Research Article
doi: 10.1590/01047760202228012985
vol(28), 2022

CERNE

Near infrared spectroscopy for estimating properties of kraft paper reinforced with cellulose nanofibrils

Lidiane Costa Lima [1], Lívia Ribeiro Costa [1], Ana Márcia Macedo Ladeira Carvalho [3], Maria Lucia Bianchi [2], Paulo Ricardo Gherardi Hein ✤[1]

[1] Forestry Science Departament. Federal University of Lavras. Lavras, Brazil
[2] Chemistry Department. Federal University of Lavras. Lavras, Brazil
[3] Forestry Science Departament. Federal University of Viçosa. Viçosa, Brazil

TECHNOLOGY OF FOREST PRODUCTS

ABSTRACT

Background: The aim of this study was to investigate near infrared (NIR) spectroscopy ability to estimate nanofibril concentration, physical and mechanical properties of Kraft paper reinforced with cellulose nanofibrils (CNF). For this purpose, paper sheets were prepared by mixing unbleached *Eucalyptus* fibers and cellulose nanofibrils. Twelve treatments result from the combination of the nominal grammages 75, 85, 95 and 105 g m^{-2} with the nanofibril concentrations of 1, 5 and 10wt%, with 7 sheets per treatment. NIR spectra measured directly on paper specimens were correlated with physical and mechanical properties values obtained through conventional laboratory analyzes.

Results: Principal component analysis (PCA) revealed no separation among specimens related to nanofibril content. Partial least squares regression (PLS-R) models for estimating nanofibril content, tensile index, stretch and resistance to air passage yielded R^2cv ranging from 0.73 to 0.98. Partial least squares - discriminant analysis (PLS-DA) correctly classified up to 93% of the paper specimens both by grammage and nanofibril content using NIR spectra.

Conclusion: This approach appears to be suitable for predicting physical and mechanical properties of Kraft papers and can detect cellulose nanofibril content in the cellulose handsheets.

Keywords: Mechanical resistance. Nanocellulose. Machine learning. Papermaking. Quality control.

HIGHLIGHTS

Models were able to successfully estimate paper properties.
NIR spectroscopy can quickly evaluate paper handsheets quality.
NIR spectrum contains information about CNF content within the paper.
Tensile index, stretch and resistance to air passage can be estimated by NIR-based models.

LIMA, L.C., COSTA, L.R., CARVALHO, A. M. M. L., BIANCHI, M.L.; HEIN, P.R.G. Near Infrared Spectroscopy For Estimating Properties Of Kraft Paper Reinforced With Cellulose Nanofibrils. CERNE, v.28, n.1, e-102985, doi: 10.1590/01047760202228012985

✤Corresponding author e-mail: paulo.hein@ufla.br Submitted: 30/07/2021 Accepted: 10/05/2022

Open Access
CERNE
ISSN 0104-7760
UFLA UNIVERSIDADE FEDERAL DE LAVRAS

INTRODUCTION

The world paper and paperboard production in 2019 was 404 million tons (FAO, 2020) and Brazil continues to be a global reference in the production of cellulose pulp. In 2019, the country maintained its position as the second-largest producer, reaching 19.7 million tons of cellulose production. Of all production, 75% was exported, totaling 14.7 million tons (IBA, 2020). The paper quality control analyses can be time consuming and expensive depending on the specific instruments (Fardim et al., 2005) to measure tearing, bursting and tensile index, stretch, tenacity, modulus of elasticity, ring crush test and etc. Thus, in the past years the increasing search for fast, reliable and low cost solutions for paper characterization is understandable.

Among monitoring alternatives, near infrared (NIR) spectroscopy has emerged as a rapid-response analytical tool, with increasing acceptance from industries related to its applicability on the production line (Tsuchikawa and Kobori 2015). Its fast, non-destructive, non-invasive, minimum sample preparation and nearly universal application, which includes any molecule containing C H, N H, S H or O H bonds (Pasquini, 2003). These characteristics explain the large number of studies in several areas that have been carried out over the years involving the technique (Tsuchikawa and Schwanninger, 2013).

According to Pasquini (2018) some frequencies will not be absorbed in a given wavelength range while other will be partially or completely absorbed, forming a complex figure of absorption intensity versus wavelength that constitutes the unique absorption spectra of a substance or sample. NIR spectra must be associated with strong information from traditional laboratory analysis and chemometric tools in order to calibrate, validate and extract information from data, allowing the prediction of properties, estimates of substance content in a sample and identification of different samples in a group by their differences or similarities, among other applications (Naes et al., 2002).

Many studies have been developed to evaluate forest products such as wood (Tsuchikawa and Kobori, 2015; Costa et al., 2018a; Arriel et al., 2019; Rosado et al., 2019), engineered panels (Via, 2010; Belini et al., 2011; Huang et al., 2019), charcoal (Costa et al., 2018b) and pulp (Costa et al., 2019). In terms of cellulosic pulp, research involving NIR spectroscopy began with the estimation of pulp kappa number (Birkett and Gambino, 1989). Then, Wright et al. (1990) have predicted the pulp yield and cellulose content from NIR spectra, Wallbacks et al. (1991) have investigated changes in chemical composition of pulp samples during Kraft pulping from NIR information and Reiter et al. (1999) used the technique to control of pulp composition and chemical additives in the paper machine. Yan and Krishnagopalan (2003) applied this technique for real-time estimation of effective alkali, lignin, and dissolved organics in the cooking liquor.

NIR spectroscopy can also be suitable to detect changes in paper mechanical properties, since they are directly related to intra and intermolecular interactions present in the material (Fardim et al., 2005). According to Samyn et al. (2018), cellulose nanofibrils (CNF) has been applied to papermaking processes for producing stronger papers with lower grammages. Here, we hypothesizes that NIR can be sensible to variations in cellulose nanofibrils (CNF) content in the paper, since they interact with the matrix through hydrogen bonds. Few studies have investigated papers reinforced with CNF by NIR spectrsocopy. Costa et al. (2019) have reported PLS-R models to estimate cellulose pulp dryness based on NIR signature of cellulosic pulp pads and reported models able to predict dryness from their NIR data with an error of 2.5%. Viana et al. (2016) have established predictive models to estimate the crystallinity indices and tensile and burst strength of cellulosic and nanocellulosic films through NIR spectroscopy. Santos et al. (2009) reported predictive models for kappa number prediction using handsheets (R^2=0.877) and disaggregated pulp (R^2=0.801) and Models for viscosity prediction using handsheets (R^2=0.857) and disaggregated pulp (R^2=0.695). Tyson et al. (2010) compared carbohydrate content calibrations for four sample preparation methods (coarse, fine, and milled pulp, and handsheets). Coarse and fine pulp preparation methods consistently yielded the best calibration and prediction statistics so they recommended collecting spectra from coarse pulp for carbohydrate content analysis by NIR spectroscopy. Sandak et al. (2015) developed models for monitoring kinetics and estimating the biodegradability and mechanical properties of paper coniferous and reported that Principal component analysis (PCA) was used for analising NIR pectra of sheets of paper infested by *C. globosum* and clear clusters were observed at various degradation stages. The models for estimating breaking length of papers presented R^2 of 0.75 in test set validation.

These previous studies indicate that NIR spectroscopy coupled with multivariate analyses can be used for monitoring pulp and paper quality. To our acknowledgement, no study has used fast and reliable methods such as NIR spectroscopy to predict the technological properties of paper handsheets reinforced with CNF. The existence of predictive models of physical and mechanical properties of pulps and papers could open new perspectives for online and real-time control. Automated systems could detect material flaws and decision-making could be taken immediately. Thus, the aim of this study was to develop NIR spectroscopic models for estimating nanofibrils content, physical and mechanical properties of paper handsheets reinforced with cellulose nanofibrils. It is expected that models with good predictive performance can be industrially applied in product quality control.

