:8200–8214. https://doi.org/10. 1039/c4cs00062e

[8] Toumpanaki E, Shah DU, Eichhorn SJ (2020) Beyond what meets the eye: imaging and imagining wood mechanical-structural properties. Adv Mater. https://doi.org/10.1002/ad ma.202001613

[9] Gierlinger N, Schwanninger M (2007) The potential of Raman microscopy and Raman imaging in plant research. J Spectrsc 21:69–89. https://doi.org/10.1155/2007/498206

[10] Thumm A, Riddell M, Nanayakkara B, Harrington J, Meder R (2010) Near infrared hyperspectral imaging applied to mapping chemical composition in wood samples. J Near Infrared Spectrosc 18:507–515. https://doi.org/10.1255/jnirs. 909

[11] Colares CJG, Pastore TCM, Coradin VTR, Marques LF, Moreira ACO, Alexindro GL, Poppi RJ, Braga JWZ (2016) Near infrared hyperspectral imaging and MCR-ALS applied for mapping chemical composition of the wood specie Swietenia Macrophylla King (Mahogany) at microscopic level. Microchem J 124:356–363. https://doi.org/10.1016/j. microc.2015.09.022

[12] Ma T, Inagaki T, Ban M, Tsuchikawa S (2019) Rapid identification of wood species by near-infrared spatially resolved spectroscopy (NIR-SRS) based on hyperspectral imaging (HSI). Holzforschung 73:323–330. https://doi.org/ 10.1515/hf-2018-0128

[13] Haddadi, A., Burger, J., Leblon, B., Pirouz, Z., Groves, K., & Nader, J. (2015). Using near-infrared hyperspectral images on subalpine fir board Part 1: Moisture content estimation. *Wood Material Science & Engineering*, *10*(1), 27-40.

[14] Haddadi A, Burger J, Leblon B, Pirouz Z, Groves K, Nader J (2013) Using near-infrared hyperspectral images on subalpine fir board. Part 2: density and specific gravity estimation. Wood Mater Sci Eng 10:41–56. https://doi.org/10.1080/ 17480272.2015.1011231

[15] Lestander T, Geladi P, Larsson SH, Thyrel M (2012) Near infrared image analysis for online identification and separation of wood chips with elevated levels of extractives. J Near Infrared Spectrosc 20:591–599. https://doi.org/10.1255/jnirs.992

[16] Inagaki T, Mitsui K, Tsuchikawa S (2015) Visualisation of degree of acetylation in beechwood by near infrared hyperspectral imaging. J Near Infrared Spectrosc 23:353–260. https://doi.org/10.1255/jnirs.1181

[17] Gierlinger N, Keplinger T, Harrington M (2012) Imaging of plant cell walls by confocal Raman microscopy. Nat Protoc 7:1694–1708. https://doi.org/10.1038/nprot.2012.092

[18] Belt T, Keplinger T, Hänninen T, Rautkari L (2017) Cellular level distributions of scots pine heartwood and knot heartwood extractives revealed by Raman spectroscopy imaging. Ind Crop Prod 108:327–335. https://doi.org/10.1016/j.indcrop.2017.06.056

[19] Merk V, Chanana M, Keplinger T, Gaan S, Burgert I (2015) Hybrid wood materials with improved fire retardance by bio-inspired mineralisation on the nano and sub-micron level. Green Chem 17:1423–1428. https://doi.org/10.1039/C4GC01862A

[20] Ermeydan MA (2018) Modification of spruce wood by UV-crosslinked PEF hydrogels inside wood cell walls. React Funct Polym 131:100–106. https://doi.org/10.1016/j.reactfunctpolym.2018.07.013

[21] Altgen M, Awais M, Altgen D, Klüppel A, Mäkelä M, Rautkari L (2020) Distribution and curing reactions of melamine formaldehyde resin in cells of impregnation-modified wood. Sci Rep 10:3366. https://doi.org/10.1038/s41598-020-60418-3

[22] Awais M, Altgen M, Mäkelä M, Altgen D, Rautkari L (2020) Hyperspectral near-infrared image assessment of surface-acetylated solid wood. ACS Appl Bio Mater 3:5223–5232. https://doi.org/10.1021/acsabm.0c00626

[23] Greenspan L (1977) Humidity fixed points of binary saturated aqueous solutions. J Res Natl Bur Stand 81A:89–96

[24] 24Strømdahl K (2000) Water sorption in wood and plant fibres. PhD thesis thesis, Technical University of Denmark.

