Food Packaging and Shelf Life 31 (2022) 100791

Contents lists available at ScienceDirect

Food Packaging and Shelf Life

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

Rapid authentication and composition determination of cellulose films by UV-VIS-NIR spectroscopy

Patricia Cazón, Daniel Cazón, Manuel Vázquez *, Esther Guerra-Rodriguez

Department of Analytical Chemistry, Faculty of Veterinary, University of Santiago de Compostela, 27002 Lugo, Spain

ARTICLE INFO

Keywords:
Chitosan
Polyvinyl alcohol
Bacterial cellulose
Regenerated cellulose
UV-VIS-NIR spectroscopy
Chemometric

ABSTRACT

In recent years, efforts to develop new materials for the food industry have focused mainly on polysaccharides- and proteins-based films or coatings. Fast and inexpensive analytical tools are needed to guarantee their compositions. This work evaluates the feasibility of a rapid and accurate method based on UV-VIS-NIR spectroscopy combined with chemometric techniques to analyze polysaccharide-based films for authentication and composition determination. As case study, cellulose-based films (vegetable and bacterial) combined with chitosan and polyvinyl alcohol were used as biocomposite models. Applying chemometric techniques, it was obtained models to predict the content of chitosan, polyvinyl alcohol and cellulose. Linear discriminant analysis was used to authenticate cellulose films, showing an accuracy of 100% to classify cellulose films as function on the cellulose source (vegetable or bacterial). It was concluded that UV-VIS-NIR spectroscopy combined with chemometrics can be used to authenticate the origin and determine the composition of polysaccharide-based films.

1. Introduction

The global production of plastics almost reached 370 million tons in 2019, assuming a production in Europe of 58 million tons. Packaging is the main plastic consumer, being the 39.6% of the total demand (https://www.plasticseurope.org/en/resources/market-data). Only 9% of the total global plastic production is recycled, being 12% incinerated, and 79% accumulated in landfills or the environment. If current production and waste management trends continue, roughly 12,000 million tons of plastic waste will be in landfills or in the natural environment by 2050 (Geyer, Jambeck, & Law, 2017). The growing demand associated with the limitations of the management of its waste, has led serious environmental problems. Awareness about the impact of currently production and consumption systems on the environment has stimulated worldwide interest in the redesign of products, processes and services, seeking the sustainable use of raw materials, the reduction of plastic waste and sustainable systems (da Silva et al., 2021; de Medeiros et al., 2021; Foschi & Bonoli, 2019).

In this sense, the European Commission is strongly committed to issues related to the production of plastic materials and the management of plastic waste, with the implementation of new European policies and strategies aimed at achieving sustainability, as has been published in the *European strategy for plastics in the circular economy local and regional dimension* (European Commission, 2018; Foschi & Bonoli, 2019). European Commission aims to achieve 100% of reusable or easily recyclable plastic packaging placed on the market by 2030 (European Commission, 2018).

Due to the environmental crisis, the increase in social environmental awareness and the change in the environmental policies of the world powers, biodegradable materials from renewable sources emerge as a viable alternative for the reduction of plastic waste.

In recent years, polysaccharides and proteins have been the most extensively studied group of biopolymers to develop biodegradable, and also edible, films and coatings to apply on food and reduce the use of synthetic polymers (Khan, Sadiq, & Mehmood, 2020). These materials are obtained from natural and renewable sources such as plants, animals or microorganisms (Khemir et al., 2020; Nisar et al., 2019). In the group of polysaccharides, the main materials studied for food packaging are cellulose and its derivatives, chitosan, starch, alginate, pectin, pullulan and carrageenan. Within the group of proteins are gelatin, whey, soy proteins, wheat gluten, corn zein, casein and keratin (Cazón & Vázquez, 2021a; Cazon, Velazquez, Ramírez, & Vázquez, 2017; Mohamed, El-Sakhawy, & El-Sakhawy, 2020). Currently, the results of the polysaccharide- and protein-based films or coatings indicate that these materials are far from being able to completely replace food synthetic plastic packaging available on the market. However, these materials and

* Corresponding author.
E-mail address: manuel.vazquez@usc.es (M. Vázquez).

https://doi.org/10.1016/j.fpsl.2021.100791
Received 26 July 2021; Received in revised form 29 October 2021; Accepted 6 December 2021
Available online 18 December 2021