MATERIAL AND METHODS

Nanofibrils production

Cellulose nanofibrils (CNF) were obtained from bleached Kraft commercial pulp of *Eucalyptus* sp. through mechanical fibrillation. The pulps are composed mainly of cellulose (84.1 ± 0.1), hemicelluloses (15.7 ± 0.3) and residual soluble lignin (0.2 ± 0.1) (Dias et al. 2019). The pulp was soaked overnight in deionized water, disintegrated in a mechanical stirrer (Fisatom 722) and fibrillated in a Masuko Supermascoloider grinder (MKCA-2J, Masuko Sangyo Ltd.), with rotation of 1500 rpm, pulp concentration of 2% (w/v) and 30 cycles with electric current kept between 4 and 6 A.

Paper sheets formation

Paper preparation

Unbleached Kraft pulp from *Eucalyptus* sp. was refined to a Schopper-Riegler of 21 degrees in order to dissociate the fibers but limiting the influence of refining on paper properties. The treatments were grammage and nanofibrils content (Table 1). The class column refers to the twelve treatments used in which there was a variation in the grammage and nanofibrils content of the papers. These values will serve as a reference base for the development of models associating these values with the spectra in the NIR.

The paper sheets were produced according to the standard T 205 sp-02 (TAPPI 2006a). The CNF (0.1 wt%) were sonicated at 450 W during ten minutes, and then mixed with the refined unbleached cellulose pulp in a homogenizer, respecting the quantities as shown in Table 1. The mixture was placed in a paper forming machine and underwent a vacuum filtration process, forming the wet specimens that were pressed in an electronic press and dried at room temperature. The resultant paper sheets presented area of 0.0201 m^2, compound 12 treatments and 7 paper sheets per treatment, totaling 84 specimens with nominal grammages of 75, 85, 95 and 105 g m^{-2} and nanofibrils content of 1, 5 and 10% in relation to the paper grammage (wt%). For this, the volume of the nanofiber solution to be poured over the fiber solution was complete for each treatment and the resulting solution contained 1, 5 and 10% nanofiber concentration. It is important to highlight that nominal grammage is different to actual (final) grammage of the evaluated paper samples as the size of CNF is much smaller than fiber and 100% of CNF retention was not reached.

Paper characterization

The tests performed on the specimens were conducted in a controlled room with a relative humidity of 50 ± 2% and a temperature of 23 ± 1°C (ISO 1990). Paper characterization (physico-mechanical properties) was measured according to standard TAPPI methods. The evaluated properties and their respective standards are listed in Table 2.

Table 1. Treatments code, nominal grammage, nanofibrils content and the class of the Kraft paper sheets based on treatments.

Treatment	Nominal grammage (g. m^{-2})	Nanofibrils content (%)	Unbleached pulp (g. m^{-2})	Nanofibrils (g. m^{-2})	Class
105-1		1	103.95	1.05	1
105-5	105	5	99.75	5.25	2
105-10		10	94.5	10.5	3
95-1		1	94.05	0.95	4
95-5	95	5	90.25	4.75	5
95-10		10	85.5	9.5	6
85-1		1	84.15	0.85	7
85-5	85	5	80.75	4.25	8
85-10		10	76.5	8.5	9
75-1		1	74.25	0.75	10
75-5	75	5	71.25	3.75	11
75-10		10	67.5	7.5	12

Table 2. Standards for characterization of the paper handsheets.

Properties	Standards
Resistance to air passage (RAP)	T536 om-07 (Tappi 2007a)
Tearing index (TR)	T414 om-04 (Tappi 2004)
Bursting index (BI)	T403 om-02 (Tappi 2002)
Tensile index (TI), Stretch (ST), Tenacity (TEA) and Modulus of elasticity (MOE)	T494 om-06 (Tappi 2006b)
Ring crush test (RCT) (N. m^{-1})	T 822 om-02 (Tappi 2007b)
Corrugating medium test (CMT) (N)	T 809 om-99 (Tappi 2006c)

NIR spectra acquisition

Near infrared spectra were recorded using a Bruker spectrometer (model MPA, Bruker Optik GmbH, Ettlingen, Germany) in diffuse reflectance mode, based on a Fourier transform and equipped with an integrating sphere. Spectral analysis was performed within the 12500-3600 cm^{-1} range, at 3.87 cm^{-1} resolutions (each spectrum consisted of 1300 absorption values), as described in Costa et al. (2019). However, wavelengths from 9000 to 4000 cm^{-1} were selected for developing predictive calibrations. Each NIR spectrum was obtained with 32 scans, means were calculated and compared to the sintered gold standard used as background to obtain the absorption spectrum of the sample. Two NIR spectra were recorded for each paper specimen (rough surface). The spectrometer was connected to a computer that stored the spectra data collected by means of the OPUS program, Version 7.5.

Multivariate data analysis

Principal Component Analysis (PCA) was carried out to previously explore the data and to evaluate the dependence of the data by means of clusters. The PCAs were calculated using up to eight latent variable (LV).

Partial Least Squares Regressions (PLS-R) were adjusted based on NIR spectra and laboratory determined nanofibril content (in terms of input addition level of CNF), physical and mechanical properties, forming the calibration set. The PLS-R calibrations were cross-validated by the leave-one-out method and the maximum number of eight latent variable (LV) were used.

The predictive models were calibrated from original spectra; the number of LV selected for each model was that which minimized the residual variance of the calibration and cross-validation. The PLS-R models were chosen based on the following statistics: coefficient of determination of cross-validation (R^2cv), root mean standard error of cross-validation (RMSEcv), ratio performance of deviation for cross validation (RPDcv) and latent variables (LV), as described in Costa et al. 2018a.

For the discriminant analysis of partial least squares (PLS-DA), the nanofibrils content was considered as a categorical variable, not showing continuous values, but discrete. In this study, the samples were grouped into twelve different classes, containing 1, 5 and 10% nanofibrils content in grammage 75, 85, 95 and 105 (g. m^{-2}) (Table 1). The values 0 or 1 were then assigned to all the samples in each class, and when the sample belonged to that category, the value 1 was assigned and when the sample did not belong to the category, the value 0 was assigned, thus defining the group to which a sample belongs. Preliminary PLS models were performed to estimate continuous values in each of twelve categories. The PLS-DA models were evaluated by the number and percentage of hits.

Calibrations for estimating all papersheet properties were performed from untreated and after first derivative (13-point filter and a second order polynomial), normalization, normal standard variate (SNV) and multiplicative scatter correction (MSC). The wavenumbers from 12500 to 3600 cm^{-1} were used for calibrations and cross-validations. The software The Unscrambler® (CAMO AS, Oslo, Norway, v. 9.7) was used for multivariate analysis.

Microscopic characterization

Field-emission gun scanning electron microscopy (FEG-SEM) was applied to the paper specimens and to a cellulose nanofibrils film produced by casting. It was investigate cellulose nanofibrils presence, paper surface and nanofibrils interaction with the paper fibers. The samples were placed on half-moon aluminum stubs, fixed with carbon tape, covered with gold and analyzed in a microscope model JSM-6701F (JEOL®).

RESULTS

Physical and mechanical properties of paper handsheets

Table 3 present the mean results for each paper handsheets property. The variation between minimum to maximum value for each paper trait (Delta, %) is very important for calibrations based on NIR signatures. The greater the variation between minimum and maximum values for a given property, the greater the chances of calibrating a model with good predictive performance. Only tensile index and modulus of elasticity had variation between extreme values lower than 100%. The combination of four levels of grammage and three levels of nanofibril content generated paper handsheets samples with strong variation in mechanical performance.

Figure 1 shows the FEG-SEM micrographs of the papers sheets. Images 1A and 1B represent treatment 75-10 rough and smooth surface, respectively, accompanied by the samples photographs. The different surfaces occur due to paper formation process, in which the material suffers vacuum filtration and pressing in contact with a flat surface, causing a paper surface to be smooth and the other rough.

