Biomass and Bioenergy 208 (2026) 108881

Contents lists available at ScienceDirect

Biomass and Bioenergy

journal homepage: www.elsevier.com/locate/biombioe

Rapid chemical characterization of loblolly pine forest residues with near-infrared hyperspectral imaging

Sameen Raut [a,1], Ighoyivwi Onakpoma [b,1], Seyedehsan Vasefi [c], Joseph Dahlen [a,*], Laurence R. Schimleck [b], Sudhagar Mani [c], Gerald Presley [b], Thomas L. Eberhardt [d]

[a] Warnell School of Forestry and Natural Resources, University of Georgia, 180 E Green St, Athens, GA, 30602, USA
[b] Wood Science and Engineering, Oregon State University, 119 Richardson Hall, Corvallis, OR, 97331, USA
[c] College of Engineering, University of Georgia, 110 Riverbend Road, Athens, GA, 30602, USA
[d] Forest Products Laboratory, USDA Forest Service, One Gifford Pinchot Drive, Madison, WI, 53726, USA

ARTICLE INFO

Keywords:
Ash
Cellulose
Extractives
Hemicellulose
Lignin
Pinus taeda
Spectral data analysis

ABSTRACT

Loblolly pine (*Pinus taeda*) is the most important commercial tree species in southeastern U.S. In recent years, there has been a growing interest in utilizing loblolly pine wood and harvest residues consisting of bark, branches, needles, and wood, as a viable renewable feedstock to meet growing energy, chemical, and biomaterial demand. Feedstock chemical compositions affect the conversion efficiency and product yields, but these properties are costly and laborious to measure. Calibration models capable of rapid and precise estimation of these properties were developed using near-infrared (NIR) hyperspectral imaging data from southern pine forest harvest residues collected from different physiographic regions and age groups. Partial least squares regression models showed good cross-validation fit statistics for most measured properties, with 6 of 12 having $R^2_{cv} > 0.9$, thus demonstrating that NIR hyperspectral imaging is a rapid and reliable tool for estimation of chemical composition. The results reported here have important implications in rapid feedstock suitability analysis for conversion of southern pine harvest residues to biofuels and bioproducts.

1. Introduction

Society can reduce its reliance on fossil fuels by utilizing sustainably sourced carbon neutral biomass sources [1,2] to which biomass from forests is the most common form used to generate energy and bioproducts [3]. In the United States, approximately 320 million dry tons of forest biomass are harvested annually and these operations generate approximately 90 million dry tons of forest residues [4,5]. Forest residues are generally left behind on the forest floor because they have high collection and transportation costs due to their low bulk density, low energy density, high moisture content, high ash content, and higher variability in their chemical composition [6–10]. The high variability in forest residues result in both logistics and process complications at a conversion facility which has resulted in limited use of forest residues outside of direct combustion as fuel for biopower and/or combined heat and power generation [11].

For forest residues to be utilized in applications beyond combustion, it is important to know the chemical composition to ensure their suitability for efficient conversion to biofuels and bioproducts [12]. The protocols to measure chemical properties require expensive equipment and have lengthy procedures limiting their use. Rapid techniques for biomass compositional analysis are needed if these variable forest residue feedstocks are to be better utilized in biorefining applications. NIR spectroscopy has been successfully employed for rapid characterization of various properties, including wood chemical properties in many tree species for the forestry industry [12–18] and biomass properties for the bioenergy industry [19,20]. Near-infrared spectroscopy hyperspectral imaging (NIR-HSI) systems are particularly suited to quantifying forest biomass variability because they combine imaging and spectroscopy [21–25]. Due to the imaging capabilities in NIR-HSI systems, segregation of different forest residue types is more feasible in an industrial setting as these sorting systems are becoming more common for different applications in agriculture [26–29] and waste sorting facilities [30]. However, these systems have not been used in forest biomaterials facilities due to technology limitations, and the inclusion of such systems would require some processing changes within a facility.

* Corresponding author. Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, 30602, USA.
E-mail address: jdahlen@uga.edu (J. Dahlen).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.biombioe.2025.108881
Received 7 November 2025; Received in revised form 11 December 2025; Accepted 19 December 2025
Available online 23 December 2025

In the southeastern U.S., loblolly pine (*Pinus taeda*) is the most widely planted and the most economically important tree species because of its fast growth and high yields [31,32]. Because of this, there has been a growing interest in utilizing loblolly pine wood and harvest residues as a renewable feedstock [4,5,33]. The objective of this study was to develop NIR-HSI calibration models capable of rapidly predicting the monomeric and polymeric sugars, lignin, extractives, and ash content in loblolly pine forest residues of both pure (bark, branches, needles, wood) and mixed fractions in two particle size classes (3 mm round 'pin chips' and 7 mm round 'small accept chips') [34,35]. This study will provide important information for non-destructive evaluation of southern pine forest residues.

2. Materials and methods

2.1. Sample locations

Green southern pine forest residues were collected from five planted loblolly pine stands. To increase the variability in the residues collected, stands were sampled from different physiographic regions (Piedmont, Upper Coastal Plain, and Lower Coastal Plain) [36] and from two age groups (1st thinning: age 10–15 years; final harvest: age 22+ years) (Table 1) [16].

2.2. Collection of forest residues

Ten to twelve trees were felled in each stand. The trees were delimbed and the main stem was cut into approximately 1 m bolts for handling and transport. In the final harvest stands' bolts were taken from heights above 12 m because the wood under 12 m was considered to be mature wood. In the 1st thinning stands the entire stem was considered juvenile wood and thus bolts were taken randomly along the stem. Approximately 400 kg of bolts (wood and bark) and approximately 100 kg of branches containing needles were collected from each stand.

2.3. Processing of forest residues

The bolts were placed in a custom-made drum debarker to separate bark from juvenile wood (referred to as wood for the remainder of the text), and any remaining bark was manually removed using a chisel. Needles were handpicked from the branches. Branches, needles, and wood were separately chipped in the green condition and then sorted into different size classes (oversized (>45 mm), 16 mm round, 7 mm round, 3 mm round, and fines) using a TMI Chip Class sorter (RDM Test Equipment LLC, New Brighton, MN, USA) [34,35]. The oversized material was rechipped and sieved again. The materials were weighed after chipping and then oven dried at 103 ± 2 °C (65 ± 2 °C for needles) and then reweighed when the weight stabilized (achieved 0 % MC). Bark was oven-dried first before chipping and sieving because chipping green bark did not yield enough 3 and 7 mm round particles to match other biomass types.

Table 1
Descriptive information of the five stands from where the forest residues were sourced.

Stand	Region	Age (yrs)	Category	Latitude	Longitude
1	Upper coastal plain	35	Final harvest	33.30799	−81.74788
2	Lower coastal plain	15	1st thinning	30.745509	−81.883691
3	Piedmont	12	1st thinning	33.4311397	−83.4579118
4	Lower coastal plain	24	Final harvest	33.2650808	−79.4235915
5	Piedmont	22	Final harvest	34.594429	−80.9371121

2.4. Forest residue processing and blending

From each stand, only the 3 mm round and the 7 mm round chipped-material size classes were used while making pure or mixed fractions because the 16 mm size class consisted mostly of wood; thus the 16 mm size class did not yield a sufficient amount of useable materialfor every forest residue type, especially for bark and needles. For each of the five forest residue types (4 pure and 1 mixed), 2 kg of material per stand were used in this study, and the rest of the materials were used for grinding energy efficiency studies. The mass distribution of the material in each forest residue type and their respective chip-size classes (3 mm round and 7 mm round) from the original collected material was calculated and then mixed fractions were made by mixing chipped pure forest residue material based on this mass distribution (Table 2). A flowchart illustrating the steps involved in collecting, processing, and blending of different forest residues is shown in Fig. 1.

2.5. Hyperspectral image collection

A push broom near-infrared (NIR) hyperspectral camera with an InGaAs sensor (Specim FX17, Specim Spectral Imaging Ltd., Oulu, Finland) and lens (Specim OLET 17.5 f/2.1) was used to scan chipped materials. The device collected spectral data across 224 wavelengths ranging from 931 to 1718 nm at 3.5 nm intervals. The working distance of the camera was 230 mm from the sample. Two sets of tungsten halogen lamps positioned 240 mm from the samples, each containing 3 bulbs, were used to illuminate samples at an angle of 45°. Sample movement and image acquisition were managed using application software (Specim Lumo) provided with system. A dark current (camera shutter closed), and a white reference were collected before acquiring each image for intensity calibration. The scanning speed was set at 12.00 mm/s while the exposure time was set to 5.25 ms. This exposure time was selected so that the white reference peak value reached approximately 85 % of the detector range (0–4095, bit depth = 12) to ensure optimum linear dynamic range utilization. All chipped materials were conditioned in a temperature and humidity-controlled room (22 °C and 52 % RH) to achieve a moisture content (MC) of about 10 % (dry-basis) before scanning. The conditioned samples were loaded to a depth of approximately 7 mm on an anodized black aluminum tray for scanning. The camera was focused at a depth of 7 mm above the aluminum tray, and the depth of field for the imaging setup is approximately 13 mm, thus material above or below the focal plane was in focus except for very tall particles. Each raw image had dimensions of 640 and 1960 pixels corresponding to approximately 187 and 574 mm in width and length captured in each image, respectively.