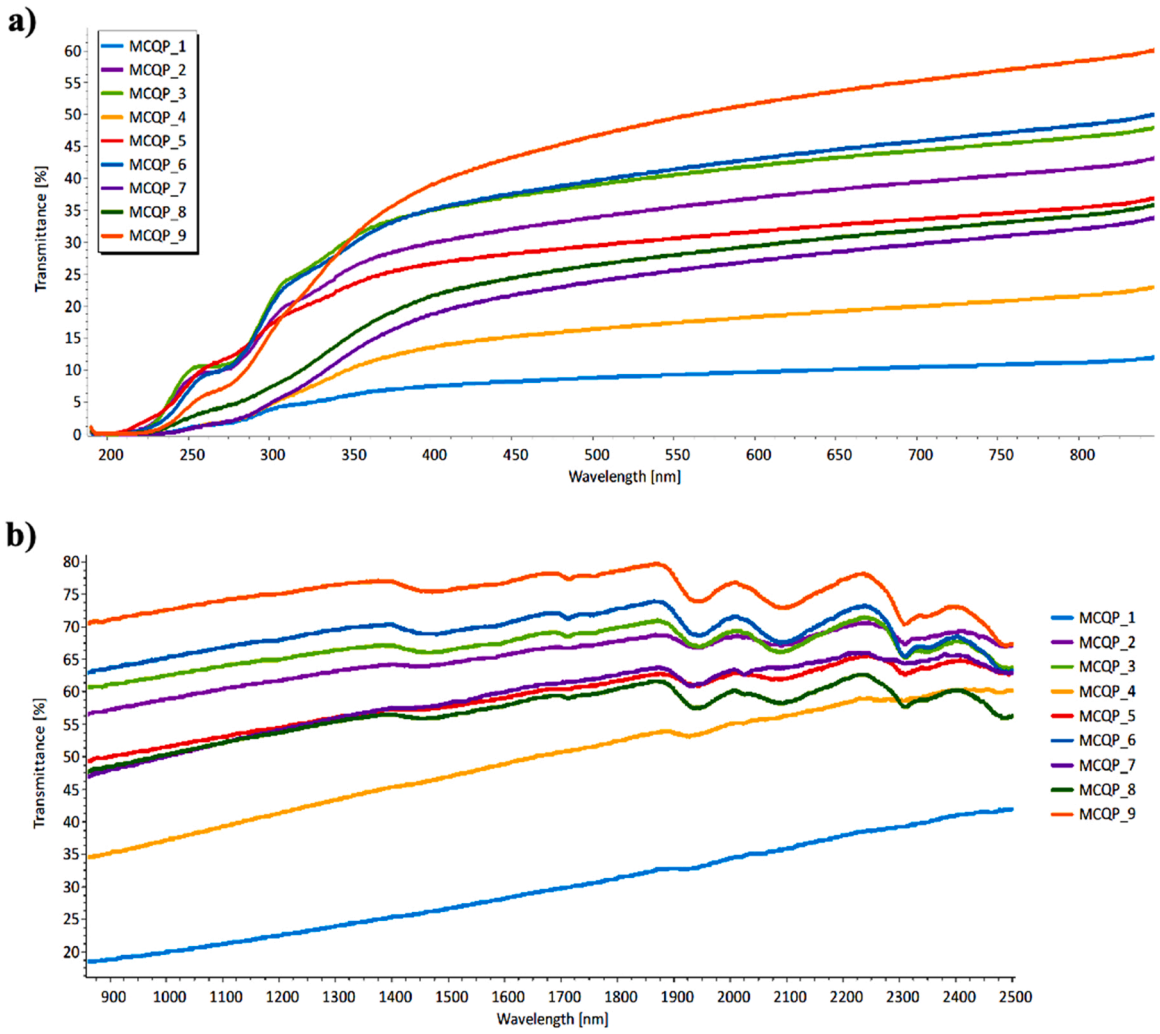

Fig. 1. UV-VIS (a) and NIR (b) region spectra of the bacterial cellulose samples with chitosan and polyvinyl alcohol. MCQP is bacterial cellulose with chitosan and polyvinyl alcohol.

their combinations as biocomposites, have been showed extraordinary properties which could achieve a partial replacement, as well as being an extra stress factor on packaging. In addition, several studies have demonstrated the feasibility of these materials to develop active and intelligent films or packaging that promise to be the next generation of food packaging (Janjarasskul & Suppakul, 2018; Krepker et al. 2017).

Many works in the literature on the characterization of protein- or polysaccharide-based materials and their biocomposites clearly demonstrate how their functional, structural and physicochemical properties have a strong dependence on the formulation of the final compound (Cazón & Vázquez, 2021b; Vázquez, Velazquez, & Cazón, 2021b; Vázquez, Velazquez, & Cazón, 2021a). For example, it was observed how a small variation of plasticizers, the relationship between the polymers used or variations in the concentration of other additives completely modify the material interactions. Even as they are hydrophilic materials, their interaction with the humidity of the surrounding atmosphere will vary, which directly affects the water activity and properties of the material (Cazon & Vázquez, 2019; Cazón & Vázquez, 2021b; Cazon et al., 2017).

On this basis, it is needed to develop a simple, rapid and low-cost analytical method that allows a precise real-time monitoring of the concentration of each component of the matrix at each stage of the manufacturing process. The concentration control system could guarantee the functionality of the film or coating and its behavior during its estimated shelf life under the storage conditions for which it has been developed.

The UV-VIS-NIR spectroscopy plays an important role to evaluate or predict the intermediate or finished product quality attributes due to their capacity to be a rapid, non-destructive and non-polluting method that requires minimal or no sample preparation (Rodionova, Fernández Pierna, Baeten, & Pomerantsev, 2021). In this method, the product is irradiated with UV-VIS-NIR radiation and the absorbed or transmitted radiation is measured. The spectral characteristics of the sample change through wavelength as function on its chemical composition, as well as on its light scattering properties which are related to the microstructure (Nicolaï et al., 2007). UV-VIS-NIR spectra process monitoring can be a very powerful tool in the biopolymer industry for the reduction of the cycle time and the cost of the analytical process (Fischer & Eichhorn, 1998). Detailed knowledge of the product composition makes possible to respond to process variations directly (Fischer, Bayer, Eichhorn, & Otto, 1997).

The great feasibility and simplicity of UV-VIS-NIR spectroscopy has meant that in recent years it has been evaluated for the quality control and product classification of several manufacturing process or heterogeneous products, such as the determination of polyphenolic compounds of red wines (Martelo-Vidal & Vázquez, 2014), cholesterol in egg yolk (Puertas & Vázquez, 2019a), volatile compounds in white grapes during ripening (Ripoll, Vazquez, & Vilanova, 2017), cellulose content in pulp (Zhou et al. 2018), mango quality attributes (Munawar, von Hörsten, Wegener, Pawelzik, & Mörlein, 2016), chemical composition of