Figure 1A shows the fibers accommodation and great porosity, that is reduced in the smooth face due to greater fiber collapse (Figure 1B). Images 1C and 1D represent treatments 75-5 and 105-10, respectively. The magnification allows the visualization of paper thickness, showing a much more compact and well formed paper with the increase in grammage and nanofibril content, both contributing to reduce porosity and improve fiber-fiber connection. Figures 1E and 1F represent treatments 105-10, 105-5, respectively. The higher amounts of nanofibrils are evidenced by the arrows, pointing to zones of nanofilm formation and aggregation between fibers by CNFs. Arrows at Image 1F point to nanofilms formed within the paper, with translucency evidenced by the fibers visible behind the films.

According to Börjesson and Westman (2015), the hydroxyl groups in the cellulose polymers can form hydrogen bonds between different cellulose polymers (intermolecular hydrogen bonds) or within the polymer itself (intramolecular hydrogen bonds). The intramolecular bonds give stiffness to the polymer chain, while the intermolecular bonds allow the linear polymers to form sheet structures, as can be seen in Figure 1F.

Table 3. Physical and mechanical properties of paper handsheets.

Properties	Mean	SD	Min	Max	Delta (%)
Resistance to air passage (s 100cm^{-3})	13.47	14.49	1.27	47.97	3677
Tearing index (TR) (mN m^{2})	716.10	166.70	416.00	1017.00	144.5
Bursting index (BI) (KPa m^{2})	163.75	48.77	82.90	264.88	219.5
Tensile index (TI) (Nm g^{-1})	43.28	7.48	31.23	56.88	82.13
Stretch (ST) (%)	2.09	0.41	1.33	2.94	121.0
Tensile energy absorption (TEA) (J m^{-2})	57.53	23.33	22.49	114.92	410.9
Modulus of elasticity (MOE) (MNm kg^{-1})	4.72	0.34	3.91	5.41	38.4
Ring crush test (RCT) (N m^{-1})	0.60	0.17	0.34	1.08	217.6
Corrugating medium test (CMT) (N)	88.51	26.81	38.17	159.10	316.8

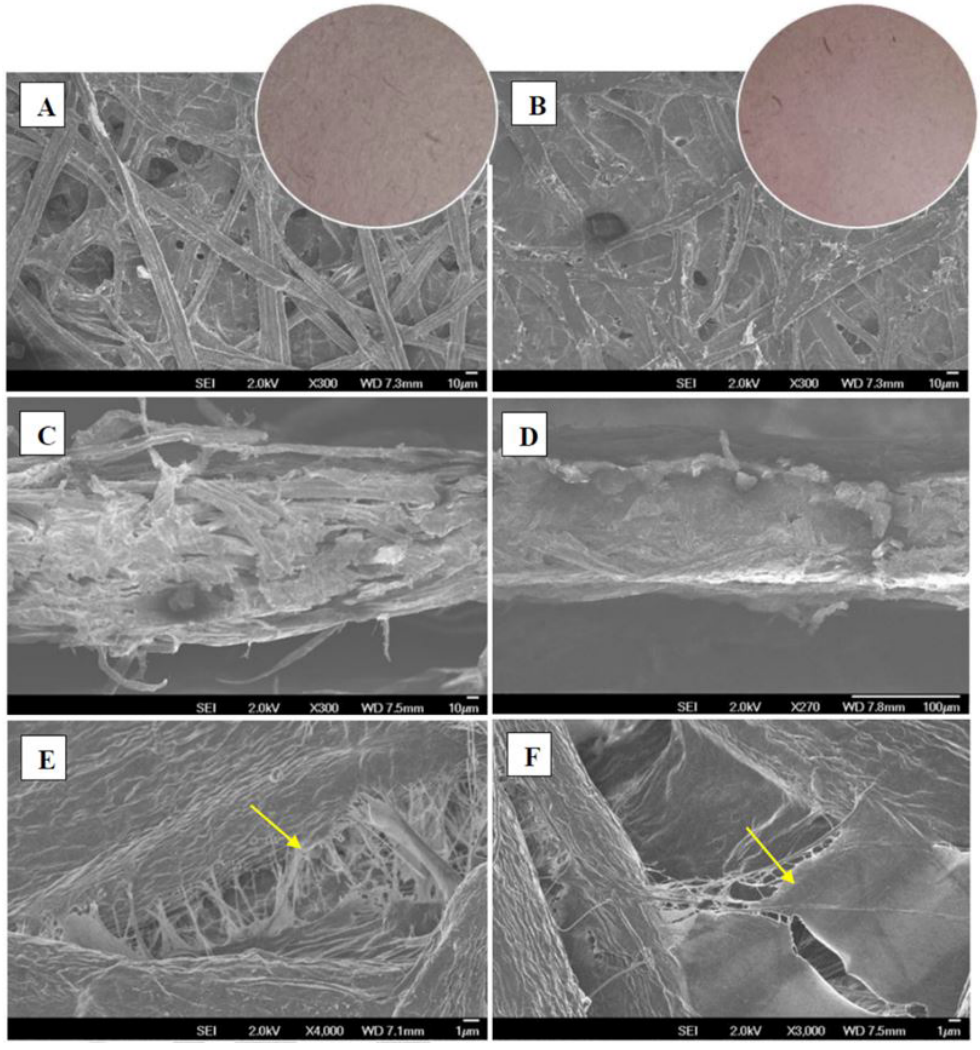

Figure 1. FEG-SEM micrographs of papers reinforced with cellulose nanofibrils showing paper surfaces rough (A) and smooth (B) with respective samples, paper thickness (C, D), formation of nanofilm within the paper (E, F). Arrows pointing to zones of nanofilm formation.

NIR spectra signatures

Figure 2 represents the spectra for the three nanofibril contents (1, 5 and 10%) and grammage (75, 85, 95 and 105 g m^{-2}), where each NIR spectrum represents the average of NIR spectra taken from several paper specimens. Figure 2C shows the first derivative spectra in the full range and Figure 2B, D and E shows detailed spectra ranges.

Principal component Analyses

Figure 3 presents the PCA performed to verify clusters with specimens presenting the same nanofibrils content. With the principal components 1 and 2 (PC 1 and PC 2) from the PCA of untreated NIR data it was possible to explain, respectively, 82 and 17% of the data variability, explaining 99% of the variance.

PCA data treated with Multiplicative Scatter Correction (MSC) were used in each nominal grammage separately (Figure 4), to avoid the problem of separation between papers reinforced with 1, 5 and 10% of CNF showed in Figure 3. According to Sandak et al. (2016), MSC is a signal processing algorithm that is particularly useful due to the non-linear scatter that is present in reflectance spectra.