Table 2
Composition of bark, branches, needles, and wood by weight in a representative forest residue sample for each stand.

Size class	Stand	Bark	Branch	Needle	Wood
3 mm	1	0.11	0.3	0.12	0.47
	2	0.18	0.29	0.14	0.4
	3	0.17	0.45	0.2	0.18
	4	0.25	0.38	0.17	0.2
	5	0.21	0.3	0.23	0.27
	Mean	0.18	0.34	0.17	0.3
7 mm	1	0.01	0.11	0.01	0.87
	2	0.02	0.17	0	0.81
	3	0.03	0.33	0.01	0.63
	4	0.05	0.24	0.05	0.66
	5	0.06	0.21	0.07	0.66
	Mean	0.03	0.21	0.03	0.73

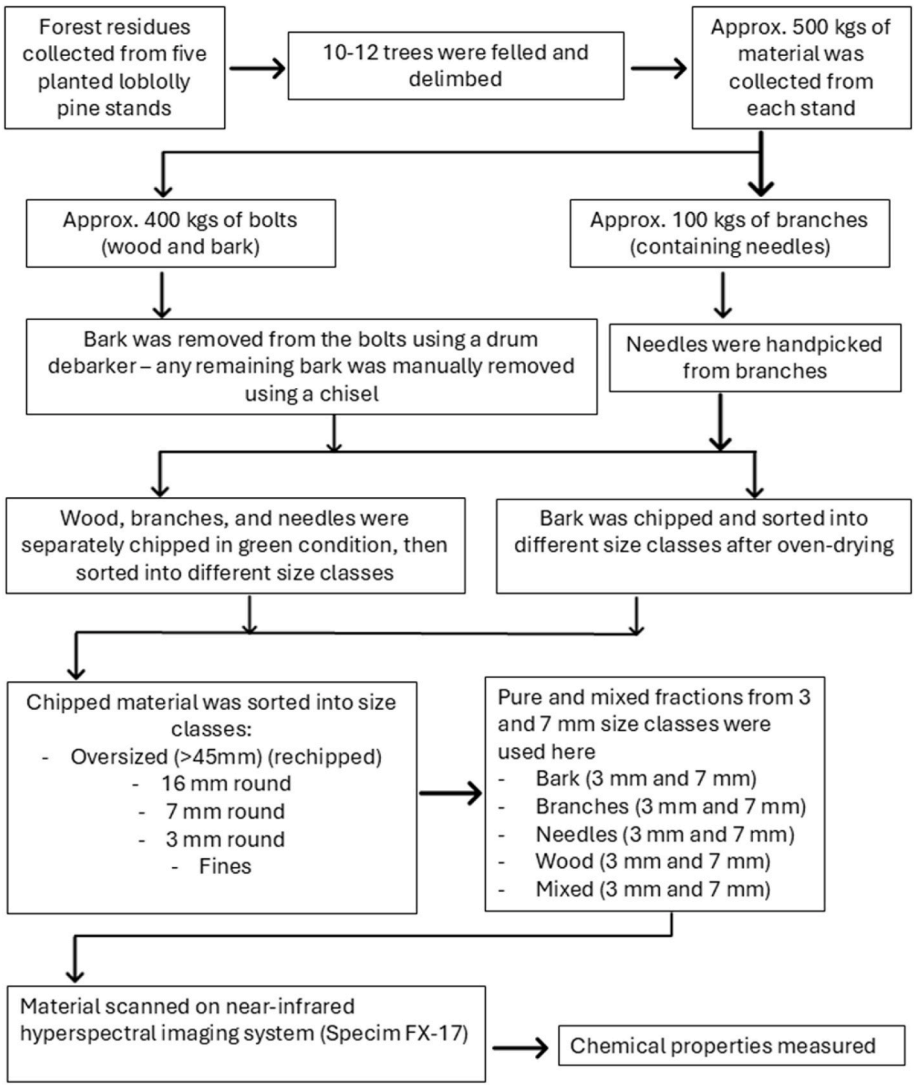

Fig. 1. A flowchart showing the steps involved in collecting, processing, and blending of different forest residues.

2.6. Hyperspectral image processing and spectral data collection

The hyperspectral images were processed in Python version 3.9 (Python Software Foundation, https://www.python.org/) on the Spyder interface [37]. Libraries used were NumPy [38], OpenCV [39], pandas [40], SciPy [41], and Spectral Python (SPy) [42]. Each image was calibrated with the measured dark and white reference signals:

$$R = \frac{R_0 - D}{W - D} \quad (1)$$

where R represents the relative diffuse reflectance, R_0 is the measured diffuse reflectance, D represents the dark current reference signal, and W is the white reference signal. Spectral data in the range 931–942 nm was removed due to poor image quality. Each image was cropped to 540 pixels in width and 1910 pixels in length to reduce the edge effects. Pseudo-color images of bark (red = 1325 nm, green = 1424 nm, blue = 1438 nm), branches (red = 1124 nm, green = 1488 nm, blue = 1537 nm), needles (red = 1343 nm, green = 1378 nm, blue = 1467 nm), wood (red = 1096 nm, green = 1286 nm, blue = 1467 nm), and the mixed fraction (red = 1096 nm, green = 1286 nm, blue = 1467 nm) are shown in Fig. 2.

To isolate the tray from the samples, the standard deviation of each pixel relative diffuse reflectance spectra was calculated, and the image was smoothed using a 5 × 5 median filter which resulted in one value per pixel. Each pixel that had the standard deviation value less than 0.04 was considered the background (tray) and set to 0. A mask was then created where pixels with a value of 0 were the background and all the other pixels were converted to a value of 1 through thresholding. Hierarchical contour retrieval models [43] were applied to detect contours in the image and contours with area less than 500 were removed. This resulted in a mask where the background was removed and only the samples remained. Pixel values from each background-removed image were averaged to provide a single diffuse reflectance spectrum for each tray scanned. A total of 100 trays (80 pure fractions and 20 mixed fractions) were scanned with duplicate scans of the different forest residue/size classes. The duplicate spectra were averaged to give a single spectrum, therefore, a total of 50 unique diffuse reflectance spectra were collected that were utilized for developing NIR models. Each diffuse reflectance spectrum (R) was transformed to pseudo absorbance (A) (referred to as absorbance for the remainder of the text) (Fig. 3) using the equation:

$$A = \log_{10}\left(\frac{1}{R}\right) \quad (2)$$

Using the Savitzky-Golay approach [44] with left and right gaps of 4 points applied to the absorbance data, a first and second derivative spectra pretreatment was calculated which generated 212 X-variables. This approach resolved peaks which overlapped substantially in the

Fig. 2. Pseudo-color images from the hyperspectral imaging setup for bark, branches, needles, wood, and the mixed fraction from the 7 mm size class. (For interpretation of the references to color, refer to manuscript text where this figure is referenced.)

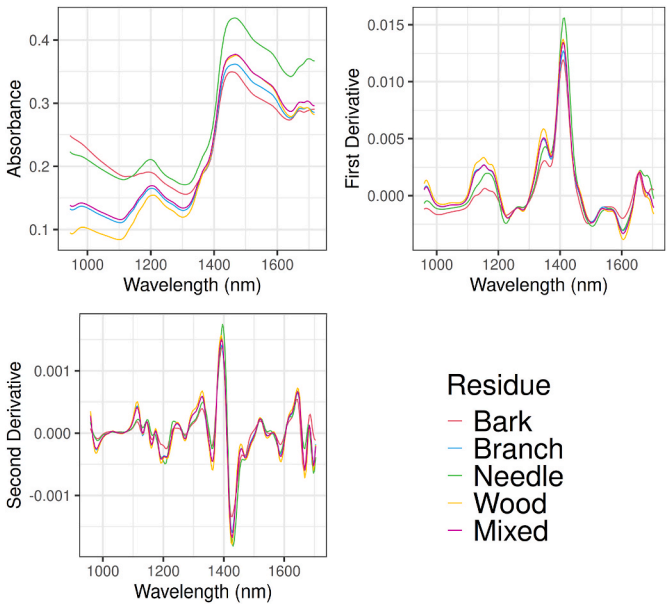

Fig. 3. Forest residue averages for absorbance, first derivative, and second derivative spectra.

original spectral data [45]. The absorbance, first derivative, and second derivative data were utilized to fit NIR models to the chemical composition reference data.