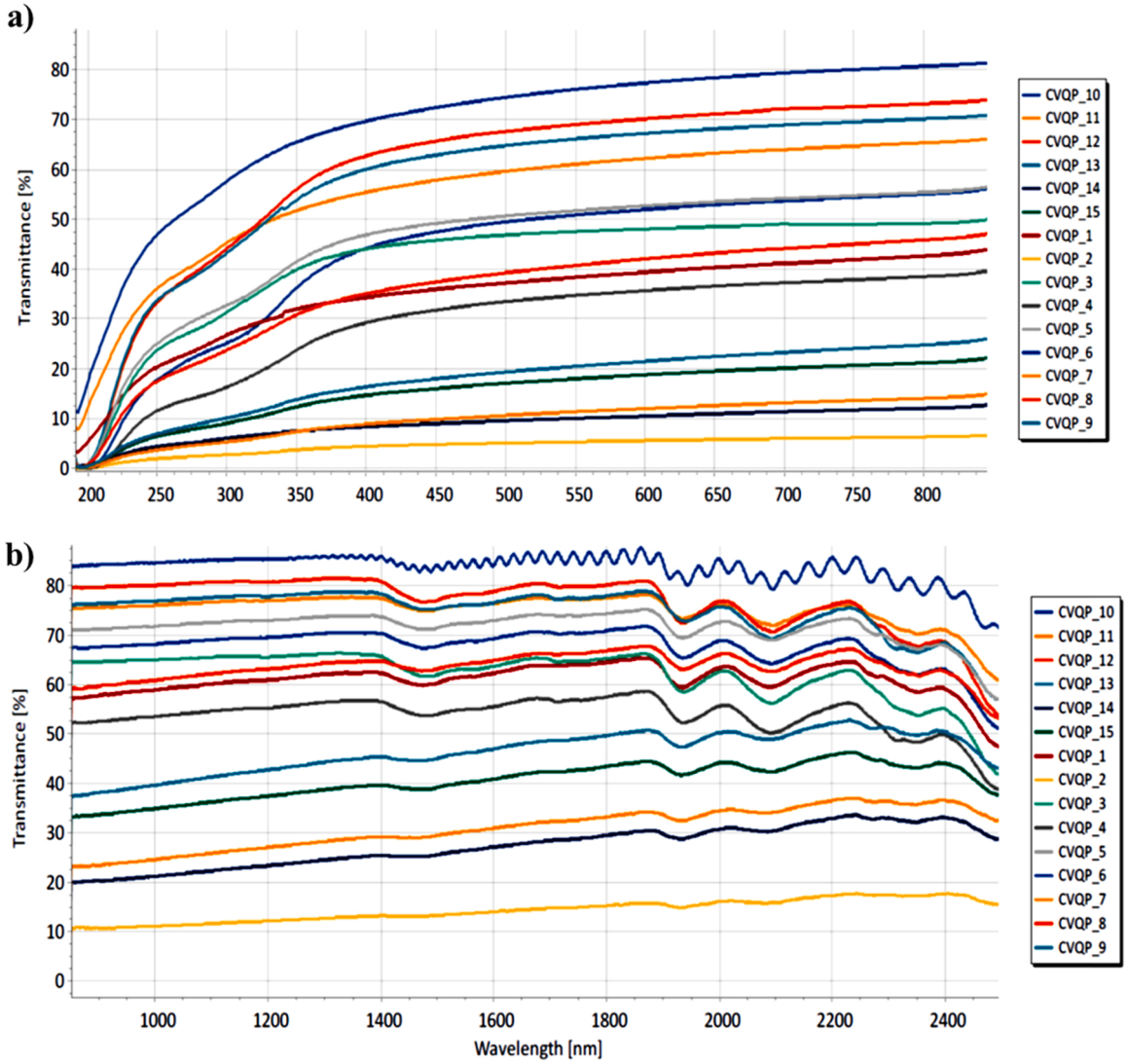

Fig. 2. UV-VIS (a) and NIR (b) region spectra of the regenerated cellulose samples with chitosan and polyvinyl alcohol. CVQP is regenerated cellulose with chitosan and polyvinyl alcohol.

plant materials (Huang & Yu, 2019), flax fiber composition (Huang & Yu, 2019), moisture content of wet granulation in a fluidized bed dryer (Barla, Kumar, Nalluri, Gandhi, & Venkatesh, 2014), discriminating geographic origin of sesame oils (Liu et al., 2019) or detection of soybean meal adulteration (Rodionova et al., 2021), among others.

Chemometric tools are required to extract the information provided by UV-VIS NIR spectra. Advanced multivariate statistical techniques are used to establish mathematical and statistical relationships between UV-VIS-NIR spectra, the chemical interpretation of data and target quality parameters (Munawar et al., 2016; Nicolaï et al., 2007).

In this work, the UV-VIS-NIR spectroscopy combine with chemometric tools has been evaluated to demonstrate its feasibility to develop qualitative and quantitative predictive models to control the composition of polysaccharide-based films. Regenerated cellulose (RC) films from vegetable origin and bacterial cellulose (BC) combined with chitosan and polyvinyl alcohol (PVOH) were used to test the UV-VIS-NIR spectra and chemometric tools to performance and validate accurate predictive components compositions and cellulose origin classification methods.

2. Materials and methods

BC films were elaborated using *Komagataeibacter xylinus* from the "Colección Española de Cultivos Tipo" (CECT), yeast extract provided by Scharlau Microbiology (Barcelona, Spain) and D(+)-glucose monohydrate (99% extra pure) and sodium hydroxide (98%) purchased from Acros organics (Geel, Belgium). Pure RC films were obtained from extra pure cellulose microcrystalline (average particle size 90 μm), urea (99.5%) for analysis, sodium hydroxide (98%) and acetic acid (99.5%) purchased from Acros organics (Geel, Belgium). Composite films were elaborated using chitosan (M_w 100,000–300,000) and acetic acid (99.5%) from Acros organics (Geel, Belgium) and polyvinyl alcohol, degree of hydrolysis (calculated on dried substance) ≥ 98% and molecular weight (M_W) approx. 30,000 obtained from Merck KGaA (Darmstadt, Germany). The samples were conditioned using extra pure anhydrous sodium bromide (99%) from Acros organics (Geel, Belgium).

2.1. Preparation of films

RC/chitosan/PVOH and BC/chitosan/PVOH were elaborated by immersion of the cellulose-based films in chitosan/PVOH baths at several concentrations published elsewhere (Cazon, Vazquez, & Velazquez, 2018b; Cazon, Velázquez, & Vázquez, 2019). Two Petri dishes were prepared for each batch. Dried films of 12 × 45 mm were cut and equilibrated in desiccators with saturated sodium bromide solution (57% relative humidity) at 25 ºC for 5 days and later measured in a

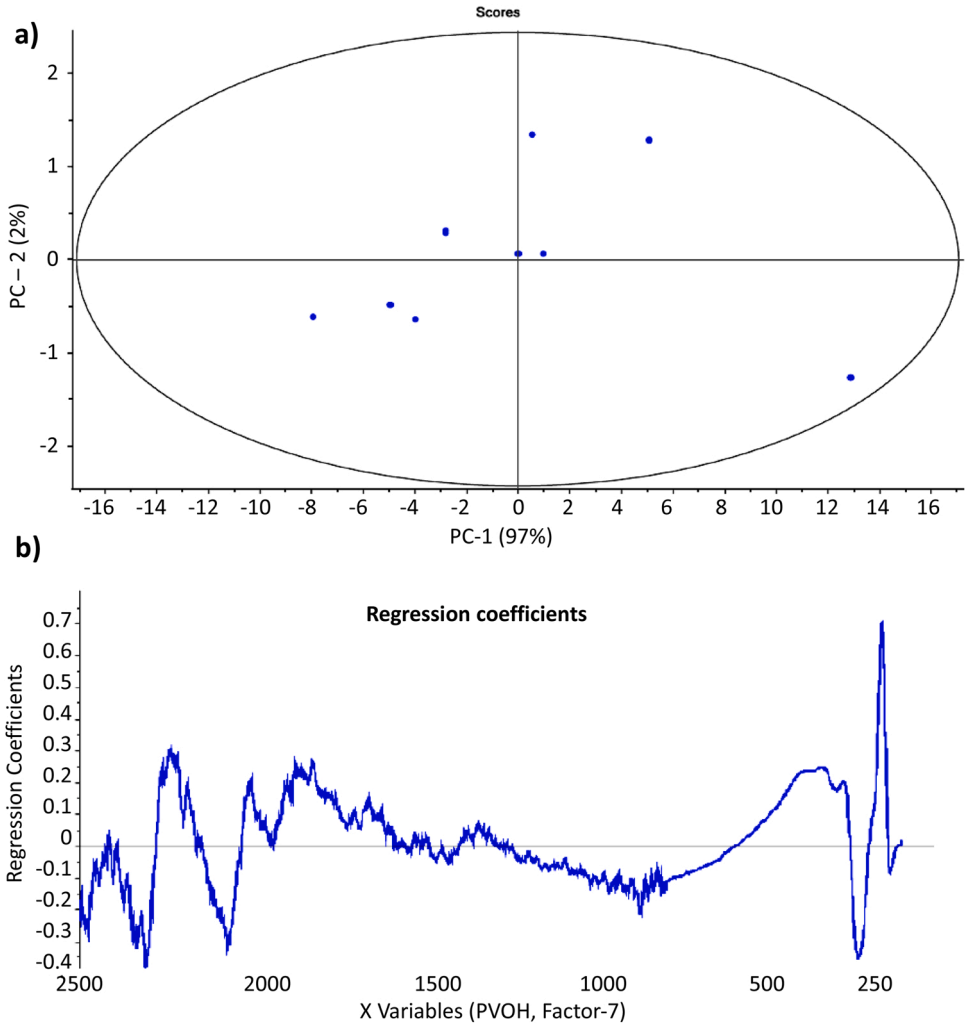

Fig. 3. a) Scores for bacterial cellulose samples combined with chitosan and polyvinyl alcohol and b) Regression coefficients of the PLS model of PVOH for each wavelength.