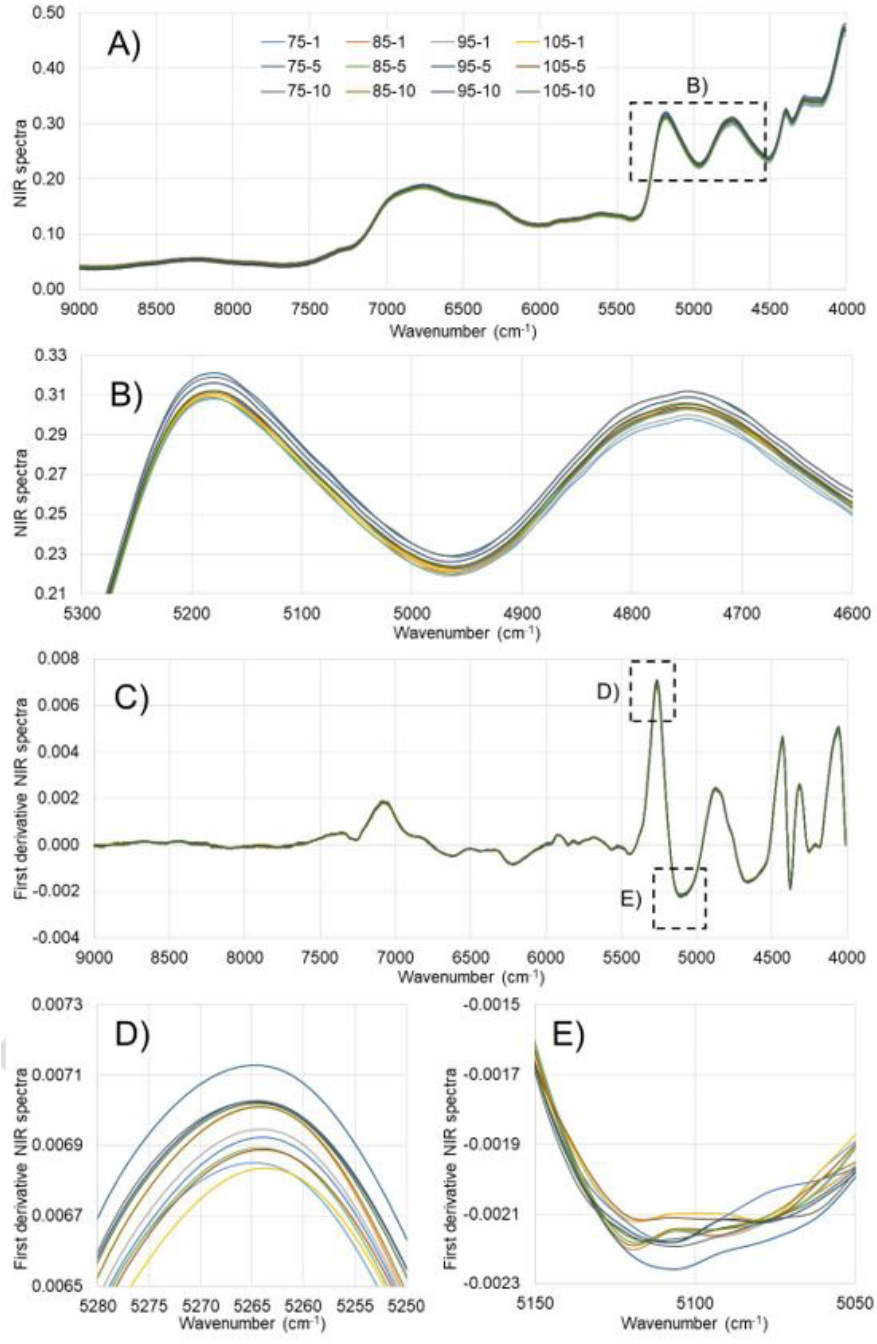

Figure 2. Mean raw (untreated, A) and first derivative (C) NIR spectra for each nanofibril content (1, 5 and 10%) and grammage (75, 85, 95 and 105 g m^{-2}) of Kraft paper handsheets reinforced with CNF and detailed spectra ranges (B, D and D).

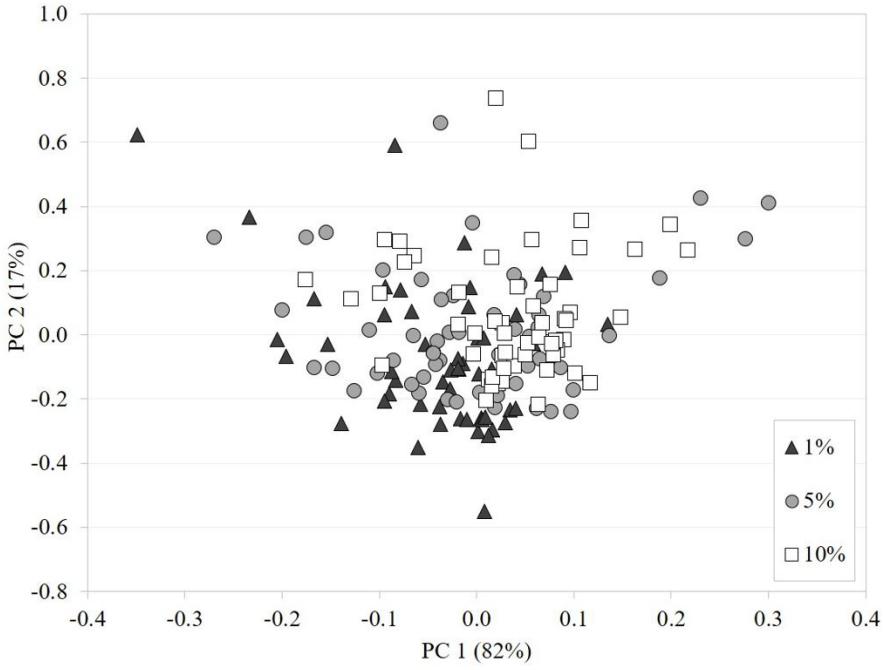

Figure 3. Scores of a PCA applied to untreated NIR spectra recorded from Kraft paper handsheets reinforced with CNF.

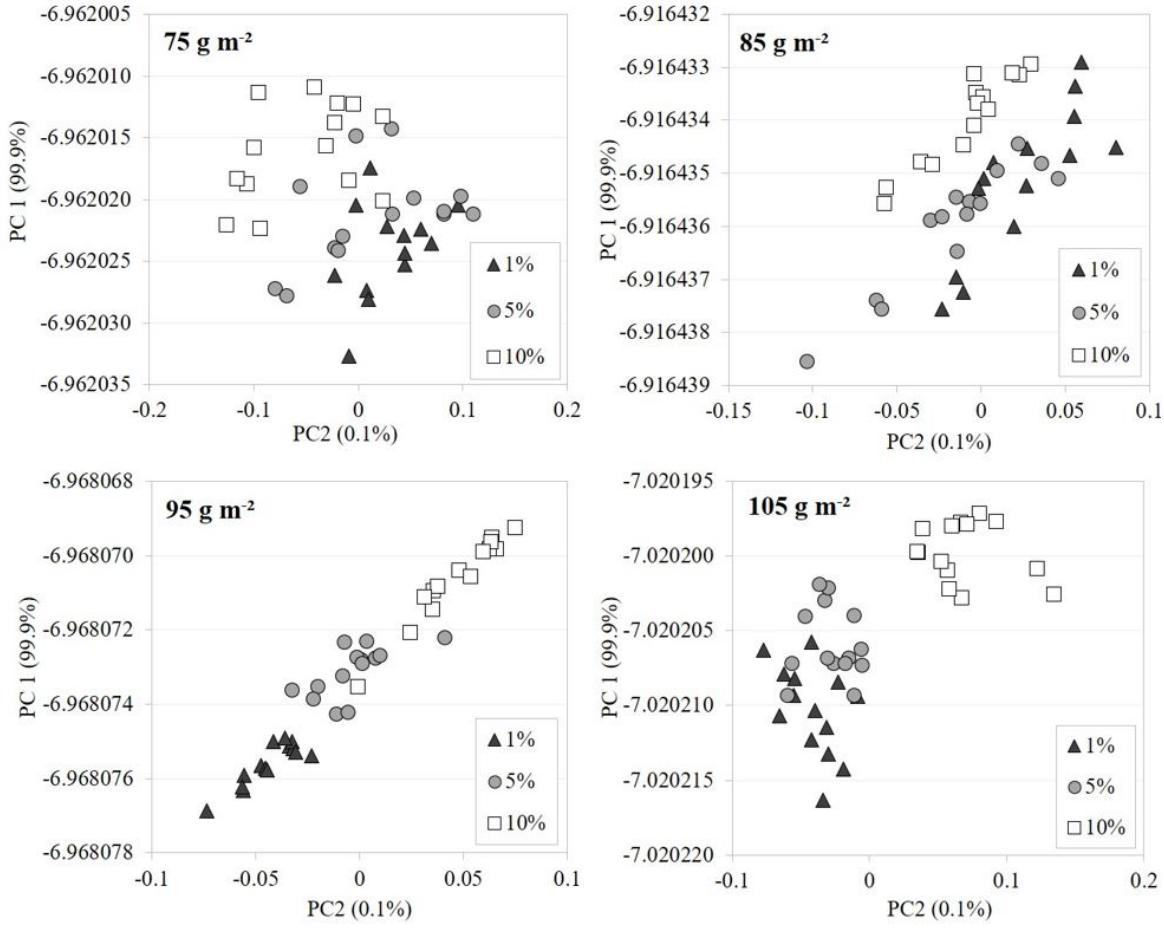

Figure 4. Scores of PCAs applied to NIR spectra after Multiplicative Scatter Correction recorded from Kraft paper handsheets reinforced with CNF with nominal grammage of 75, 85, 95 and 105 g. m^{-2}.