2.7. Chemical composition

The pure and mixed chipped materials were ground in a SM Retsch 2000 cutting mill (Retsch GmbH, Haan, Germany) to pass through a 1-mm sieve size. Approximately 1.5 g of ground material was placed in an F57 filter bag (ANKOM Technology, Macedon, NY, USA) and heat sealed using an impulse sealer. Four replicates were made for each forest residue of each size class. The net oven dry weight (0 % MC) of the ground material inside the filter bags was recorded using a digital scale. The filter bags containing samples were then placed in a Soxhlet extractor and exhaustively extracted, sequentially in hexane, acetone, and ethanol. All samples were extracted for at least one day in each solvent. The post-extraction net oven-dry weight was recorded and used to calculate extractives content (%) on a dry-basis:

$$\text{Extractives content (\%)} = \frac{\text{Weight}_{\text{OD}} - \text{Weight}_{\text{ODE}}}{\text{Weight}_{\text{OD}}} \times 100 \quad (3)$$

where $\text{Weight}_{\text{OD}}$ is the net oven-dry weight of the ground material before extraction and $\text{Weight}_{\text{ODE}}$ is the net oven-dry weight of the ground material after extraction. Determination of structural carbohydrates (glucan, xylan, galactan, arabinan, and mannan) and total lignin was done following the procedures outlined in Ref. [46] and reported on a dry-basis. All structural carbohydrate and lignin measurements were done in duplicate. Cellulose content was then calculated using the formula provided by Easty and Malcom [47]:

$$\text{Cellulose} = \text{Glucan} - \left(\frac{1}{3} \times \text{Mannan}\right) \quad (4)$$

Hemicellulose content was calculated by subtracting the cellulose quantity from the sum of the sugars. Holocellulose content was calculated as the sum of all sugars. Cellulose content was also determined using the diglyme-HCl method with a supplemental bleaching step as outlined in Cullen and Macfarlane [48]. Ash content was determined from five replicates at 600 °C [49,50].

2.8. Statistical analysis and near-infrared spectroscopy models

A correlogram was constructed to display significant Pearson correlations. Partial least squares regression (PLSR) calibration models with a maximum of 10 factors considered were developed to predict variation in each measured property using the NIR absorbance, first derivative, and second derivative data. The PLSR models are reported two ways, the first being that cross-validation was conducted by reserving the spectral and reference data for all forest residues from one stand (five forest residue types, two size classes; hence, 10 data points kept as cross-validation) in each iteration which continued until data from all five stands had been used as the cross-validation set. The coefficient of determination of cross-validation (R^2_{cv}) and root mean square error of

cross-validation (RMSECV) are reported for the first approach. The best model (first derivative, second derivative) was used to make the plots for each property shown in the results section; note that absorbance never provided the best model. In the second approach, data was split into a training set containing four stands (stands 2–5) with an independent test set containing one stand (stand 1). Here the R^2_{cv}, RMSECV, coefficient of determination of prediction (R^2_p) and root mean square error of prediction (RMSEP) are reported for the first and second derivative data.

The variable importance on projection (VIP) for each wavelength was calculated from the best models from the first approach to modeling (cross-validation was conducted by reserving the spectral and reference data for all forest residues from one stand). The peaks in the VIP values were found and the four highest peaks from each model are reported here. For each reported peak, the bond vibration and forest residue cell wall polymer (if known) are also reported, because smoothing may result in subtle shifts in the band position of the peaks, we considered peaks within 5 nm of what has been reported in the literature, and immediately neighboring peaks were grouped together for different properties [45,51]. The correlogram and the PLSR models were developed in R statistical software [52] using the R Studio interface [53]. The packages utilized were the tidyverse collection of packages [54], pls [55], pracma [56], plsVarSel [57], signal [58], and gridExtra [59].

2.9. Chemical properties visualization

The best models for each property were used to visualize the spatial variation on the NIR-HSI images. Each non-background pixel in the NIR-HSI image was converted to the first or second derivative using the procedures previously described, and then the model was used to predict the chemical content. The predictions were constrained by 90 % of the minimum and 110 % of the maximum measured values, and then saved as a png file.

3. Results

The Pearson correlations ranged from −0.72 to 0.98 with non-significant correlations (p-value >0.05) being shown with a 'X' in their box (Fig. 4). The strongest positive correlation was cellulose determined from the HPLC [46,47] with holocellulose, given

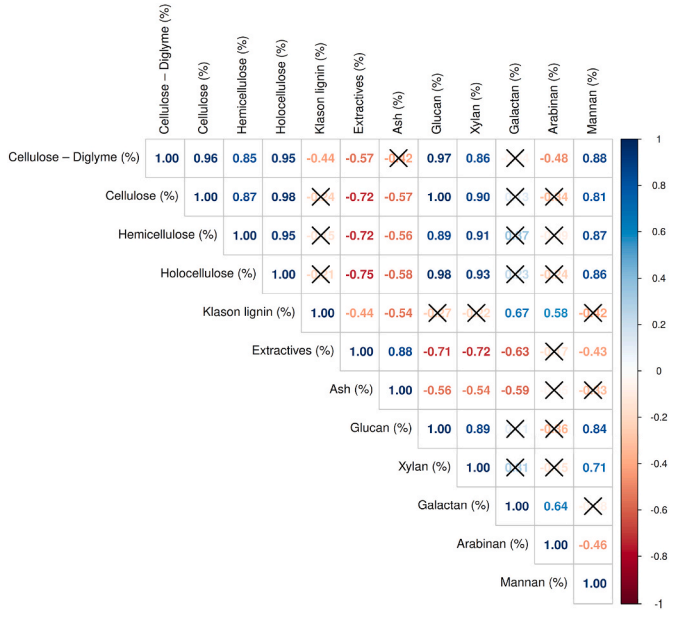

Fig. 4. A correlogram for the chemical properties measured in this study. Significant Pearson correlations (p-value ≤0.05) shown as text with non-signficant correlations (p-value >0.05) shown with a 'X' in their box.

holocellulose is the sum of cellulose and hemicellulose from HPLC this finding was logical. Cellulose via diglyme and cellulose via HPLC was very strongly correlated (r = 0.96) which suggests that measuring cellulose via diglyme-HCl with a supplemental bleaching step can provide reliable estimates of cellulose measured via HPLC in the same samples, which is otherwise more challenging to measure as it requires costly equipment. The strongest negative correlation was found between extractives and HPLC cellulose (r = −0.72), glucan (r = −0.71), and xylan (r = −0.72), this finding is logical because needles have the highest amount of extractives (~30 %) but relatively low cellulose content (~22 %).

The PLSR models developed for chemical composition data measured from different forest residue types and size classes yielded models of varying quality (Tables 3 and 4). There were subtle differences between the fit statistics developed using absorbance, first derivative, and second derivative spectra for the cross validation only models. The cross-validation only models differed marginally from those developed using separate training and testing datasets, as such we will only discuss the cross-validation models moving forward. Models with R^2_{cv} values above 0.9 were considered very good/excellent, those between 0.80 and 0.90 were considered good, those between 0.6 and 0.79 were considered fair/satisfactory, and any model with an R^2_{cv} value below 0.6 was considered poor [60]. For the structural compounds (Fig. 5), cellulose had the best fit statistics, with the diglyme-HCl measured cellulose being similar to HPLC cellulose (R^2_{cv} 0.97 vs 0.96, RMSECV 1.3 % vs 1.6 %). Models for total lignin (R^2_{cv} 0.94, RMSECV 1.7 %) and holocellulose (R^2_{cv} 0.93, RMSECV 3.5 %) also had excellent fit statistics. The lignin model performed better than holocellulose across the different forest residue fractions with some additional variation present in the prediction of bark lignin content. In contrast the hemicellulose model (R^2_{cv} 0.75, RMSECV 2.8 %) was considerably noisier due to some individual sugars having excellent or good models, while others (arabinan, galactan, and mannan) were noticeably poorer.

The individual sugar models had varying fit statistics ranging from excellent to fair (Fig. 6). Glucan had excellent fit statistics (R^2_{cv} = 0.97, RMSECV = 1.6 %) with the performance essentially mirroring the HPLC cellulose model. The glucan model performed well across the different forest residue fractions with some variation in the prediction of bark glucan content. Xylan had good fit statistics (R^2_{cv} = 0.89, RMSECV = 0.7 %) but relative to glucan had more variation across the different forest residue fractions. In contrast, the models for galactan (R^2_{cv} = 0.65, RMSECV = 0.6 %) and mannan (R^2_{cv} = 0.65, RMSECV = 1.7 %) can only be considered fair and used for preliminary screening, and both had considerable variation across the different forest residue fractions. The arabinan model (R^2_{cv} = 0.55, RMSECV = 0.6 %) was poor and generally under predicted the mixed fraction values relative to the measured values.

The extractives model had the best R^2_{cv} (0.98) for all of the measured properties with only minor variation in the prediction of the extractives content of needles (Fig. 7). The excellent performance of the extractives model would make it suitable for quality control purposes. Ash had good fit statistics with R^2_{cv} of 0.81 and RMSECV of 0.5 %. There was some variation across the different forest residue fractions for ash, with the mixed fraction predictions being underestimated relative to the measured values.