UV-VIS-NIR spectrophotometer.

2.2. Spectral analysis

Spectral measurements were carried out in a V-670 spectrophotometer (Jasco Inc., Hachioji, Tokyo, Japan) in transmittance mode at 2 nm intervals with a scan speed of 1000 nm/min and "medium" speed response. The samples were placed in the sample compartment using a special cell to measure the films transmittance. The sample must completely cover the circular section of this accessory where the light beam will pass through it. The spectrum region evaluated in this study included UV (190–380 nm), VIS (380–780 nm) and NIR (780–2500 nm) ranges. Measurements were performed in duplicate for each film formulation and at room temperature. The spectral data was recorded using Spectra Manager™ II software (Jasco Inc., Hachioji, Tokyo, Japan) and then exported to The Unscrambler® X software Version 10.5 (Camo, Oslo, Norway) to perform the pretreatments and obtain the different classification models.

The Unscrambler® X software supports multivariate data sets and incorporates the most common chemometric calibration algorithms. It was used for quantification and authentication. Smoothing (Savitzky-Golay) for preprocessing and variable selection was used since the optimized model with less variables require less time for calibration and prediction (Andersen & Bro, 2010).

Prior to the setting up of the regression models to obtain the predictive quantification and classifications models, the Principal Component Analysis (PCA) of the spectra was performed.

Calibration models were obtained to quantify the chitosan, PVOH and RC concentrations of the cellulose-based films using Principal Component Regression (PCR) or Partial Least Squares Regression (PLS). The validation of predictive models was performed by full cross-validation. Predictive capabilities of the calibration models and their validation were evaluated by five important statistical indicators, namely the root mean square error of calibration (RMSEC), root mean square error of cross-validation (RMSECV), the coefficient of determination of calibration (r^2_c), the coefficient of determination of cross-validation (r^2_{cv}) and the residual predictive deviation of cross-validation (RPD). The prediction error of a calibration model is defined by RMSEC and RMSECV when cross validation is used. The optimal number of latent variables was selected by a criterion of minimum RMSECV value. RPD is defined as the ratio of the standard deviation of the response variable and the standard error of prediction performance RMSECV (Nicolaï et al., 2007). The supervised Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM) approaches were performed to classify samples based on the cellulose source.

3. Results and discussion

BC/chitosan/PVOH and RC/chitosan/PVOH biocomposite samples

Table 1
Statistical values of the models obtained by PCR and PLS methods for bacterial cellulose (BC) or regenerated cellulose (RC) films combined with chitosan and polyvinyl alcohol (PVOH). PC is principal components, LF is latent factors, RMSECP is mean square prediction error, RMSECV is mean square validation error, RPD is the value of the ratio of performance to deviation, r^2_c is the coefficient of calibration determination, and r^2_{cv} is the coefficient of determination of the validation.

BC	**PCR - Calibration**			**PCR - Cross Validation**		
	PC	RMSEC (mg/g)	r^2_c	RMSECV (mg/g)	r^2_{cV}	RPD
PVOH	7	0.046	0.999	0.081	0.998	20.79
Chitosan	7	0.048	0.986	0.082	0.964	5.117
BC	**PLS - Calibration**			**PLS - Cross Validation**		
	LF	RMSEC (mg/g)	r^2_c	RMSECV (mg/g)	r^2_{cV}	RPD
PVOH	7	0.030	0.999	0.057	0.999	29.63
Chitosan	7	0.040	0.991	0.069	0.974	6.071
RC	**PCR - Calibration**			**PCR - Cross Validation**		
	PC	RMSEC (mg/g)	r^2_c	RMSECV (mg/g)	r^2_{cv}	RPD
PVOH	8	0.273	0.969	0.393	0.940	3.999
Chitosan	8	0.126	0.863	0.182	0.731	1.897
Cellulose	8	0.290	0.842	0.367	0.763	2.021
RC	**PLS - Calibration**			**PLS - Cross Validation**		
	LF	RMSEC (mg/g)	r^2_c	RMSECV (mg/g)	r^2_{cv}	RPD
PVOH	8	0.197	0.984	0.276	0.970	5.695
Chitosan	8	0.102	0.909	0.141	0.839	2.447
Cellulose	8	0.285	0.847	0.360	0.773	2.063

were characterized in previous work (Cazon et al., 2018b; Cazon, Vazquez, & Velazquez, 2018a), showing excellent properties to apply as active film in food packaging.

BC/chitosan/PVOH and RC/chitosan/PVOH biocomposite samples were obtained by dipping the BC or RC films in baths with chitosan (0–1% w/w) and/or PVOH (0–4% w/w). The BC concentration was considered constant since it depends on the composition of the culture medium and the fermentation conditions, which have been kept constant in each batch. RC-based films, the microcrystalline cellulose concentration ranged from 3% to 5% w/w. The BC/chitosan/PVOH samples formulation was established following a full factorial experiment design, analyzing all possible combinations at lower, intermediate and higher concentrations for each film component (chitosan and PVOH). In the case of RC/chitosan/PVOH samples, the formulation was established following a Box-Behnken experiment design, which is obtained by combining two-level factorial designs with incomplete block designs (Box & Behnken, 1960). The formulation of the films as well as their complete characterization can be consulted in previous works (Cazon et al., 2018b; Cazon, Vázquez, & Velazquez, 2019).

Transmission spectra in the UV-VIS region and in the NIR region of BC/chitosan/PVOH and RC/chitosan/PVOH are showed in Figs. 1 and 2, respectively. Both transmission spectra showed high differences between the samples as function of their composition. The NIR spectra exhibited intense absorption bands around 1490 nm, 1940 nm, 2100 nm from the combination of stretching and deformation of the O–H group (Barla et al., 2014). Samples with higher PVOH and chitosan content showed spectra with more obvious absorption bands at 1440 nm from the first O–H overtone (Barla et al., 2014). Samples with higher chitosan content showed a soft peak around 2050 nm, related with the combination of the N-H stretching and the second overtone of the C═O stretching (Rathke & Hudson, 1993; Tran, Duri, Delneri, & Franko, 2013).

The spectra of the both cellulose-based films indicated that the greatest variability of the samples was observed in the UV region around 250 nm and in the NIR region between 1400 and 2500 nm wavelength.

3.1. Bacterial cellulose combined with chitosan and PVOH: predictive model for the film composition

The PCA involved a complete analysis of the main components, reducing variables, redundant information and noise if required. PCA transform a set of correlated response variables into principal components (PC), generating a new set of non-correlated variables. The PC represents the pattern of observations and provides information about data structure. Hence, it was applied to raw spectra to detect outliers and patterns by subjecting a Hotelling T^2 ellipse with confidence level of 95%. Data outside of this ellipse, illustrated in the score graphic, was marked as spectral outliers and removed (Martelo-Vidal & Vázquez, 2014; Munawar et al., 2016; Puertas & Vázquez, 2019a). The scores plot (Fig. 3a) revealed a relationship between the data and outliers were not detected. Therefore, all samples were used to perform the further calibration models. Besides, the plot indicated that 99% of the variation of the spectra can be explained with two PC. PC-1 and PC-2 explain the 97% and 2% of the variation of the spectra, respectively. PCA results indicated that the relationship between the spectra and the composition of BC/chitosan/PVOH composite films can be interpreted with a high level of confidence.