PLS-R and PLS-DA models to predict paper properties

The statistics associated to the predictive NIR-based models are listed in Table 4. The models were previously developed from NIR spectra after different mathematical treatments, but first derivative enhanced the quality of models, increasing their R^2 and reducing their RMSE (root mean standard error). Thus, models listed in Table 4 were developed from first derivative spectra. Figure 5 shows the plots of measured versus predicted values for the calibrations with higher R^2cv.

Partial Least Squares Discriminant Analysis (PLS-DA) was carried out in order to classify the paper sheets according to the twelve treatments studied; the main models are shown in Table 5. Classes were assigned according to the grammage and nanofibrils contents added in the handsheets.

Studies in the literature found better statistics when excluding the region of the spectrum from 12500 to 9000 cm^{-1} and 4000 to 3600 cm^{-1} for presenting noise and difficulty in acquiring information (Belini et al., 2011; Costa et al., 2018b; Arriel et al., 2019). Thus, this region was excluded and the analysis performed only in the region from 9000 to 4000 cm^{-1} (Table 5 - models 6 to 10).

The percentages of correct classifications for the calibration and also for the validation were satisfactory, with accuracy ranging from 72.0% to 97.0% and from 59.5% to 83.9%, respectively for calibration and validation. After calibration, validation must be carried out to validate and test the accuracy of the calibration, which is an essential step in the model construction (Gemperline, 2006). In cross-validation some specimens are separated, a model is built with the remaining samples and the prediction is made in relation to the initially separated specimens. Therefore, the statistics associated with validation are inferior to those of calibration.

In order to better explore model 3, Table 6 presents the confusion matrix that presents the same approach used in the models in Table 5 but includes the correct and incorrect classifications for each class by PLS-DA model.

The models in Table 5 are able to classify paper sheet samples according to their grammage and nanofibril content. Model 3 presented the best classification performance in calibration and therefore its classifications are detailed in Table 6. The calibration model 3 correctly classified 160 of the 168 paper samples (95.2% of hits). Only eight (8) of the 168 samples were misclassified: the model misclassifying only three (3) samples in terms of NCF concentration but correct grammage (highlighted in dark gray); Four (4) samples were incorrectly classified for grammage but at the correct NCF concentration (highlighted in bold) and only 1 sample was incorrectly classified for both weight and NCF concentration (highlighted in italics and underlined). In most cases, the error referred to only one level above or below the correct category.

Table 4. Calibration and cross-validation of PLS-R models to estimate nanofibrils concentration, nominal grammage, physical and mechanical properties of the Kraft paper sheets reinforced with cellulose nanofibrils based on first derivative NIR spectra.

Properties	R^2c	RMSEc	R^2cv	RMSEcv	LV	RPDcv
Nanofibrils concentration (%)	0.98	0.01	0.98	0.01	8	3.70
Nominal grammage (g. m^{-2})	0.68	6.35	0.59	7.18	8	1.57
Resistance to air passage (s $100cm^{-3}$)	0.79	6.55	0.73	7.52	6	1.93
Tearing index (TR) (mN. m^2)	0.63	100.9	0.29	139.8	7	1.19
Bursting index (BI) (KPa. m^2)	0.75	24.46	0.65	28.74	8	1.70
Tensile index (TI) (Nm. g^{-1})	0.93	1.90	0.91	2.27	8	3.29
Stretch (ST) (%)	0.80	0.18	0.74	0.21	7	1.94
Tensile energy absorption (TEA) (J. m^{-2})	0.70	12.72	0.62	14.76	7	1.58
Modulus of elasticity (MOE) (MNm. kg^{-1})	0.68	0.19	0.60	0.22	5	1.55
Ring crush test (RCT) (N. m^{-1})	0.53	0.10	0.15	0.13	8	1.32
Corrugating medium test (CMT) (N)	0.55	16.42	0.17	22.27	8	1.20

R^2c - coefficient of determination for calibration; RMSEc – root mean square error for calibration; R^2cv - coefficient of determination for cross validation; RMSEcv – root mean square error for cross validation, LV – latent variables.

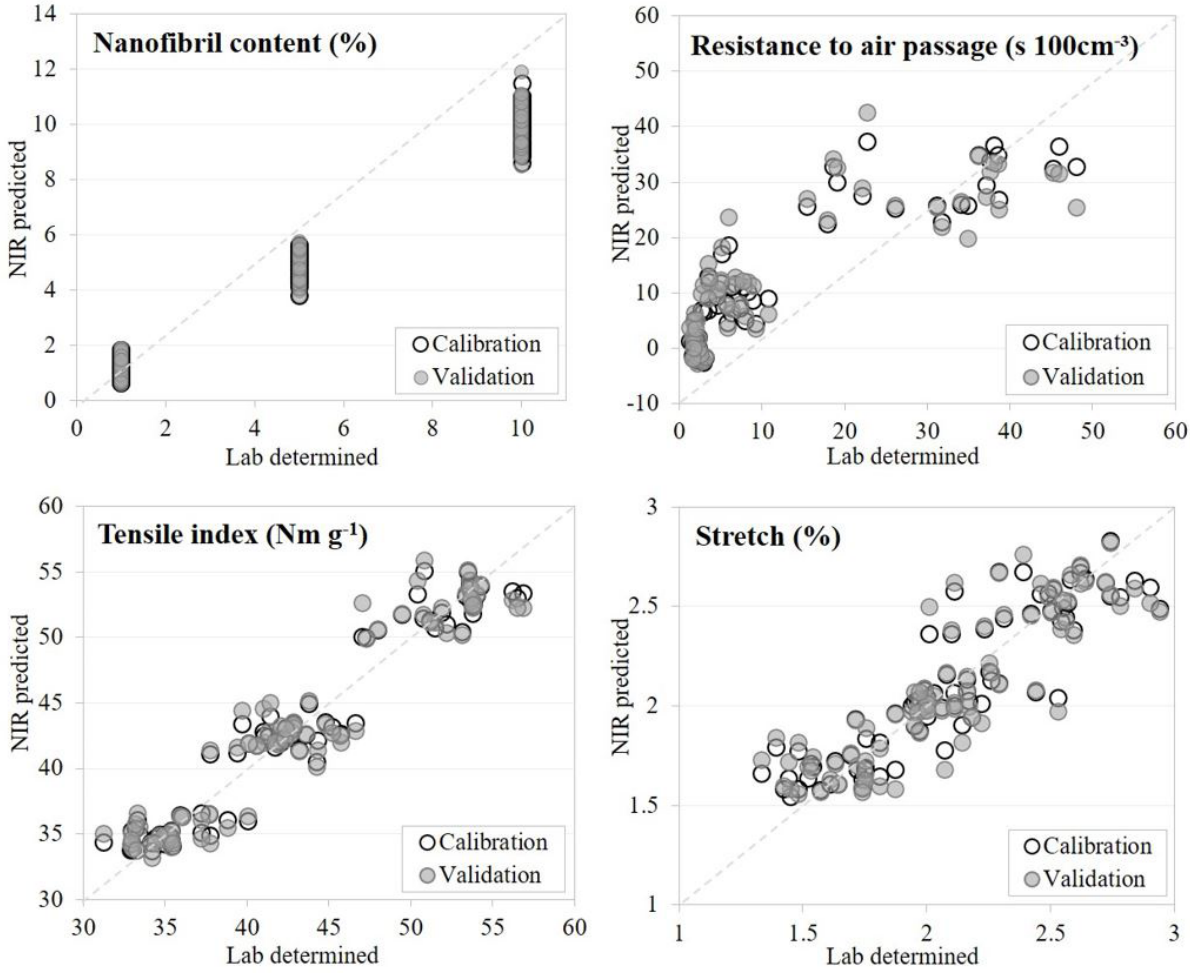

Figure 5. Lab determined versus NIR Spectroscopy predicted values to estimate physical and mechanical properties of paper reinforced with CNF.