The wavelengths of importance for the NIR models as determined by variable importance on projection (VIP) is shown in Table 5, from each model the top four peaks are reported here. Eleven out of seventeen of the identified wavelengths have known bond vibrations, and ten of these have a known cell wall polymer or analyte associated with the bond vibration. The model for lignin had the most direct link to wavelengths of importance identified to lignin, as all of the wavelengths are either solely associated with lignin (1170, 1417, and 1683 nm) or associated with all cell wall polymers (1346 nm). Fourteen of the identified wavelengths are relatively short (<1486 nm) and these wavelengths are associated with differences in particle size and color [45,51], both of

Table 3
Partial least squares regression (PLSR) model fit statistics for chemical composition data with the cross validation applied at the stand level; the best model for each analyte is highlighted in bold (lowest RMSECV).

	Absorbance			First derivative			Second derivative		
Cell wall polymer or analyte (%)	No. of factors	R^2_{cv}	RMSECV (%)	No. of factors	R^2_{cv}	RMSECV (%)	No. of factors	R^2_{cv}	RMSECV (%)
Cellulose – Diglyme (%)	6	0.97	1.3	**6**	**0.97**	**1.3**	5	0.97	1.5
Cellulose (%)	5	0.96	1.6	**5**	**0.96**	**1.6**	2	0.96	1.7
Hemicellulose (%)	3	0.73	2.9	2	0.74	2.9	**3**	**0.75**	**2.8**
Holocellulose (%)	3	0.92	3.8	3	0.92	3.7	**3**	**0.93**	**3.5**
Klason lignin (%)	6	0.92	1.9	**4**	**0.94**	**1.7**	3	0.92	1.9
Extractives (%)	5	0.95	2.3	**10**	**0.98**	**1.4**	9	0.97	1.7
Ash (%)	5	0.76	0.5	4	0.81	0.5	**4**	**0.81**	**0.5**
Glucan (%)	6	0.97	1.6	**5**	**0.97**	**1.6**	5	0.96	1.8
Xylan (%)	5	0.86	0.8	4	0.86	0.8	**4**	**0.89**	**0.7**
Galactan (%)	4	0.40	0.8	10	0.49	0.7	**8**	**0.65**	**0.6**
Arabinan (%)	9	0.52	0.6	7	0.54	0.6	**7**	**0.55**	**0.6**
Mannan (%)	3	0.64	1.7	**1**	**0.65**	**1.7**	3	0.61	1.7

Table 4
Partial least squares regression (PLSR) model fit statistics for chemical composition data with the training set (stands 2–5) and testing set (stand 1); the best model for each analyte is highlighted in bold (lowest RMSEP).

	First derivative					Second derivative				
Cell wall polymer or analyte (%)	No. of factors	R^2_{cv}	RMSECV (%)	R^2_p	RMSEP (%)	No. of factors	R^2_{cv}	RMSECV (%)	R^2_p	RMSEP (%)
Cellulose – Diglyme (%)	**4**	**0.96**	**1.6**	**0.98**	**1.2**	4	0.96	1.7	0.97	1.2
Cellulose (%)	3	0.96	1.6	0.94	1.8	**1**	**0.95**	**1.8**	**0.96**	**1.6**
Hemicellulose (%)	**3**	**0.74**	**2.6**	**0.81**	**2.9**	3	0.74	2.6	0.79	3.0
Holocellulose (%)	**3**	**0.93**	**3.4**	**0.94**	**3.4**	3	0.93	3.4	0.94	3.5
Klason lignin (%)	**4**	**0.94**	**1.7**	**0.90**	**1.4**	4	0.93	1.9	0.90	1.4
Extractives (%)	3	0.95	2.3	0.97	1.8	**7**	**0.96**	**2.1**	**0.99**	**1.0**
Ash (%)	4	0.87	0.4	0.45	0.5	**4**	**0.87**	**0.4**	**0.51**	**0.5**
Glucan (%)	**3**	**0.96**	**1.8**	**0.95**	**1.9**	3	0.96	1.9	0.95	1.9
Xylan (%)	7	0.93	0.6	0.90	0.6	**7**	**0.94**	**0.5**	**0.95**	**0.5**
Galactan (%)	10	0.77	0.5	0.85	0.4	**8**	**0.80**	**0.5**	**0.75**	**0.5**
Arabinan (%)	7	0.58	0.6	0.40	0.6	**5**	**0.54**	**0.6**	**0.34**	**0.7**
Mannan (%)	**1**	**0.66**	**1.5**	**0.60**	**2.1**	3	0.66	1.5	0.60	2.1

which can be quite different for forest residues. Three of the four wavelengths were the same for cellulose - diglyme and HPLC cellulose (966, 1128, 1339 nm), which again illustrates the similarities in the measurements.

Fig. 8 shows the predictions of cellulose (HPLC) and lignin from one sample of the mixed fraction in the 7 mm size class. The top image is the pseudo color image, note that pieces of bark appear more blue in color due to the wavelengths selected for the pseudo color image. The middle image is the cellulose prediction, and here both bark pieces as well as needles have low predicted cellulose content. Lignin content is shown at the bottom of the figure and here bark pieces have high lignin content in contrast to the branches, needles and wood. Earlywood and latewood can be differentiated from each other on some of the wood chips based on the predicted values. Additionally, the predicted values vary with the different surfaces (transverse vs radial or tangential) on some of the wood chips. A model that was focused on accurate prediction of individual chips would need further work to be accurate as the transverse surfaces of the wood chips sometimes appear to show a high cellulose (>45 %) and lignin content (>50 %) which is not realistic.

4. Discussion

In this paper, we developed chemometric models using a push-broom NIR-HSI system for forest residues with both pure forest residues of bark, branches, needles, and wood, as well as mixed fractions which contained a mixture of the different forest residues. To our knowledge, this is the first paper that uses NIR-HSI to predict chemical properties on forest residues that contain the major forest residue fractions produced by trees.

There are a wide variety of papers published that utilize NIR and NIR-HSI to predict properties from different energy crops [61–64] and forest residues. For bark, Schimleck and Yazaki [65] used a benchtop NIR system to predict extractives of the bark in radiata pine (*Pinus radiata*) and found good relationships (standard error of predictions (SEP) ranged from 1.5 to 2.1 %). Also working with radiata pine, Bridson et al. [66] used a benchtop NIR system to predict the hot water extractives (RMSECV 2.6 %), ash (RMSECV = 0.4 %), gross calorific value (RMSECV = 0.2 MJ/kg), and total phenolics (RMSECV = 39 mg/g) from bark and found good relationships. So and Eberhardt [67] used a benchtop NIR system to predict the amount of inner bark to total bark content and found reasonable results (RMSEP = 19.5 %) which could then be used to predict chemical components. A literature search on branches and NIR resulted in no results for chemical properties, however Barotto et al. [68] utilized a benchtop NIR on *Eucalyptus globulus* and *Eucalyptus viminalis* branches to predict anatomy, hydraulic, and physical properties. For both bark and branches, we were not able to find any studies using NIR-HSI to predict chemical properties, however Xue et al. [69] used NIR-HSI for species identification using bark. Considerably more NIR studies have been conducted on needles, with most work being done using NIR-HSI systems in remote sensing applications to study plant stress as it relates to water content [70,71], nutrient content [72–74], chlorophyll concentration [75], blight [76], and as a surrogate to fusiform rust resistance [77].

With respect to wood, considerably more research has been conducted using NIR and to a lesser extent NIR-HSI for chemical property variability [78–81]. Schimleck and Tsuchikawa [82] provide a comprehensive summary of studies utilizing NIR, and Schimleck et al. [78] provide a similar study of NIR-HSI applications. Axrup et al. [83] scanned wood samples on a moving conveyer belt with a low cost NIR diode system and concluded it could be used to monitor material properties. Specific to NIR-HSI, Thumm et al. [13] mapped the chemical composition of radiata pine using NIR-HSI from samples collected from