Once the PCA analysis was obtained, the quantitative analysis was carried out by PCR and PLS. RPD is a commonly used to interpret and compare calibration models. Higher RPD values indicate a greater probability of the model to predict the desired chemical component in settled samples (Munawar et al., 2016). In specific agricultural applications, RPD larger than 1.5 is regarded as good for preliminary screenings and initial predictions; a value between 2.0 and 2.5 is considered satisfactory for prediction, value greater than 3.0 corresponds to an excellent prediction accuracy (Huang & Yu, 2019; Nicolaï et al., 2007). The statistical parameters r^2_c and r^2_{cv} represent the proportion of explained variance of the response variable in de calibration and validation set, respectively. The predictions were considered excellent when r^2 is greater than 0.91, good when r^2 is ranged between 0.82 and 0.90, approximate when r^2 is ranged between 0.66 and 0.81, and poor when r^2 is lower than 0.66 (Puertas & Vázquez, 2019a).

Table 1 shows the results obtained from the statistical analysis of the calibration and cross-validation model for the determination of the content of chitosan and PVOH in BC-based films by PCR and PLS methods. The complete range spectra of all samples were used to perform the calibration model. Due to the good correlation between data, it was not required any pretreatment of the spectra to improve the model accuracy.

The number of PCs and LFs (latent factors) were considered in the PLS and PCR method to obtain the predictive model. The number of PCs and LFs selected were 7, being the smallest number that provide the lowest RMSECV and the highest correlation coefficient. Therefore, 7-PCs and 7-LFs were needed to make up the optimal model.

The statistical results of the calibration and cross-validation models (Table 1) of the studied components showed a good relationship of the spectral data with the concentrations of chitosan and PVOH on BC-based films. Regarding PVOH quantification predictive models, both PCR and PLS methods showed r^2 values of 0.99 and RPD much higher than 2.5. For chitosan models, the lowest r^2 value obtained was 0.96, and RPD values were also much higher than 2.5.

Fig. 3b shows the regression coefficients for each wavelength for the PVOH concentration model obtained following the PLS analysis by means of 7 LFs. Results indicated that the UV region was the most influent region to build the predictive model.

Results can conclude that both methods, PCR and PLS, allow to obtain excellent calibration models, showing the PLS model a regression coefficient closer to 1 (Table 1). These results show that it is feasible to use the PLS and PCR models to quantify the concentration of chitosan and PVOH in BC-based films.

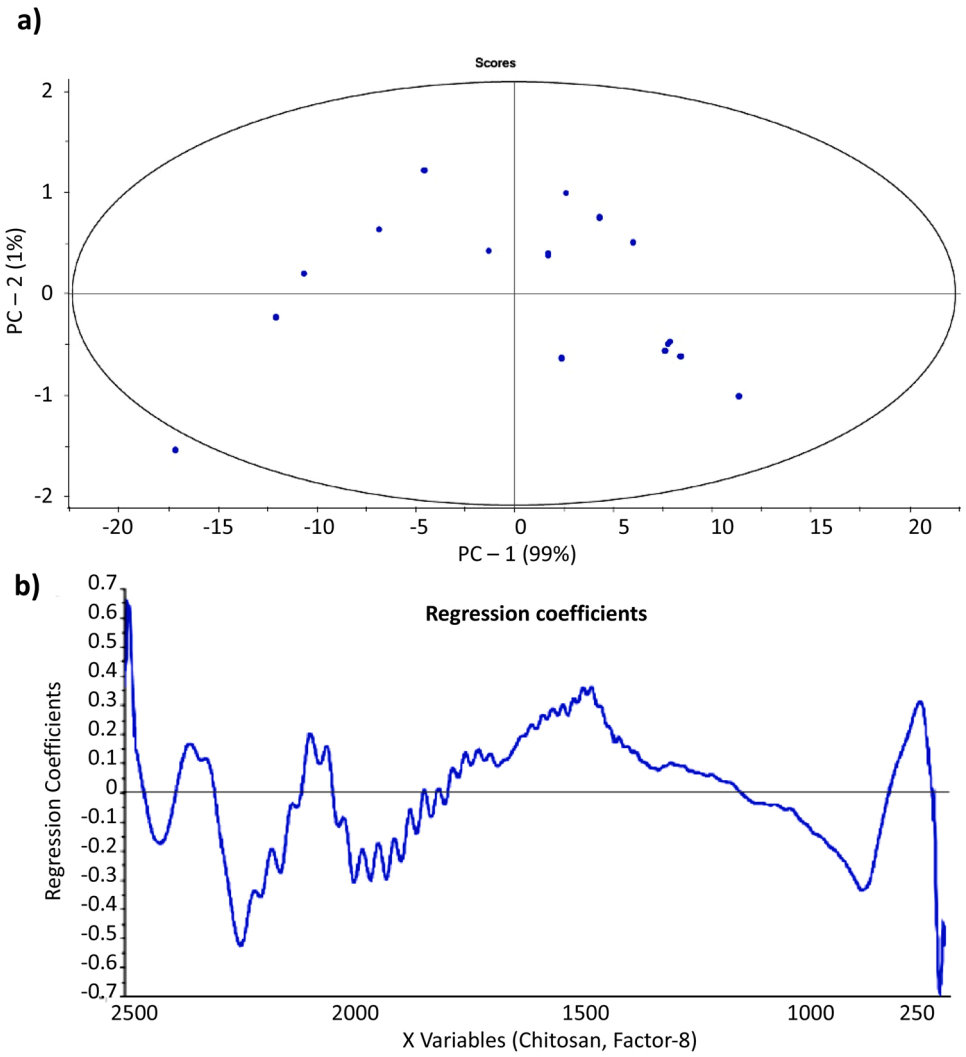

Fig. 4. a) Scores for regenerated cellulose samples combined with chitosan and polyvinyl alcohol and b) Regression coefficients of the PLS model of chitosan concentration for each wavelength.

3.2. Regenerated cellulose combined with chitosan and PVOH: predictive model for the film composition

The PCA of the spectra of the samples was performed, obtaining the scores plot showed in Fig. 4a. The preliminary analysis indicated that 100% of the variation of the spectra can be explained with two PC, despite one outlier was detected. PC-1 and PC-2 explain 99% and 1% of the variation of the spectra of the films, respectively. Therefore, the relationship between the spectra and the composition of the RC-based films can be interpreted with high level of confidence.