Table 5. PLS-DA models developed from NIR signatures for classifying the papers in classes according to their grammage and nanofibrils content, including percentage (%) of correct classifications in calibration and cross-validation.

Model	Wavenumbers (cm^{-1})	Math. treat.	LV	Correct classification (%)	
				Calibration	Cross-validation
1		No	11	91.7	79.2
2		Norm	10	94.6	75.0
3	12500 - 3600	1D	10	95.2	83.9
4		SNV	10	97.0	82.7
5		MSC	10	97.0	83.3
6		No	9	72.0	61.9
7		Norm	10	75.0	59.5
8	9000 - 4000	1D	10	94.0	79.2
9		SNV	10	82.7	65.5
10		MSC	10	82.1	66.1

Table 6. Confusion matrix for estimating the paper class by PLS-DA model calibration based on the NIR signature.

Classification (reference)	Classification by NIR and PLS-DA												Correct classifications	
	105-01	105-05	105-10	95-01	95-05	95-10	85-01	85-05	85-10	75-01	75-05	75-10	N	%
105-01	13	1											13	92.9
105-05	1	13											13	92.9
105-10			14										14	100
95-01				13			**1**						13	92.9
95-05				1	12						**1**		12	85.7
95-10			**1**			13							13	92.9
85-01							12			**1**	*1*		12	85.7
85-05								14					14	100
85-10									14				14	100
75-01										14			14	100
75-05											14		14	100
75-10												14	14	100

Numbers highlighted in light gray - number of paper samples correctly classified; Numbers highlighted in dark gray - number of paper samples incorrectly classified by concentration (but correctely classified by grammage); Numbers highlighted in bold - number of paper samples incorrectly classified by grammage (but correctely classified by concentration); Numbers highlighted in italic and underline - number of paper samples incorrectly classified by grammage and concentration)

DISCUSSION

NIR spectra signatures

Figure 2 demonstrates the homogeneity of the spectral profile, with no differentiated signals among nanofibril contents. Characteristic bands of lignocellulosic materials can be observed in the spectra: OH of alcohol (6711 cm^{-1} or 1490 nm), OH of water (5181 cm^{-1} or 1930 nm) and OH of phenol (4761 cm^{-1} or 2100 nm), as observed by Fardim et al. (2005), C-H stretching and CH_2 deformation combination (4283-4386 cm^{-1} or 2280-2330 nm), as described by Workman Jr. and Weyer (2008).

The spectra of each treatment do not differ visually. Although the differences between the NIR spectra were very small, the multivariate statistical analysis was sensitive to variations between the NIR signatures. Small spectral variations are associated with differences in the NCF/fiber ratio of the papers, as only the amount of nanofibrillated material used in each treatment was different.

The papers evaluated in this study consisted of fiber, fines and nanocellulose fibrils. The NIR peaks highlighted in Figure 2B probably correspond to hydroxyl group (Workman Jr. and Weyer 2008). The information contained in NIR spectra, especially concerning -OH interactions, makes possible to predict the moisture content, cellulose content or tensile strength (hydrogen bonds).

Principal component Analyses

The PCA scores (Figure 3) had not displayed a clear separation among papers reinforced with 1, 5 and 10% of CNF demonstrating that it was not possible to separate papers through the NIR spectroscopy technique from the specimens set and experimental procedure used in this study. This difficulty can be associated to the unequal amount of CNF in each treatment (Table 1), since the percentages are associated with the paper grammage and the real value added to the paper increases for each grammage, even in the same nanofibril content.

Formation of clusters was more evident when the PCA was performed according to the nominal grammage (Figure 4). There was the formation of clusters for the grammage 105 g. m^{-2} just for the 10% nanofibrils concentration, with no separation between 1 and 5% concentrations (Figure 4). The nominal grammage 95 g. m^{-2} presented separation for all three nanofibrils concentrations. No cluster formation was observed for nominal grammage 75 and 85 g. m^{-2}, demonstrating that not only the differences in the real nanofibrils concentration with the same percentage represents an obstacle for the NIR technique to accurately identify and separate samples.

PLS-R and PLS-DA models to predict paper properties

According to Gemperline (2006), the root mean standard error (RMSE) describes the degree of agreement between the model estimated concentration values for the samples and the accepted true values for the samples used to obtain the model parameters. In other words, the RMSE measures the efficiency of the calibration model to predict the property of interest in unknown samples. Therefore, only the models that yielded the lowest RMSEcv were selected and presented in Table 4.

According to the data presented in Table 4, the calibration for nanofibrils concentration had the highest

R^2cv, demonstrating that NIR spectroscopy technique can be used to detect the presence of CNF in the cellulose handsheets. The same behavior could be inferred for tensile index (TI), stretch (ST) and resistance to air passage (RAP), which presented satisfactory R^2cv (0.91, 0.74 and 0.73, respectively) and low RMSEcv (2.27 Nm. g^{-1}, 0.21% and 7.52 s 100 cm^{-3}, respectively).

The nanofibrils concentration can be satisfactorily predicted by NIR Model (R^2cv = 0.98) while the estimates of grammage presented lower relationship with reference values (Model 2, R^2cv = 0.59). Thus, it is possible to infer that the better R^2cv of TI, ST and RAP is related to their strong dependence of nanofibrils concentration. On the other hand, TR, BI, TEA, RCT and CMT are influenced by grammage and nanofibrils concentration, with inferior NIR prediction capacity.

Few studies have reported NIR-based calibrations to estimate some paper properties. For example, Fardim et al. (2005) have analyzed hand-sheets produced from pulps refined to different levels via NIR spectroscopy and reported models with very good ability to predict TR, TI, BI, MOE and ST, with RMSEcv ranging from 0.13 to 3.64 and the performance maintaining itself when a new set of samples were analyzed. The RMSEcv of tensile index of papers of the present study was 2.27 Nm. g^{-1} while the RMSEcv of Fardim et al. (2005) results was 3.64 Nm. g^{-1}. For Stretch, our RMSEcv was 0.21% while that of Fardim et al. (2005) was 0.13%.

Most researchers supported the in-line applicability of the NIR spectroscopy technique in their studies. Samistraro et al. (2009) have studied Kraft paper with grammages between 115 and 440 g. m^{-2} by means of NIR spectroscopy and reported models with R^2cv ranging from 0.87 to 0.94 and RMSEcv from 1.5 to 8.2%, for TI, BI, TR and RCT. Although the TI found in this study is as good as that found by Samistraro et al. (2009), the superiority in the NIR prediction of the other properties can be associated with the larger range of grammages investigated by the authors, that may have provided more variability to construct a robust calibration. Tyson et al. (2012) have studied mechanical properties of *Eucalyptus* unbeaten bleached hand-sheets, with a grammage of 60 g. m^{-2}, collected directly from the production line. The authors found R^2cv for tensile index of 0.51, resistance to air passage of almost 0.40 and stretch lower than 0.10, all calibrations considered poor. The authors justified by stating that the calibration models were constructed from data with little variability, reflection of the reality of a commercial pulping operation, where the pulping process is tightly regulated and the variability of the final product is minimized. Viana et al. (2016) have analyzed cellulosic nanostructured films and found an R^2cv for bursting index of 0.93 and a tensile index of 0.83, confirming the NIR spectroscopy technique ability to predict nanocomposites properties.

Despite the existence of the studies reported above in which some paper properties were evaluated by NIR spectroscopy, to our knowledge, no studies have investigated properties such as grammage, corrugating medium test (CMT) or Tensile energy absorption (TEA). In the present study, TR, RCT and CMT presented the poorer R^2cv while grammage, TEA and MOE presented regular values of R^2cv (Table 4). No literature research analyzing Corrugating medium test, grammage and Tensile energy absorption by NIR spectroscopy was found for comparison purposes.