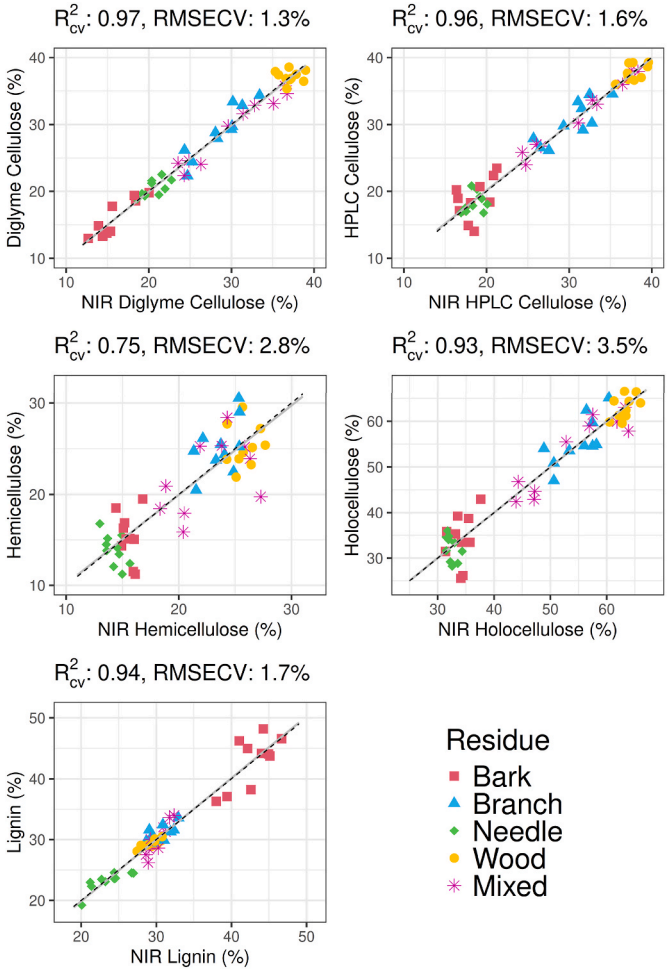

Fig. 5. Measured vs partial least squares regression (PLSR) model predicted diglyme-HCl cellulose, HPLC cellulose, hemicelluloses, holocellulose, and total lignin using spectral data collected from samples of different residue types. The dashed black line denotes the line of equivalence.

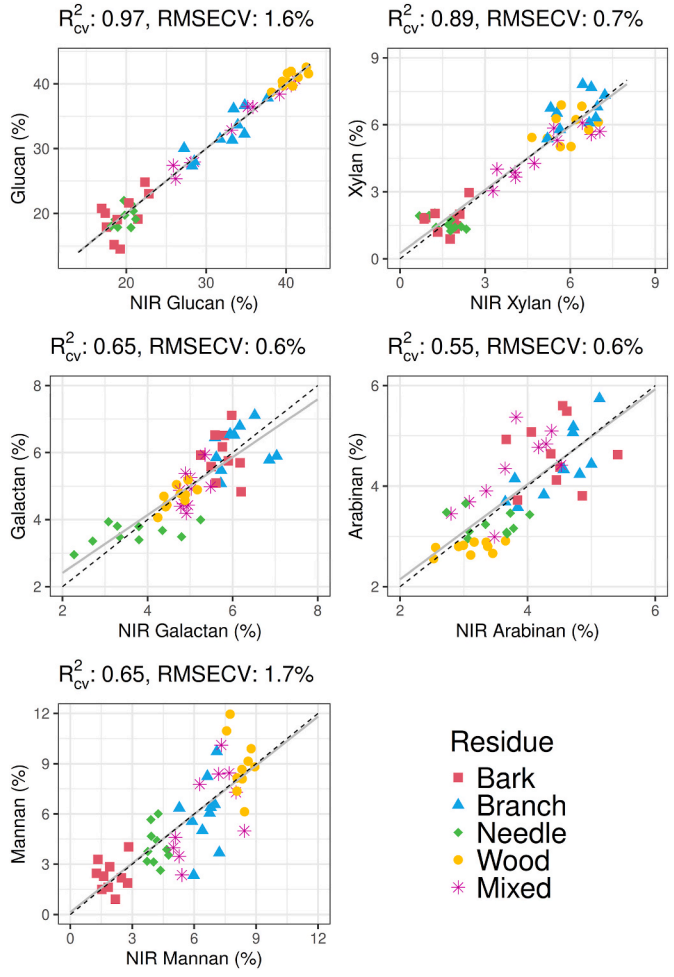

Fig. 6. Measured vs partial least squares regression (PLSR) model predicted glucan, xylan, galactan, arabinan, and mannan content using spectral data collected from samples from different residue types. The dashed black line denotes the line of equivalence.

disks. They found strong relationships for Klason lignin ($R^2_p = 0.78$, SEP = 0.96 %) and reasonable relationships for cellulose, glucose, and some of the hemicellulose sugars. For loblolly pine wood and using a benchtop NIR system, Acquah et al. [12] measured the same chemical properties as reported here and found similar relationships, with some models having better performance while others having worse. Given the different forest residues and instruments used, comparisons between studies can be challenging.

The use of NIR-HSI systems in biomass facilities could significantly improve quality control procedures for a facility that utilizes forest residues. Current quality control procedures for facilities take several forms, with some form of visual inspection of loads happening to qualitatively evaluate the types of forest residues found within each load. Facilities may also take random samples and evaluate the presence or absence of different forest residues and their particle size. NIR-HSI could be substituted for these tasks such that all loads entering a facility can be screened and their properties predicted, loads that do not meet standards could be rejected, or their price per unit be decreased. A major technical barrier for scanning loads entering a facility would be handling both closed top and open top trailers, and the variability in the height to which material is loaded within a trailer. Unlike a conveyor within a plant where the height of the material can be reasonably controlled, material within a trailer will have a wide variability in the surface topography, and thus it would be important to determine which materials within the image are reasonably within the depth of field. This

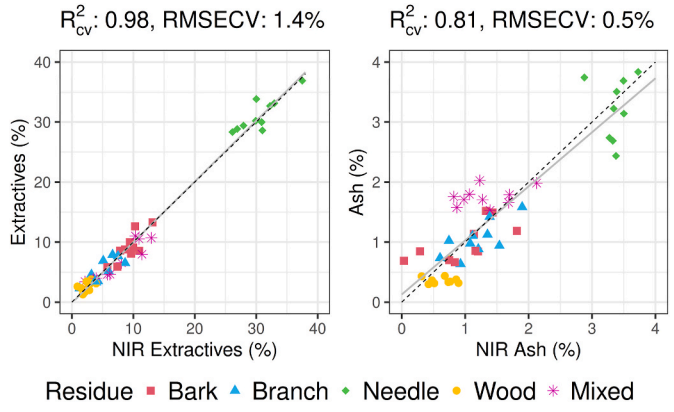

Fig. 7. Measured vs partial least squares regression (PLSR) model predicted extractives and ash content using spectral data collected from samples from different residue types. The dashed black line denotes the line of equivalence.

could be accomplished by measuring the surface depth and combining the measurements with the NIR-HSI to determine particles in and out of focus.

Once material enters a facility the most common method for sorting is with respect to size. For example, pulp mills screen wood chips for size

Table 5
Wavelengths of importance for the partial least squares regression (PLSR) models determined with variable importance on projection (VIP). Bond vibration and associated cell wall polymer (- means not identified) if identified [45, 51].

Wavelength	Bond Vibration	Cell wall polymer or analyte	Properties
966	O-H str. second overtone	–	Cellulose (both), Glucan, Mannan
1114	–	–	Holocellulose, Xylan
1128	–	–	Cellulose (both), Extractives, Glucan, Mannan
1170	C-H str. second overtone	Lignin	Lignin
1177	–	–	Cellulose – Diglyme, Mannan
1339	–	–	Cellulose (both), Extractives, Glucan, Mannan
1346	1st OT C-H str. + C-H def.	Hemicellulose & All	Lignin
1364	1st OT C-H str. + C-H def.	Cellulose	Holocellulose, Ash, Xylan
1396	–	–	Arabinan
1403	–	–	Ash, Xylan, Galactan
1417	1st OT C-H str. + C-H bend	Lignin	Lignin
1428	1st OT O-H str. + H20	Cellulose & Water	Extractives, Arabinan
1435	1st OT O-H str.	Water	Galactan
1445	1st OT O-H str.	Lignin & Extractives	Holocellulose, Ash, Xylan
1672	1st OT Car-H str.	Lignin	Holocellulose, Ash, Galactan, Arabinan,
1683	1st OT C-H str. & 1st OT Car-H str.	Hemicellulose & Lignin	Lignin, Extractives
1697	1st OT C-H str.	Lignin	Cellulose, Glucan, Galactan, Arabinan

with fines being burned while oversized material is rechipped to further reduce their size to acceptable dimensions [84]. While the larger sizes contain a higher amount of wood relative to needles and bark (Table 2), it is clear that separation by size is unlikely to provide forest residue rich fractions of materials and thus an air classification system could first be used to help separate bark and needles from branch and stem wood [85, 86]. For example, Saha et al. [86] sorted forest residues into a light and heavy sort using air classification and they found that the 'light' sort whereby low density material was screened contained most of the bark, needles, and consequently the ash. After air classification, NIR-HSI could be used to further segregate the material. What gets screened would ultimately depend on the products produced by the facility. Lestander et al. [87] used NIR-HSI (1000–2500 nm) to predict the extractives content in wood chips with the intent to demonstrate that chips with high extractives content could be identified using NIR-HSI and thus segregated such that they could be processed as a feedstock for chemical products. They measured extractives content from Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) wood chips and found strong relationships (RMSE 1.2 %). For forest residues, this same logic could be applied such that forest residues that have beneficial (e.g. high extractives content) or alternatively negative properties (e.g. high ash content) could be segregated.