Once the PCA had been obtained and it was established that all the samples will be used to perform the calibration models, the quantitative analysis was carried out by means of PCR and PLS. An initial analysis showed statical coefficient values far from the desire values to get an optimal predictive model. Noise is the undesirable effect presents in the signals obtained in spectroscopy which affects the accuracy of the model. This can occur due to several reasons, such as variations in temperature and humidity (ambient noise) or the very nature of the sample. The pre-treatment of the spectra is carried out to avoid the noise and unreliable results. Spectral smoothing is applied in cases where the signal-to-noise ratio is small because the averaging of spectra is not sufficient. As a solution, there are some mathematical algorithms that can be applied to the spectrum and reduce the noise by smoothing the signal. Within smoothing, the most used method is based on Savitzky-Golay algorithm (Nicolaï et al., 2007). Thus, the pre-treatment Savitzky-Golay smoothing was performed to eliminate noise. Besides, as shown Fig. 2 the highest variability between the samples took place in the UV (200–300 nm) and NIR region between 1400 and 2500 nm wavelength. Therefore, it was decided to perform the PCR and PLS analysis using the spectral data reduced to these regions with higher variations.

Table 1 shows the statistical results of the calibration and cross-validation models, as well as the number of PCs and LFs used for the quantification of RC, chitosan and PVOH in the films by the PCR and PLS methods. The number of PCs and LFs selected was 8, being the smallest number of components or factors that provide the lowest RMSECV and the highest correlation coefficient.

The statistical results of the calibration and cross-validation models showed a good relationship between the spectral data and compound concentrations. Specifically, excellent methods have been obtained for PVOH quantification, showing r^2 and RPD values higher than 0.94 and 2.5, respectively, using both PCR and PLS methods. Better prediction model was obtained by the PLS method than PCR for chitosan quantification model, giving excellent r^2_c, good r^2_{cv} and RPD close to 2.5. However, in the case of predictive cellulose quantification models, both PLS and PCR, the obtained r^2 showed good calibration models (around 0.84) and approximate validation models ($\approx$ 0.76). The statical parameters indicated that both models could very roughly quantify

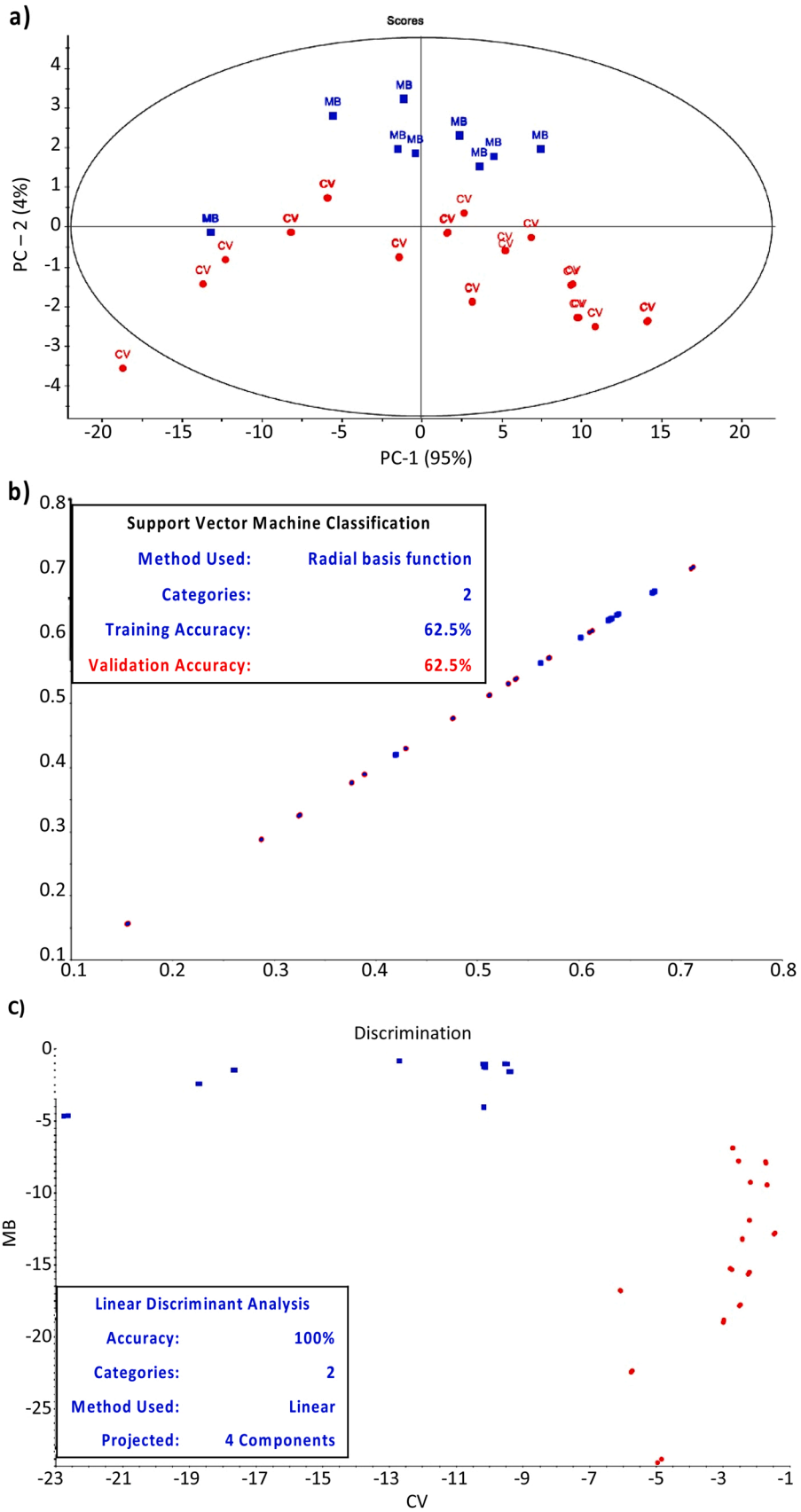

Fig. 5. a) Scores of the type of cellulose used in the production of the films; b) Classification of samples following the SVM analysis method and c) Result of the LDA classification method for regenerated and bacterial cellulose-based samples.

cellulose content. Probably, the reduced concentration range evaluated (from 3% to 5% w/w) was not enough to manifest significant transmittance variations among the samples to obtain a more accurate predictive model.

Fig. 4b shows the regression coefficients for each wavelength for the chitosan concentration model obtained following the PLS analysis by means of 8 LFs.

Overall, PLS demonstrated better model performance than PCR, as proven by the larger R^2_c, R^2_{cv} and RPD. Thus, PLS could be applied as a more powerful tool to quantify the chitosan, PVOH and cellulose concentration in RC-based films.

In the literature, no previous studies have been found using UV-VIS-NIR spectroscopy techniques to quantify the components of biodegradable films. Only previous studies have been found to quantify the cholesterol content in eggs (Puertas & Vázquez, 2019a) or components of wines (Martelo-Vidal & Vázquez, 2014) that have established as valid predictive models developed using quantitative analysis methods PLS and PCR with coefficients of determination around 0.73.

3.3. Identification of cellulose origin used to produce biodegradable films: cellulose from vegetable or bacterial sources

The samples were classified according to the cellulose origin: RC as vegetable cellulose or BC. The set of all spectral data were used. Data were analyzed using linear chemometric classification tools, LDA and SVM, to establish a method that allows the identification and classification of the type of cellulose in the films.