The strong association between the nanofibril concentration (R^2cv=0.98), resistance to air passage (R^2cv=0.73) and tensile index (R^2cv=0.91) determined in laboratory and that predicted by the NIR model, indicate the possibility of using the NIR spectroscopy technique to estimate nanofibril concentration and physical and mechanical properties of unknown papers (Figure 5). It is interesting to note that no clear cluster was formed by scores of PCA calculated from untreated data. Nevertheless, PLS regression algorithm was able to detect small differences in NIR signatures and associate them to variations in nanofibril concentration and key paper properties.

To our acknowledgement, no study reported PLS-DA models to obtain evaluate handsheet quality. However, it is possible to find in the literature studies that used this technique for other lignocellulosic materials, such as charcoal, engineered panels and cellulose pulps. Costa et al. 2018b used PLS-DA to classify charcoals by final temperature of carbonization, with one mistake in just one of the four categories, resulting in correct classifications of 97.8 to 100%. Belini et al. (2011) used PLS-DA to classify the percentage of sugarcane bagasse in MDF panels, with a 94% of the samples classified correctly by cross-validation. These studies prove that the partial least squares discriminant analysis (PLS-DA) can be successfully used to classify variables based on NIR spectra for the classification of lignocellulosic materials. There are differences between the spectra of different lignocellulosic materials due to the chemical difference between them. Meder and Meglen (2012) observed that the spectra of pure cellulose, hemicelluloses (galactose) and lignin are different and that the amount of these compounds interfered in the analyses. These studies reinforce the ability of PLS-DA to analyze lignocellulosic materials.

In the present study, the highest percentages of correct classifications for calibration (95.2%) and validation (83.9%) were obtained using the wavenumbers from 12500 to 3600 cm^{-1}, with the mathematical treatment of the first derivative (1D) and with ten latent variables (Table 5 - model 3).

The NIR range from 9000 to 4000 cm^{-1} performed better (Table 5) because the spectral range extending from 9,000 to 12,500 cm^{-1} has high noise level and present poor information quality (Hein et al. 2009). This specific range corresponds to the third harmonic region and the absorption is low (Workman Jr. and Weyer 2008).

In Table 6, all specimens of classes 3, 8, 9, 10, 11 and 12 were currently classified, with classes 10, 11 and 12 belonging to the same weight. That is, the model achieves 100% accuracy to classify the amount of CNF in the paper when the weight is 75 g m^{-2} (Table 1 and Table 6). Some samples were classified incorrectly, but the percentage of nanofibrils was the same; what changed was the weight of the paper. For example, a class 4 sample was classified as class 7, both classes contained the same amount of nanofibrils (1%) but with different weights (95 and 85 g.

m^{-2}, respectively) as shown in Tables 1 and 6. There were also incorrect classifications of the nanofibril concentrations within the same weight, for example class 1 (weight of 105 g m^{-2} and 1% CNF) missed only one sample as belonging to class 2 (weight of 105 g m^{-2} and 5% CNF).

Many studies have been developed to make possible to use NIR spectroscopy for pulp and paper control (Fardim et al. 2005; Santos et al. 2009; Tyson et al. 2010; Sandak et al. 2015; Viana et al. 2016; Costa et al. 2019). The studies dealing with paper properties is less numerous; However, the results reported in the literature point to a secure applicability of the technique to in-line control of paper and paperboard production. Fardim et al. (2005) recommended their NIR based models for mechanical properties of paper for immediate applications in "at-line" conditions.

In short, the promising findings of this study indicates that the quality of nano-based paper products can benefit from intense researches and increasing development of less expensive and faster methods to produce cellulose nanostructures. This approach has the potential to be developed as an on-line sensor to measure CNF content in paper and paper strength.

Limitations of this study

In this experiment, the added amount of CNF and the amount of the retained CNF in the paper could be different. The size of CNF is much smaller than fiber and 100% of CNF retention is probably not possible. In addition, the first pass retention of CNF in the paper may be affected by the grammage of the paper. Thus, the actual remained amount of CNF in the paper could be different in all prepared samples and it is very difficult to measure the retention of CNF.

Furthermore, pulp itself normally contains fines which is much smaller than fiber. In addition, lots of fines or microfibrils are produced during refining. Another limitation of this study is that no control treatment (without adding nanofibrils) was performed for each grammage. The initial idea of the study was to adopt grammage and CNF addition models that would generate papers with highly varied quality, as the objective was to generate predictive models based on NIR.

However, although we do not know the exact amount of nanofiber retained not even if there are fines in the pulp, the results show that NIR spectroscopy was sensitive enough to distinguish the levels of nanofiber in the cellulose handsheets. The statistics associated with the models indicate that it is possible to generate adequate estimates for monitoring and controlling the quality of papers.

CONCLUSIONS

The results found in this study show that NIR spectroscopy has potential to predict physical and mechanical properties of paper sheets. PLS-R models were capable to predict the nanofibrils concentration on paper with R^2cv of 0.98, although the PCA was not able to separate samples based on the nanofibrils concentration.

PLS-DA models for paper reinforced with CNF correctly classified the nanofibrils concentration or the grammage with more than 83% of hits.

AUTHORSHIP CONTRIBUTION

Project Idea: LCL, MLB, PRGH
Database: LCL
Processing: LCL, AMMC
Analysis: LCL, LRC, PRGH
Writing: LCL, LRC, PRGH
Review: LRC, PRGH, MLB, AMMC

ACKNOWLEDGMENTS

The authors would like to thank Klabin S.A., the Wood Science and Technology Laboratory and the Nanotechnology Laboratory of the Universidade Federal de Lavras (UFLA), the Pulp and Paper Laboratory of the Universidade Federal de Viçosa (UFV), Ph.D. Mônica Bezerra de Oliveira and Ph.D. Fernanda Maria Guedes Ramalho for providing material, structure and knowledge to carry out this experimental work. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001; by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq Brazil: grants n. 405085/2016-8 and 406593/2021-3) and by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, Brazil). PRG Hein was supported by CNPq grants (process no. 309620/2020-1).

REFERENCES

ARRIEL, T.G.; RAMALHO, F.M.G.; LIMA, R.A.B.; SOUSA, K.I.R.; HEIN, P.R.G.; TRUGILHO, P.F. Developing near infrared spectroscopic models for predicting density of *Eucalyptus* wood based on indirect measurement. Cerne, v. 25, p. 294-300, 2019.

BELINI, U.L.; HEIN, P.R.G.; FILHO, M.T.; RODRIGUES, J.C.; CHAIX, G. Near infrared spectroscopy for estimating sugarcane bagasse content in medium density fiberboard. BioResources. v. 6, p. 1816-1829, 2011.

BIRKETT, M.; GAMBINO, M. Estimation of pulp kappa number with near-infrared spectroscopy. Tappi Journal, v. 72, p. 193-197, 1989.

BÖRJESSON, M.; WESTMAN, G. Crystalline Nanocellulose - Preparation, Modification, and Properties. In: POLETTO, M. Cellulose - Fundamental Aspects and Current Trends. London: InTech, 2015. Chapter 7. p. 159-191.

COSTA, E.V.S.; ROCHA, M.F.V.; HEIN, P.R.G.; AMARAL, E.A.; SANTOS, L.M.; BRANDÃO, L.E.V.S.; TRUGILHO, P.F. Influence of spectral acquisition technique and wood anisotropy on the statistics of predictive near infrared–based models for wood density. J Near Infrared Spectroscopy, v. 26, p. 106-116, 2018a.

COSTA, L.R.; TRUGILHO, P.F.; HEIN, P.R.G. Evaluation and classification of eucalypt charcoal quality by near infrared spectroscopy. Biomass and Bioenergy, v. 112, p. 85-92, 2018b.

COSTA, L.R.; TONOLI, G.H.D.; MILAGRES, F.R.; HEIN, P.R.G. Artificial neural network and partial least square regressions for rapid estimation of cellulose pulp dryness based on near infrared spectroscopic data. Carbohydrate Polymers, v. 224, p. 115-186, 2019.