5. Conclusions

Forest residues are widely available yet remain underutilized mostly due to their high compositional variability. Biomass facilities processing forest residues currently rely on visual load inspections and basic particle size measurements for quality control which results in inconsistent feedstock quality and inefficient processing. Near-infrared hyperspectral imaging as a rapid, non-destructive screening tool offers opportunities

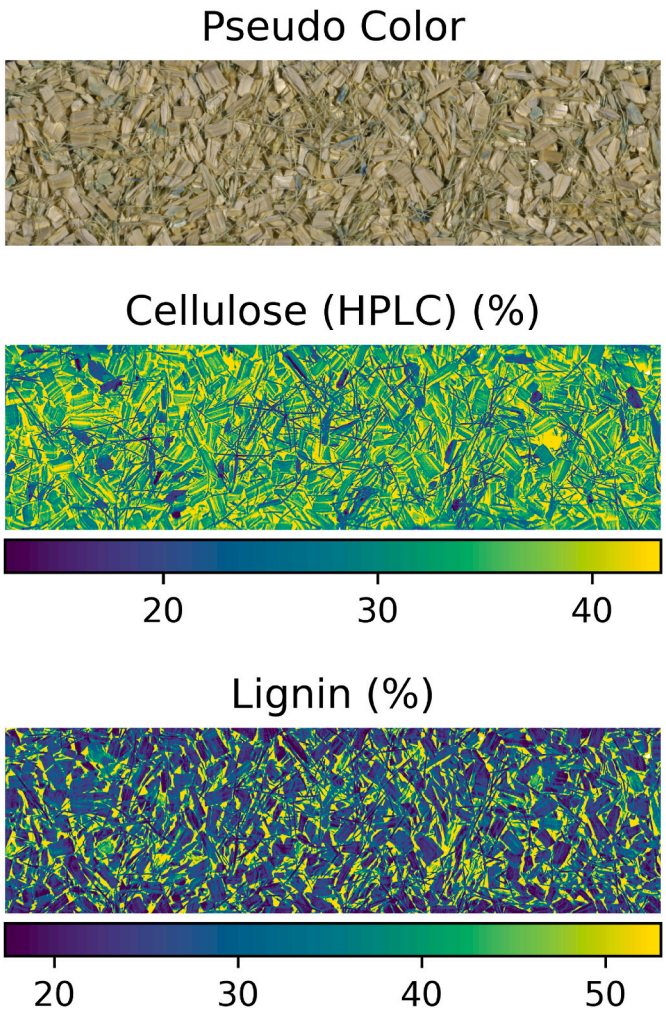

Fig. 8. Pseudo color image of a 7 mm size class from the mixed fraction (top image), and the predictions for cellulose measured via HPLC (middle image) and lignin content (bottom image). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

for real-time characterization of biomass whereby all loads entering a facility can be assessed for their physical and chemical properties and loads that do not meet standards could be rejected or their price per unit be decreased.

In this study we developed calibration models capable of rapid and precise estimation of chemical composition in loblolly pine harvest residues collected from different physiographic regions and age groups using NIR spectral data. Partial least squares regression models showed strong cross-validation fit statistics ($R^2_{cv} > 0.9$) for 6 out of 12 properties measured, and thus NIR-HSI can be a rapid and reliable tool for estimation of chemical properties of forest residues. Among the structural components, cellulose had the best cross-validation fit statistics followed by total lignin and holocellulose. The hemicellulose model was noticeably poorer as some individual sugars (glucan, xylan) had excellent or good predictive models while others (arabinan, galactan, and mannan) did not. The extractives model achieved the best cross-validation fit statistics across all measured properties, showing only minor variation in predicting extractives content in needles. The results provided here have important implications in rapid and real-time feedstock analysis in a facility utilizing southern pine harvest residues. Chemometric models such as the ones developed in this study can also be integrated with process parameters in plants to predict product yield and compositions. Future research could focus on expanding calibration datasets to capture regional and species-level variability in feedstocks as well as pairing

NIR-HSI with advanced machine learning and deep learning models to improve prediction capabilities or be used to segment individual particles.

CRediT authorship contribution statement

Sameen Raut: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Ighoyivwi Onakpoma:** Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Seyedehsan Vasefi:** Writing – review & editing, Data curation. **Joseph Dahlen:** Writing – original draft, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Laurence R. Schimleck:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Sudhagar Mani:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Gerald Presley:** Writing – review & editing, Methodology, Investigation, Data curation. **Thomas L. Eberhardt:** Writing – review & editing, Methodology, Data curation.

Acknowledgments

This work was partially supported by a cooperative agreement between the University of Georgia (UGA) and the US Department of Energy (DOE) Award Number: DE-EE0008911, and the NIFA McIntire-Stennis project 1023340.

Data availability

Data will be made available on request.