The supervised methods for classifying require the training set of samples that is used to locate the information responsible for splitting the samples into different groups. Within the supervised methods, discriminant methods divide the space into as many regions as there are classes in the calibration set, creating limits shared by the spaces. At least two classes are needed for these methods to be defined, which implies prior knowledge of the classes. An unknown sample is classified as belonging to one of the classes (Martelo-Vidal & Vázquez, 2014).

The PCA of the complete spectra data was performed to analyze the PC of BC and RC-based films. The scores graph (Fig. 5a) showed points clustering and the outlier corresponding to the Exp. 10 of RC/chitosan/PVOH. Despite this outlier, PCA results indicate that 99% of the spectra variation can be explained with two PCs. PC-1 and PC-2 explain 95% and 4% of the spectra data variation, respectively. Therefore, the relationship between the spectral data and the cellulose origin can be interpreted with a high degree of accuracy, allowing the classification by cellulose type.

SVM can work with linear and non-linear multivariate analysis and aims to seek the optimal hyperplane which minimize misclassification in the training process. It transforms the input data into higher dimensional space using kernel function known to influence its performance. Among the kernel functions, this study chose the structure of the radial basis function (RBF), which is the simplest and fastest computation and performances well to large numbers of training samples or large numbers of features in the input space (Puertas & Vázquez, 2019b; Wang et al., 2018).

The results obtained by the SVM method is shown in Fig. 5b. SVM method results exhibited a correct classification percentage of 62.5%. Consequently, SVM method was not able to discriminate between RC and BC. However, the LDA method achieved a classification with an accuracy percentage of 100% using 4 PCs (Fig. 5c). LDA algorithm aims to find a linear function that maximizes between groups variances and minimizes within groups variances to characterize or separate two or more classes of samples (Wang et al., 2018).

4. Conclusion

Chemometric analysis is a useful tool to obtain both quantitative and qualitative methods to determine the cellulose-based films composition and cellulose origin.

Results showed that it is possible to obtain fast and reliable prediction models for the quantification of the biodegradable films composition by chemometric analysis of the UV-VIS-NIR spectral data.

The spectra analysis of BC/chitosan/PVOH samples provided excellent predictive composition determination models by both PCR and PLS methods to quantify chitosan and PVOH.

PLS method showed better determination coefficient values for RC/chitosan/PVOH samples, being recommended to obtain a predictive composition model of chitosan and PVOH. However, both PCR and PLS methods were useful to obtain approximate predictive regenerated cellulose content models.

A fast classification method applying the LDA classification method has been obtained to classify the cellulose origin (vegetable or microbial) of films.

Thus, this study presents a fast method to quantify the polysaccharide films formulation with potential application in the food industry.

Ethical statement

Not Applicable. This research work does not carry out human or animal trials experiment.

CRediT authorship contribution statement

Patricia Cazón: Visualization, Methodology, Writing – review & editing. **Daniel Cazón:** Methodology, Investigation, Formal analysis, Writing – original draft. **Manuel Vázquez:** Formal analysis, Writing – review & editing. **Esther Guerra-Rodriguez:** Methodology, Writing – review & editing.

Conflict of interest

There is no conflict of interest regarding this manuscript.

Data availability

The whole data of the present manuscript is available; it will be provided if asked.

References

Andersen, C. M., & Bro, R. (2010). Variable selection in regression-a tutorial. *Journal of Chemometrics, 24*(11–12), 728–737. https://doi.org/10.1002/cem.1360

Barla, V. S., Kumar, R., Nalluri, V. R., Gandhi, R. R., & Venkatesh, K. (2014). A practical evaluation of qualitative and quantitative chemometric models for real-time monitoring of moisture content in a fluidised bed dryer using near infrared technology. *Journal of Near Infrared Spectroscopy, 22*(3), 221–228. https://doi.org/10.1255/jnirs.1095

Box, G. E. P., & Behnken, D. W. (1960). Some new three level designs for the study of quantitative variables. *Technometrics, 2*(4), 455. https://doi.org/10.2307/1266454

Cazon, P., & Vázquez, M. (2019). Applications of chitosan as food packaging materials. In G. Crini, & E. Lichtfouse (Eds.), *Sustainable Agriculture Reviews 36: Chitin and Chitosan: Applications in Food, Agriculture, Pharmacy, Medicine and Wastewater Treatment* (pp. 81–123). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-16581-9_3.

Cazón, P., & Vázquez, M. (2021a). Bacterial cellulose as a biodegradable food packaging material: a review. *Food Hydrocolloids, 113*, Article 106530. https://doi.org/10.1016/j.foodhyd.2020.106530

Cazón, P., & Vázquez, M. (2021b). Improving bacterial cellulose films by ex-situ and in-situ modifications: a review. *Food Hydrocolloids.* , Article 106514. https://doi.org/10.1016/j.foodhyd.2020.106514

Cazon, P., Vazquez, M., & Velazquez, G. (2018a). Composite films of regenerate cellulose with chitosan and polyvinyl alcohol: evaluation of water adsorption, mechanical and optical properties. *International Journal of Biological Macromolecules, 117*, 235–246. https://doi.org/10.1016/j.ijbiomac.2018.05.148

Cazon, P., Vazquez, M., & Velazquez, G. (2018b). Novel composite films based on cellulose reinforced with chitosan and polyvinyl alcohol: effect on mechanical properties and water vapour permeability (June) *Polymer Testing, 69*, 536–544. https://doi.org/10.1016/j.polymertesting.2018.06.016.

Cazon, P., Vázquez, M., & Velazquez, G. (2019). Composite films with UV-barrier properties based on bacterial cellulose combined with chitosan and poly(vinyl

alcohol): study of puncture and water interaction properties. *Biomacromolecules, 20* (5), 2084–2095. https://doi.org/10.1021/acs.biomac.9b00317

Cazon, P., Velazquez, G., Ramírez, J. A., & Vázquez, M. (2017). Polysaccharide-based films and coatings for food packaging: a review. *Food Hydrocolloids, 68*, 136–148. https://doi.org/10.1016/j.foodhyd.2016.09.009

Cazon, P., Velázquez, G., & Vázquez, M. (2019). Characterization of bacterial cellulose films combined with chitosan and polyvinyl alcohol: evaluation of mechanical and barrier properties. *Carbohydrate Polymers, 216*, 72–85. https://doi.org/10.1016/J.CARBPOL.2019.03.093

da Silva, C. J. G., de Medeiros, A. D. M., de Amorim, J. D. P., do Nascimento, H. A., Converti, A., Costa, A. F. S., & Sarubbo, L. A. (2021). Bacterial cellulose biotextiles for the future of sustainable fashion: a review. *Environmental Chemistry Letters, 19*(4), 2967–2980. https://doi.org/10.1007/s10311-021-01214-x

de Medeiros, A. D. M., da Silva Junior, C. J. G., de Amorim, J. D. P., do Nascimento, H. A., Converti, A., Costa, A. F., … Sarubbo, L. A. (2021). Biocellulose for treatment of wastewaters generated by energy consuming industries: a review. *Energies, 14*(16), 5066. https://doi.org/10.3390/en14165066

European Commission, 2018. A European strategy for plastics in the circular economy local and regional dimension. https://doi.org/10.2863/734820.