DIAS, M. C.; MENDONÇA, M. C.; DAMÁSIO, R. A. P.; ZIDANES, U. L.; MORI, F. A.; FERREIRA, S. R.; TONOLI, G. H. D. Influence of hemicellulose content of *Eucalyptus* and Pinus fibers on the grinding process for obtaining cellulose micro/nanofibrils. Holzforschung, v. 73, p. 1035-46, 2019.

FAO - FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. FAOSTAT: Forestry Production and Trade. Available at: http://www.fao.org/faostat/en/#data/FO. Accessed in: November 10th 2021.

FARDIM, P.; FERREIRA, M.M.C.; DURÁN, N. Determination of mechanical and optical properties of *Eucalyptus* Kraft pulp by NIR spectrometry and multivariate calibration. Journal of Wood Chemistry and Technology, v. 25, p. 267-279, 2005.

GEMPERLINE, P. Practical guide to chemometrics. 2ª edition. Boca Raton: Taylor and Francis, 2006, 520p.

HEIN, P.R.G.; CAMPOS, A.C.M.; TRUGILHO, P.F.; LIMA J.T.; Chaix, G. Near infrared spectroscopy for estimating wood basic density in *Eucalyptus urophylla* and *E. grandis*. Cerne, v.15, n. 2, 133-141, 2009.

HUANG, Y.; MENG, S.; ZHAO, P.; LI, C. Wood quality of Chinese zither panel based on convolutional neural network and near-infrared spectroscopy, Appl. Opt. v. 58, p. 5122-5127, 2019.

INDÚSTRIA BRASILEIRA DE ÁRVORES (IBÁ). Relatório Anual 2020. Available at: https://iba.org/datafiles/publicacoes/relatorios/relatorio-iba-2020.pdf. Accessed in: December 7th 2021.

ISO (1990) ISO 187: 1990, Paper, board and pulps—standard atmosphere for conditioning and testing and procedure for monitoring the atmosphere and conditioning of samples.

MEDER, R.; MEGLEN, R. R. Near Infrared Spectroscopic and Hyperspectral Imaging of Compression Wood in Pinus Radiata D. Don. Journal of Near Infrared Spectroscopy, v. 20, p. 583–589, 2012.

NAES, T.; ISAKSSON, T.; FEARN, T.; DAVIES, T. A user-friendly guide to multivariate calibration and classification. NIR Publications, Chichester, UK, 2002, 572p.

PASQUINI, C. Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. Journal of the Brazilian Chemical Society, v. 14, p. 198-219, 2003.

PASQUINI, C. Near infrared spectroscopy - A mature analytical technique with new perspectives - A review. Analytica Chimica Acta, v. 1026, p. 8-36, 2018.

REITER, C.; REINHARDT, B.; BORCHERS, B.; PLEW, P. Characterization of pulp and chemical additives with the help of FT-NIR technique. Das Papier, v. 53, p. 20–24, 1999.

ROSADO, L.R.; TAKARADA, L.M.; ARAÚJO, A.C.C.; SOUZA, K.R.D.; HEIN, P.R.G.; ROSADO, S.C.S.; GONÇALVES, F.M.A. Near infrared spectroscopy: rapid and accurate analytical tool for prediction of non-structural carbohydrates in wood. Cerne, v. 25, p. 84-92, 2019.

SAMISTRARO, G.; MUNIZ, G.I.B.; ZAMORA, P.P.; CORDEIRO, G.A. Previsão das propriedades físicas do papel Kraft por espectroscopia no infravermelho próximo (NIR) e regressão por mínimos quadrados parciais (PLS). Química Nova. v. 32, p. 1422-1425, 2009.

SAMYN, P.; BARHOUM, A.; ÖHLUND, T.; DUFRESNE, A. Review: nanoparticles and nanostructured materials in papermaking. Journal of Materials Science, v. 53, p. 146-184, 2018.

SANDAK A.; JASZCZUR, A.; SANDAK, J.; MODZELEWSKA, I. Near infrared assessment of biodegradability and mechanical properties of paper made of cellulose sulfate bleached coniferous pulp with addition of cationic starch and resinous adhesive. International Biodeterioration & Biodegradation, v. 97, p. 31-39, 2015.

SANDAK, J.; SANDAK, A.; MEDER, R. Assessing trees, wood and derived products with near infrared spectroscopy: hints and tips. J. Near Infrared Spectrosc., v.24, p. 485–505, 2016.

SANTOS, R.B.; GOMIDE, J.L.; SOUSA, L.C. Prediction of wood and kraft pulp quality by near infrared spectroscopy (NIRS). Rev. Árvore, v.33, p. 759-767, 2009.

TAPPI Standard (2002) T403. Om-02, Bursting strength of pulp.

TAPPI standard (2004) T414. Om-04, Internal tearing resistance of paper (Elmendorftype method).

TAPPI standard (2006a) T205. Sp-02, Forming handsheets for physical tests of pulp.

TAPPI standard (2006b) T494. Om-06, Tensile properties of paper and paperboard (using constant rate of elongation apparatus).

TAPPI standard (2006c) T 809. Om-99, Flat crush of corrugating medium (CMT test).

TAPPI standard (2007a) T536. Om-07, Resistance of paper to passage of air (high-pressure gurley method).

TAPPI standard (2007b) T 822. Om-02, Ring crush of paperboard (rigid support method).

TSUCHIKAWA, S.; KOBORI, H. A review of recent application of near infrared spectroscopy to wood science and technology. Journal Wood Science, v. 61, n. 3, p. 213-220, 2015.

TSUCHIKAWA, S.; SCHWANNINGER, M. A review of recent near-infrared research for wood and paper (Part 2). Applied Spectroscopy Review, v. 48, p. 560-587, 2013.

TYSON, J.A.; SCHIMLECK, L.R.; AGUIAR, A.M.; ABAD, J.I.M.; REZENDE, G.D.S.P.; FILHO, O.M. Development of near infrared calibrations for physical and mechanical properties of *Eucalyptus* pulps of mill-line origin. J. Near Infrared Spectroscopy, v. 20, p. 287-294, 2012.

TYSON, J.A.; SCHIMLECK, L.R.; AGUIAR, A.M.; MURO ABAD, J.I.; REZENDE, G.D.S.P.; FILHO, O.M.; MORA, C.R. Comparison of Sample Preparation Methods for NIR Analysis of Carbohydrate Content of Unbleached *Eucalyptus* Pulps. Journal of Wood Chemistry and Technology, v.30, p. 283-298, 2010.

VIA, B.K. Prediction of oriented strand board wood strand density by near infrared and Fourier transform infrared reflectance spectroscopy. J. Near Infrared Spectroscopy, v. 18, p. 491–498, 2010.

VIANA, L.C.; MUNIZ, G.I.B.; HEIN, P.R.G.; MAGALHÃES, W.L.E.; CARNEIRO, M.E. NIR spectroscopy can evaluate the crystallinity and the tensile and burst strengths of nanocellulosic films. Maderas: Ciencia y Tecnología, v. 18, p. 493-504, 2016.

WALLBACKS, L.; EDLUND, U.; NORDEN, B.; BERGLUND, I. Multivariate characterization of pulp using solid-state 13C NMR, FTIR, and NIR. Tappi Journal, v. 74, p. 201-206, 1991.

WORKMAN Jr, J.; WEYER, L. Practical guide to interpretive near-infrared spectroscopy. Boca Raton, CRC Press, Taylor & Francis Group, 2008.

WRIGHT, J.A.; BIRKETT, M.D.; GAMBINO, M.J.T. Prediction of pulp yield and cellulose content from wood samples using near-infrared reflectance spectroscopy. Tappi Journal, v. 73, p. 164-166, 1990.

YAN, D.B., KRISHNAGOPALAN, G.A. Dynamic modelling of carbohydrate degradation in Kraft pulping of softwood. Appita Journal, v. 56, p. 391-396, 2003.