References

[1] F. Johnsson, J. Kjärstad, J. Rootzén, The threat to climate change mitigation posed by the abundance of fossil fuels, Clim. Policy 19 (2019) 258–274, https://doi.org/10.1080/14693062.2018.1483885.
[2] P. McKendry, Energy production from biomass (part 1): overview of biomass, Bioresour. Technol. 83 (2002) 37–46.
[3] H. Ghaderi, M.S. Pishvaee, A. Moini, Biomass supply chain network design: an optimization-oriented review and analysis, Ind. Crops Prod. 94 (2016) 972–1000, https://doi.org/10.1016/j.indcrop.2016.09.027.
[4] U.S. Department of Energy, U.S. billion-ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry, Oak Ridge National Laboratory, Oak Ridge, TN, USA, 2011, https://doi.org/10.1089/ind.2011.7.375. R.D. Perlack and B.J. Stokes (Leads), ORNL/TM-2011/224.
[5] U.S. Department of Energy, 2023 billion-ton Report: an Assessment of U.S. Renewable Carbon Resources. M.H. Langholtz (Lead), ORNL/SPR-2024/3103, Oak Rodge National Laboratory, Oak Ridge, TN, USA, 2024, https://doi.org/10.23720/BT2023/2316165.
[6] M.A.W. Hinchee, L.N. Mullinax, W.H. Rottmann, Woody biomass and purpose-grown trees as feedstocks for renewable energy, in: P.N. Mascia, J. Scheffran, J. M. Widholm, T. Nagata (Eds.), Plant Biotechnol. Sustain. Prod. Energy Co-products, Biotechnol. Agric. for., vol. 66, Springer-Verlag, Berlin Heidelberg, 2010, pp. 155–208, https://doi.org/10.1007/978-3-642-13440-1_7.
[7] C. Whittaker, N. Mortimer, R. Murphy, R. Matthews, Energy and greenhouse gas balance of the use of forest residues for bioenergy production in the UK, Biomass Bioenergy 35 (2011) 4581–4594, https://doi.org/10.1016/j.biombioe.2011.07.001.
[8] Z. Pásztory, I.R. Mohácsiné, G. Gorbacheva, Z. Börcsök, The utilization of tree bark, Bioresources 11 (2016) 7859–7888, https://doi.org/10.15376/biores.11.3.Pasztory.
[9] E. Sjöström, Wood Chemistry: Fundamentals and Applications, second ed., Academic Press, San Diego, CA, USA, 1993.
[10] P. Gujjula, N. Kumar, J.G. Lynam, Pretreatment of loblolly pine tree needles using deep eutectic solvents, Biomass 1 (2021) 1–10, https://doi.org/10.3390/biomass1010001.
[11] K. Sahoo, R. Bergman, S. Alanya-Rosenbaum, H. Gu, S. Liang, Life cycle assessment of forest-based products: a review, Sustainability 11 (2019) 4722.
[12] G.E. Acquah, B.K. Via, T. Gallagher, N. Billor, O.O. Fasina, L.G. Eckhardt, High throughput screening of elite loblolly pine families for chemical and bioenergy traits with near infrared spectroscopy, Forests 9 (2018) 418, https://doi.org/10.3390/f9070418.
[13] A. Thumm, M. Riddell, B. Nanayakkara, J. Harrington, R. Meder, Mapping within-stem variation of chemical composition by near infrared hyperspectral imaging, J. Near Infrared Spectrosc. 24 (2016) 605–616, https://doi.org/10.1255/jnirs.1206.
[14] A. Thumm, M. Riddell, B. Nanayakkara, J. Harrington, R. Meder, Near infrared hyperspectral imaging applied to mapping chemical composition in wood samples, J. Near Infrared Spectrosc. 18 (2010) 507–515, https://doi.org/10.1255/jnirs.909.
[15] N. Gierlinger, M. Schwanninger, B. Hinterstoisser, R. Wimmer, Rapid determination of heartwood extractives in Larix sp. by means of Fourier transform near infrared spectroscopy, J. Near Infrared Spectrosc. 10 (2002) 203–214, https://doi.org/10.1255/jnirs.336.
[16] L.R. Schimleck, F. Antony, C. Mora, J. Dahlen, Mapping and modeling within-tree variation for loblolly pine pulp yield and lignin content, SN Appl. Sci. 3 (2021) 468, https://doi.org/10.1007/s42452-021-04443-5.
[17] B.K. Via, C.L. So, T.F. Shupe, L.G. Eckhardt, M. Stine, L.H. Groom, Prediction of wood mechanical and chemical properties in the presence and absence of blue stain using two near infrared instruments, J. Near Infrared Spectrosc. 13 (2005) 201–212, https://doi.org/10.1255/jnirs.538.
[18] C.L. So, T.L. Eberhardt, Chemical and calorific characterisation of longleaf pine using near infrared spectroscopy, J. Near Infrared Spectrosc. 18 (2010) 417–423, https://doi.org/10.1255/jnirs.889.
[19] C.C. Fagan, C.D. Everard, K. McDonnell, Prediction of moisture, calorific value, ash and carbon content of two dedicated bioenergy crops using near-infrared spectroscopy, Bioresour. Technol. 102 (2011) 5200–5206, https://doi.org/10.1016/j.biortech.2011.01.087.
[20] T.A. Lestander, B. Johnsson, M. Grothage, NIR techniques create added values for the pellet and biofuel industry, Bioresour. Technol. 100 (2009) 1589–1594, https://doi.org/10.1016/j.biortech.2008.08.001.
[21] M.W. Ahmed, V. Singh, M. Kamruzzaman, Near-infrared spectroscopy as a green analytical tool for sustainable biomass characterization for biofuels and bioproducts: an overview, Bioresour. Technol. 433 (2025) 132722, https://doi.org/10.1016/j.biortech.2025.132722.
[22] X. Feng, C. Yu, X. Liu, Y. Chen, H. Zhen, K. Sheng, Y. He, Nondestructive and rapid determination of lignocellulose components of biofuel pellet using online hyperspectral imaging system, Biotechnol. Biofuels 11 (2018) 88, https://doi.org/10.1186/s13068-018-1090-3.
[23] L. Pitak, K. Laloon, S. Wongpichet, P. Sirisomboon, J. Posom, Machine learning-based prediction of selected parameters of commercial biomass pellets using line scan near infrared-hyperspectral image, Processes 9 (2021) 316, https://doi.org/10.3390/pr9020316.
[24] M. Mäkelä, P. Geladi, Hyperspectral near infrared imaging quantifies the heterogeneity of carbon materials, Sci. Rep. 8 (2018) 10442, https://doi.org/10.1038/s41598-018-28889-7.
[25] S. Raut, J. Dahlen, T.L. Eberhardt, L. Schimleck, D.M. Johnson, K.D. Klepzig, S. Mani, Light gradient-boosting machine model provides biologically realistic predictions of southern pine wood extractives content using near infrared spectroscopy hyperspectral imaging compared to artificial neural network and partial least squares regression m, J. Near Infrared Spectrosc. 33 (2025) 77–90, https://doi.org/10.1177/09670335251369048.
[26] Y. Jiang, C. Li, F. Takeda, Nondestructive detection and quantification of blueberry bruising using near-infrared (NIR) hyperspectral reflectance imaging, Sci. Rep. 6 (2016) 35679, https://doi.org/10.1038/srep35679.
[27] M.S. Kim, Y.-R. Chen, B.-K. Cho, K. Chao, C.-C. Yang, A.M. Lefcourt, D. Chan, Hyperspectral reflectance and fluorescence line-scan imaging for online defect and fecal contamination inspection of apples, Sens. Instrum. Food Qual. Saf. 1 (2007) 151–159, https://doi.org/10.1007/s11694-007-9017-x.
[28] A. Vicens-Sans, X. Pascari, F. Molino, A.J. Ramos, S. Marín, Near infrared hyperspectral imaging as a sorting tool for deoxynivalenol reduction in wheat batches, Food Res. Int. 178 (2024) 113984, https://doi.org/10.1016/j.foodres.2024.113984.
[29] Y. Lu, W. Saeys, M. Kim, Y. Peng, R. Lu, Hyperspectral imaging technology for quality and safety evaluation of horticultural products: a review and celebration of the past 20-year progress, Postharvest Biol. Technol. 170 (2020) 111318, https://doi.org/10.1016/j.postharvbio.2020.111318.
[30] Y. Jeon, W. Seol, S. Kim, K.S. Kim, Robust near-infrared-based plastic classification with relative spectral similarity pattern, Waste Manag. 166 (2023) 315–324, https://doi.org/10.1016/j.wasman.2023.04.033.
[31] R. Schultz, Loblolly-the pine for the twenty-first century, New For 17 (1999) 71–88.
[32] S. McKeand, T. Mullin, T. Byram, T. White, Deployment of Genetically Improved Loblolly and Slash Pines in the South, J. for, vol 101, 2003, pp. 32–37, https://doi.org/10.1093/jof.101.3.32.
[33] P. Jernigan, T. Gallagher, D. Mitchell, M. Smidt, L. Teeter, High tonnage harvesting and skidding for loblolly pine energy plantations, For. Prod. J. 66 (2016) 185–191, https://doi.org/10.13073/FPJ-D-14-00055.
[34] TAPPI, Useful Method UM-21: Sieve Analysis of Pulpwood Chips, Peachtree Corners, GA, USA, 1984.
[35] SCAN-CM 40, 01, Wood Chips for Pulp Production. Size Distribution, Scandinavian Pulp, Paper and Board Testing Committee, Stockholm, Sweden, 2001.
[36] L. Jordan, A. Clark, L.R. Schimleck, D.B. Hall, R.F. Daniels, Regional variation in wood specific gravity of planted loblolly pine in the United States, Can. J. For. Res. 38 (2008) 698–710, https://doi.org/10.1139/X07-158.
[37] P. Raybaut, Spyder-documentation, Available from, https://www.spyder-ide.org, 2009. (Accessed 10 May 2023).
[38] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus,