Fischer, D., Bayer, T., Eichhorn, K. J., & Otto, M. (1997). In-line process monitoring on polymer melts by NIR-spectroscopy. *Fresenius' Journal of Analytical Chemistry, 359* (1), 74–77. https://doi.org/10.1007/s002160050538

Fischer, D., & Eichhorn, K.-J. (1998). Process analysis of polymers by NIR spectroscopy. *Analusis, 26*(4), 58–60. https://doi.org/10.1051/analusis:199826040058

Foschi, E., Bonoli, A., 2019. The commitment of packaging industry in the framework of the european strategy for plastics in a circular economy. doi: 10.3390/admsci9010018.

Geyer, R., Jambeck, J., & Law, K. (2017). Production, use, and fate of all plastics ever made. *Science Advances, 3*(7), 25–29.

Huang, J., & Yu, C. (2019). Determination of cellulose, hemicellulose and lignin content using near-infrared spectroscopy in flax fiber. *Textile Research Journal, 89*(23–24), 4875–4883. https://doi.org/10.1177/0040517519843464

Janjarasskul, T., & Suppakul, P. (2018). Active and intelligent packaging: the indication of quality and safety. *Critical Reviews in Food Science and Nutrition, 58*(5), 808–831. https://doi.org/10.1080/10408398.2016.1225278

Khan, M. R., Sadiq, M. B., & Mehmood, Z. (2020). Development of edible gelatin composite films enriched with polyphenol loaded nanoemulsions as chicken meat packaging material. *CYTA – Journal of Food, 18*(1), 137–146. https://doi.org/10.1080/19476337.2020.1720826

Khemir, M., Besbes, N., Khemis, I., Ben, Di Bella, C., Lo Monaco, D., & Sadok, S. (2020). Determination of shelf-life of vacuum-packed sea bream (Sparus aurata) fillets using chitosan-microparticles-coating. *CYTA – Journal of Food, 18*(1), 51–60. https://doi.org/10.1080/19476337.2019.1696893

Krepker, M., Shemesh, R., Danin Poleg, Y., Kashi, Y., Vaxman, A., & Segal, E. (2017). Active food packaging films with synergistic antimicrobial activity. *Food Control, 76*, 117–126. https://doi.org/10.1016/J.FOODCONT.2017.01.014

Liu, Y., Xia, Z., Yao, L., Wu, Y., Li, Y., Zeng, S., & Li, H. (2019). Discriminating geographic origin of sesame oils and determining lignans by near-infrared spectroscopy combined with chemometric methods. *Journal of Food Composition and Analysis, 84*, Article 103327. https://doi.org/10.1016/j.jfca.2019.103327

Martelo-Vidal, M. J., & Vázquez, M. (2014). Determination of polyphenolic compounds of red wines by UV–VIS–NIR spectroscopy and chemometrics tools. *Food Chemistry, 158*, 28–34. https://doi.org/10.1016/j.foodchem.2014.02.080

Mohamed, S. A. A., El-Sakhawy, M., & El-Sakhawy, M. A. M. (2020). Polysaccharides, protein and lipid-based natural edible films in food packaging: a review. *Carbohydr. Polym.*. https://doi.org/10.1016/j.carbpol.2020.116178

Munawar, A. A., von Hörsten, D., Wegener, J. K., Pawelzik, E., & Mörlein, D. (2016). Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics. *Engineering in Agriculture, Environment and Food, 9*(3), 208–215. https://doi.org/10.1016/j.eaef.2015.12.004

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biol. Technol.*. https://doi.org/10.1016/j.postharvbio.2007.06.024

Nisar, T., Wang, Z.-C., Alim, A., Iqbal, M., Yang, X., Sun, L., & Guo, Y. (2019). Citrus pectin films enriched with thinned young apple polyphenols for potential use as bio-based active packaging. *CyTA – Journal of Food, 17*(1), 695–705. https://doi.org/10.1080/19476337.2019.1640798

Puertas, G., & Vázquez, M. (2019a). Cholesterol determination in egg yolk by UV–VIS–NIR spectroscopy. *Food Control, 100*, 262–268. https://doi.org/10.1016/j.foodcont.2019.01.031

Puertas, G., & Vázquez, M. (2019b). Fraud detection in hen housing system declared on the eggs' label: an accuracy method based on UV–VIS–NIR spectroscopy and chemometrics. *Food Chemistry, 288*, 8–14. https://doi.org/10.1016/j.foodchem.2019.02.106

Rathke, T. D., & Hudson, S. M. (1993). Determination of the degree of N-deacetylation in chitin and chitosan as well as their monomer sugar ratios by near infrared spectroscopy. *Journal of Polymer Science Part A: Polymer Chemistry, 31*(3), 749–753. https://doi.org/10.1002/pola.1993.080310320

Ripoll, G., Vazquez, M., & Vilanova, M. (2017). Ultraviolet - visible - near infrared spectroscopy for rapid determination of volatile compounds in white grapes during ripening. *Ciencia e Tecnica Vitivinicola, 32*(1), 53–61. https://doi.org/10.1051/ctv/20173201053

Rodionova, O. Y., Fernández Pierna, J. A., Baeten, V., & Pomerantsev, A. L. (2021). Chemometric non-targeted analysis for detection of soybean meal adulteration by near infrared spectroscopy. *Food Control, 119*, Article 107459. https://doi.org/10.1016/j.foodcont.2020.107459

Tran, C. D., Duri, S., Delneri, A., & Franko, M. (2013). Chitosan-cellulose composite materials: preparation, characterization and application for removal of microcystin. *Journal of Hazardous Materials, 252–253*, 355–366. https://doi.org/10.1016/j.jhazmat.2013.02.046

Vázquez, M., Velazquez, G., & Cazón, P. (2021a). UV-Shielding films of bacterial cellulose with glycerol and chitosan. Part 1: equilibrium moisture content and mechanical properties. *CyTA – Journal of Food, 19*(1), 105–114. https://doi.org/10.1080/19476337.2020.1870566

Vázquez, M., Velazquez, G., & Cazón, P. (2021b). UV-Shielding films of bacterial cellulose with glycerol and chitosan. Part 2: structure, water vapor permeability, spectral and thermal properties. *CyTA – Journal of Food, 19*(1), 115–126. https://doi.org/10.1080/19476337.2020.1870565

Wang, L., Hui, Y., Jiang, K., Yin, G., Wang, J., Yan, Y., & Wang, T. (2018). Potential of near infrared spectroscopy and pattern recognition for rapid discrimination and quantification of Gleditsia sinensis thorn powder with adulterants. *Journal of Pharmaceutical and Biomedical Analysis, 160*, 64–72. https://doi.org/10.1016/j.jpba.2018.07.036

Zhou, C., Han, G., Gao, S., Xing, M., Song, Y., & Jiang, W. (2018). Rapid determination of cellulose content in pulp using near infrared modeling technique. *BioResources* (Vol 13)(No 3). ⟨https://ojs.cnr.ncsu.edu/index.php/BioRes/article/view/BioRes_13_3_6122_Zhou_Rapid_Determination_Cellulose_Content_Pulp⟩.