[25] Winston PW, Bates DH (1960) Saturated solutions for the control of humidity in biological research. Ecology 41:232–2387. https://doi.org/10.2307/1931961

[26] Mäkelä M, Volpe M, Volpe R, Fiori L, Dahl O (2018) Spatially resolved spectral determination of polysaccharides in hydrothermally carbonized biomass. Green Chem 20:1114–1120. https://doi.org/10.1039/c7gc03676k

[27] Bro R, Smilde AK (2014) Principal component analysis. Anal. Methods 6:2812–2831. https://doi.org/10.1039/c3ay41907j

[28] Amigo JM, Babamoradi H, Elcoroaristizabal S (2015) Hyperspectral image analysis. A tutorial Anal Chim Acta 896:34–51. https://doi.org/10.1016/j.aca.2015.09.030

[29] Fearn T, Riccioli C, Garrido-Varo A, Guerrero-Ginel JE (2009) On the geometry of SNV and MSC. Chemometrics Intell Lab Syst 96:22–26. https://doi.org/10.1016/j.chemolab.2008.11.006

[30] Geladi P, Kowalski B (1986) Partial least-squares regression: a tutorial. Anal Chim Acta 185:1–17. https://doi.org/10.1016/0003-2670(86)80028-9

[31] Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst 58:109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

[32] 32de Jong S, (1993) SIMPLS: an alternative approach to partial least squares regression. Chemometrics Intell Lab Syst 18:251–263. https://doi.org/10.1016/0169-7439(93)85002-X

[33] Gowen A, Burger J, Esquerre C, Downey G, O'Donell C (2014) Near infrared hyperspectral image regression: on the use of prediction maps as a tool for detecting model over-fitting. Near Infrared Spectrosc 22:261–270. https://doi.org/10.1255/jnirs.1114

[34] Schwanninger M, Rodriques JC, Fackler K (2011) A review of band assignments in near infrared spectra of wood and wood components. J Near Infrared Spectrosc 19:287–308. https://doi.org/10.1255/jnirs.955

[35] Bertaud F, Holmbom B (2004) Chemical composition of earlywood and latewood in Norway spruce heartwood, sapwood and transition zone wood. Wood Sci Technol 38:245–256. https://doi.org/10.1007/s00226-004-0241-9

[36] Andersen CM, Bro R (2010) Variable selection in regression - a tutorial. J Chemometr 24:728–737. https://doi.org/10.1002/cem.1360

[37] Agarwal UP, McSweeny JD, Ralph SA (2011) FT-Raman investigation of milled-wood lignins: softwood, hardwood, and chemically modified black spruce lignins. J Wood Chem Technol 31:324–344. https://doi.org/10.1080/02773813.2011.562338

[38] Olaniran SO, Etienne C, Keplinger T, Olufemi B, Rüggeberg M (2019) Mechanical behaviour of acetylated rubber wood subjected to artificial weathering. Holzforschung 73:1005–1016. https://doi.org/10.1515/hf-2018-0274

[39] Gierlinger N, Luss S, König C, Konnerth J, Eder M, Fratzl P (2010) Cellulose microfibril orientation of Picea abies and its variability at the micron-level determined by Raman imaging. J Exp Bot 61:587–595. https://doi.org/10.1093/jxb/erp325

[40] Rowell RM, Simonson R, Hess S, Plackett DV, Cronshaw D, Dunningham E (1994) Acetyl distribution in acetylated

whole wood and reactivity of isolated wood cell-wall components to acetic anhydride. Wood Fiber Sci 26:11–18

[41] Hill CAS, Jones D, Strickland G (1998) Cetin NS (1998) Kinetic and mechanistic aspects of the acetylation of wood with acetic anhydride. Holzforschung 52:623–629. https://doi.org/10.1515/hfsg.1998.52.6.623

[42] Adebawo FG, Sadeghifar H, Tilotta D, Jameel H, Liu Y, Lucia L (2019) Spectroscopic interrogation of the acetylation selectivity of hardwood polymers. Starch 71:1900086. https://doi.org/10.1002/star.201900086

[43] Fredriksson M, Bjerregaard Pedersen N, Garbrecht Thygesen L (2018) The cell wall composition of Norway spruce earlywood and latewood revisited. Int Wood Prod J 9:80–85. https://doi.org/10.1080/20426445.2018.1479680

[44] Moon RJ, Wells J, Kretschmann DE, Evans J, Wiedenhoeft AC, Frihart CR (2010) Influence of chemical treatments on moisture-induced dimensional change and elastic modulus of earlywood and latewood. Holzforschung 64:771–779. https://doi.org/10.1515/HF.2010.106

[45] Sadeghifar H, Dickerson JP, Argyropoulos DS (2014) Quantitative 31p NMR analysis of solid wood offers an insight into the acetylation of its components. Carbohydr Polym 113:552–560. https://doi.org/10.1016/j.carbpol.2014.07.046