S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature 585 (2020) 357–362, https://doi.org/10.1038/s41586-020-2649-2.
[39] G. Bradski, The OpenCV library, Dr. Dobb's J. Softw. Tool. 25 (2000) 122–125.
[40] W. McKinney, Data structures for statistical computing in Python, in: Proc. 9th Python Sci. Conf. June 28-July 3, 2010, Austin, Texas, USA, 2010, pp. 56–61, https://doi.org/10.25080/majora-92bf1922-00a.
[41] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. Pietro Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C.N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D.A. Nicholson, D.R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G.A. Price, G.L. Ingold, G.E. Allen, G.R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J.T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J.V. de Miranda Cardoso, J. Reimer, J. Harrington, J.L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N.J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P.A. Brodtkorb, P. Lee, R.T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T.J. Pingel, T.P. Robitaille, T. Spura, T.R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y.O. Halchenko, Y. Vázquez-Baeza, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (2020) 261–272, https://doi.org/10.1038/s41592-019-0686-2.
[42] T. Boggs, Spectral Python (SPy), Available from, https://www.spectralpython.net/, 2014. (Accessed 10 May 2023).
[43] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, New York, NY, USA, 1982.
[44] A. Savitzky, M.J. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639.
[45] B.G. Osborne, T. Fearn, P.H. Hindle, Practical NIR Spectroscopy with Applications in Food and Beverage Analysis, Longman Scientific and Technical, Harlow, U.K., 1993.
[46] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, D. Crocker, Determination of Structural Carbohydrates and Lignin in Biomass. Laboratory Analytical Procedure (LAP), National Renewable Energy Laboratory, U.S. Department of Energy, Golden, CO, USA, 2012. NREL/TP-510-426-18.
[47] D.B. Easty, E.W. Malcom, Estimation of pulping yield in continuous digesters from carbohydrate and lignin determinations, Tappi J. 65 (1982) 78–80.
[48] L.E. Cullen, C. Macfarlane, Comparison of cellulose extraction methods for analysis of stable isotope ratios of carbon and oxygen in plant material, Tree Physiol. 25 (2005) 563–569, https://doi.org/10.1093/treephys/25.5.563.
[49] ASTM International, D7582-24, Standard Test Methods for Proximate Analysis of Coal and Coke by Macro Thermogravimetric Analysis, ASTM International, West Conshohocken, PA, USA, 2024.
[50] ASTM International, D1102-84. Standard Test Method for Ash in Wood, ASTM International, West Conshohocken, PA, USA, 2021.
[51] M. Schwanninger, J.C. Rodrigues, K. Fackler, A review of band assignments in near infrared spectra of wood and wood components, J. Near Infrared Spectrosc. 19 (2011) 287–308, https://doi.org/10.1255/jnirs.955.
[52] R Core Team, R: a Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2023. Available from, https://www.R-project.org/. (Accessed 10 May 2023).
[53] Rstudio, Rstudio: Integrated Development for R, RStudio, Inc., Boston, MA, 2023. Available from, http://www.rstudio.com/. (Accessed 10 May 2023).
[54] H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse, J. Open Source Softw. 4 (2019) 1686, https://doi.org/10.21105/joss.01686.
[55] K.H. Liland, B.-H. Mevik, R. Wehrens, Pls: partial least squares and principal component regression, R package version 2.8-1. Available from, https://cran.r-project.org/package=pls, 2022. (Accessed 10 May 2023).
[56] H.W. Borchers, _pracma: practical numerical math functions_. R package version 2.4.6. https://CRAN.R-project.org/package=pracma, 2025. (Accessed 5 December 2025).
[57] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in partial least squares regression, Chemom. Intell. Lab. Syst. 118 (2012) 62–69.
[58] Signal Developers, signal: signal processing. http://r-forge.r-project.org/projects/signal/, 2013. (Accessed 10 May 2023).
[59] B. Auguie, gridExtra: miscellaneous functions for "grid" graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra, 2017. (Accessed 10 May 2023).
[60] E. Ostertagová, Modelling using polynomial regression, Procedia Eng. 48 (2012) 500–506, https://doi.org/10.1016/j.proeng.2012.09.545.
[61] C.D. Everard, K.P. McDonnell, C.C. Fagan, Prediction of biomass gross calorific values using visible and near infrared spectroscopy, Biomass Bioenergy 45 (2012) 203–211, https://doi.org/10.1016/j.biombioe.2012.06.007.
[62] J. Liu, S. Jin, C. Bao, Y. Sun, W. Li, Rapid determination of lignocellulose in corn stover based on near-infrared reflectance spectroscopy and chemometrics methods, Bioresour. Technol. 321 (2021) 124449, https://doi.org/10.1016/j.biortech.2020.124449.
[63] X. Jin, X. Chen, C. Shi, M. Li, Y. Guan, C.Y. Yu, T. Yamada, E.J. Sacks, J. Peng, Determination of hemicellulose, cellulose and lignin content using visible and near infrared spectroscopy in Miscanthus sinensis, Bioresour. Technol. 241 (2017) 603–609, https://doi.org/10.1016/j.biortech.2017.05.047.
[64] M.A. Sanderson, F. Agblevor, M. Collins, D.K. Johnson, Compositional analysis of biomass feedstocks by near infrared reflectance spectroscopy, Biomass Bioenergy 11 (1996) 365–370.
[65] L.R. Schimleck, Y. Yazaki, Analysis of Pinus radiata D. Don bark by near infrared spectroscopy, Holzforschung 57 (2003) 520–526, https://doi.org/10.1515/HF.2003.077.
[66] J.H. Bridson, A. Thumm, M. Cooke-Willis, S. Campion, K. Tutt, M. West, Comparison of near infrared and mid infrared spectroscopy for the prediction of Pinus radiata bark chemical properties relevant to a biorefinery, Biomass Bioenergy 185 (2024) 107235, https://doi.org/10.1016/j.biombioe.2024.107235.
[67] C.L. So, T.L. Eberhardt, Rapid analysis of inner and outer bark composition of Southern Yellow Pine bark from industrial sources, Holz als Roh- Werkst. 64 (2006) 463–467, https://doi.org/10.1007/s00107-006-0117-9.
[68] A.J. Barotto, A. Martínez-Meier, V. Segura, S. Monteoliva, J.P. Charpentier, J. Gyenge, A.S. Sergent, F. Millier, P. Rozenberg, M.E. Fernández, Use of near-infrared spectroscopy to estimate physical, anatomical and hydraulic properties of Eucalyptus wood, Tree Physiol. 43 (2023) 501–514, https://doi.org/10.1093/treephys/tpac132.
[69] X. Xue, Z. Chen, H. Wu, H. Gao, Identification of Guiboutia species by NIR-HSI spectroscopy, Sci. Rep. 12 (2022) 11507, https://doi.org/10.1038/s41598-022-15719-0.
[70] S. Dengel, J. Grace, T. Aakala, P. Hari, S.L. Newberry, T. Mizunuma, Spectral characteristics of pine needles at the limit of tree growth in subarctic Finland, Plant Ecol. Divers. 6 (2013) 31–44, https://doi.org/10.1080/17550874.2012.754512.
[71] Y. Zhang, A. Wang, J. Li, J. Wu, Water content estimation of conifer needles using leaf-level hyperspectral data, Front. Plant Sci. 15 (2024) 1428212, https://doi.org/10.3389/fpls.2024.1428212.
[72] R. Hiukka, A multivariate approach to the analysis of pine needle samples using NIR, Chemom. Intell. Lab. Syst. 44 (1998) 395–401, https://doi.org/10.1016/S0169-7439(98)00067-7.
[73] Y. Li, H. Sun, F. Tomasetto, J. Jiang, Q. Luan, Spectrometric prediction of nitrogen content in different tissues of slash pine trees, Plant Phenomic. 2022 (2022) 9892728, https://doi.org/10.34133/2022/9892728.
[74] T. Long, X.L. Che, W. Guo, Y. Lan, Z. Xie, W. Liu, J. Lv, Y. Long, T. Liu, J. Zhao, Visible-near-infrared hyperspectral imaging combined with ensemble learning for the nutrient content of Pinus elliottii × P. caribaea canopy needles detection, Front. For. Glob. Chang. 6 (2023) 1203626, https://doi.org/10.3389/ffgc.2023.1203626.
[75] I. Moorthy, J.R. Miller, T.L. Noland, Estimating chlorophyll concentration in conifer needles with hyperspectral data: an assessment at the needle and canopy level, Remote Sens. Environ. 112 (2008) 2824–2838, https://doi.org/10.1016/j.rse.2008.01.013.
[76] N. Coops, M. Stanford, K. Old, M. Dudzinski, D. Culvenor, C. Stone, Assessment of Dothistroma needle blight of Pinus radiata using airborne hyperspectral imagery, Phytopathology 93 (2003) 1524–1532, https://doi.org/10.1094/PHYTO.2003.93.12.1524.
[77] S. Lim-Hing, A.O. Conrad, C.R. Montes, K.J.K. Gandhi, K.G. Payn, T.D. Walker, C. Villari, Near-infrared spectroscopy as a high-throughput phenotyping method for fusiform rust resistance in loblolly pine, Plant Phenomic. 7 (2025) 100066, https://doi.org/10.1016/j.plaphe.2025.100066.
[78] L. Schimleck, T. Ma, T. Inagaki, S. Tsuchikawa, Review of near infrared hyperspectral imaging applications related to wood and wood products, Appl. Spectrosc. Rev. 58 (2022) 585–609, https://doi.org/10.1080/05704928.2022.2098759.
[79] S. Tsuchikawa, A review of recent near infrared research for wood and paper, Appl. Spectrosc. Rev. 42 (2007) 43–71, https://doi.org/10.1080/05704920601036707.
[80] S. Tsuchikawa, T. Inagaki, T. Ma, Application of near-infrared spectroscopy to forest and wood products, Curr. For. Reports 9 (2023) 401–412, https://doi.org/10.1007/s40725-023-00203-3.
[81] S. Tsuchikawa, M. Schwanninger, A review of recent near-infrared research for wood and paper (part 2), Appl. Spectrosc. Rev. 48 (2013) 560–587, https://doi.org/10.1080/05704928.2011.621079.
[82] L.R. Schimleck, S. Tsuchikawa, Application of NIR spectroscopy to wood and wood-derived products, in: E.W. Ciurczak, B. Igne, J. Workman, D.A. Burns (Eds.), Handb. Near-Infrared Anal, fourth ed., CRC Press, Boca Raton, FL, USA, 2021, pp. 759–780, https://doi.org/10.1201/b22513-42.
[83] L. Axrup, K. Markides, T. Nilsson, Using miniature diode array NIR spectrometers for analysing wood chips and bark samples in motion, J. Chemom. 14 (2000) 561–572, https://doi.org/10.1002/1099-128X(200009/12)14:5/6<561::AID-CEM608>3.0.CO;2-2.
[84] G. Smook, Handbook for Pulp and Paper Technologists, fourth ed., TAPPI Press, Peachtree Corners, GA, USA, 2016.
[85] J.A. Lacey, J.E. Aston, T.L. Westover, R.S. Cherry, D.N. Thompson, Removal of introduced inorganic content from chipped forest residues via air classification, Fuel 160 (2015) 265–273, https://doi.org/10.1016/j.fuel.2015.07.100.
[86] N. Saha, J. Klinger, T. Bhattacharjee, Y. Xia, V. Thompson, O.A. Oyedeji, J. Parks, M. Shahnam, Y. Xu, The effect of air separations on fast pyrolysis products for forest residue feedstocks, Fuel 375 (2024) 132572, https://doi.org/10.1016/j.fuel.2024.132572.
[87] T.A. Lestander, P. Geladi, S.H. Larsson, M. Thyrel, Near infrared image analysis for online identification and separation of wood chips with elevated levels of

extractives, J. Near Infrared Spectrosc. 20 (2012) 591–599, https://doi.org/10.1255/jnirs.992.