Determination of the Recovered Fiber Content in Paperboard Samples by Applying Mid-Infrared Spectroscopy

Jordi-Roger Riba, Trini Canals, and Rosa Cantero

Escola d'Enginyeria d'Igualada, Universitat Politècnica de Catalunya, Av. Pla de la Massa 8, 08700 Igualada (Spain)

Abstract— Paperboard is widely applied in different applications, such as packaging or graphic printing among others. There is a growing consumption of recycled paper which has led paper mills packaging industry to apply strict quality controls. This means that it is very important to dispose of methods to test the quality of the recycled products. This paper is focused to determine the recovered fiber content in paperboard samples by applying Fourier transform mid-infrared spectroscopy in combination with multivariate statistical methods. To this end, two very fast nondestructive approaches have been applied, i.e. the classification and quantification approaches. The first approach is based on classifying unknown paperboard samples into two groups, namely high and low recovered fibers content. Conversely, under the quantification approach, the content of recovered fiber of the incoming paperboard samples is determined. Experimental results presented in this paper show that the accuracy of the classification approach in classifying unknown incoming paperboard samples is very high, whereas when applying the quantification approach the root mean square error of prediction is about 4.1.

Keywords— Infrared spectroscopy, recovered fiber, pulp and paper, multivariate analysis, paperboard, classification, quantification.

I. INTRODUCTION

The use of recovered paper products has expanded considerably over the last decades[1], mainly because of environmental, economic and social benefits[2]. According to the European Recovered Paper Council[3], Europe reached a recycling rate of 71.7% in 2012. In addition, paperboard is the most recycled packaging in Europe, exceeding the recycling rate of steel, glass or aluminum. Because of the economic crisis, paper consumption in Europe has been reduced by 13% since 2007 whereas recovering has dropped by only 3.5%. Therefore, it is essential to ensure the quality of the recycled material in order to guarantee the sustainability of the recycling proces[2].

According to the U.S. Environmental Protection Agency[4], paper fiber types are usually defined as either recovered or virgin. *Virgin fibers* are defined as cellulosic elements obtained directly from trees (hardwood and softwood) and other plants. It is worth noting that virgin fibers are newly pulped, so they never have been previously used. Instead, *recovered fibers* are defined as post-consumer fibers derived from diverse origins, including paper, paperboard and other fibrous materials which have been collected mainly from manufacturing processes or municipal solid waste. Recovered fibers can also include pre-consumer material, i.e. waste material recuperated from a manufacturing process.

The use of recovered fiber has several environmental benefits, since it reduces the demand of virgin fiber from forest products, thus putting less pressure on the forests. This also allows saving energy, reducing greenhouse gas emissions and extending available fiber supply. Recycling also minimizes landfill disposal of a valuable resource, since it allows reducing the amount of waste and rejected materials.

Paperboard is used to meet different needs. According to the final application, paperboard requires specific properties, including brightness, smoothness or strength among others, which can be achieved by using suitable blends of both virgin and recovered fibers. However, the use of recovered fibers in paperboard formulations used as packaging materials to be in contact with foodstuffs are of special concern[1]. This is because some of the chemical components included in the recovered materials are harmful to human health and can migrate from the packaging into food.

It is required some processing to obtain usable fibers from recovered fibrous materials, the extent of which and the amount of energy required depend on the final product requirements. Therefore, the use of too higher amounts of recovered fiber can reduce environmental returns beyond a threshold percentage. It means that the final manufactured product determines the maximum amount of recovered fiber in their formulations.

Manual and automatic paper sorting systems are being commonly applied in many countries to recover usable fibers from a waste stream[5]. Sorting methods pursue to recover the highest purity raw material from the waste stream, since by this way chemicals addition and energy requirements are minimized while facilitating the manufacture of high quality products[6]. However, manual sorting often faces several drawbacks including unpredictable end product quality, relatively high costs (especially in developed countries) or the exposition to dust, microorganisms or other pathogenic agents which may cause infections to the work team[7]. Therefore automated paper sorting systems are acquiring importance in the paper industry today and are constantly subjected to technological improvements.

There is a growing interest to develop automatic sorting systems. For example, a sorting system based on NIR spectral imaging has been described for paper classification of different paper types, i.e. raw and colored cardboard, newspaper and printer paper[8]. In Rahman et al.[5] it is described a paper sorting technique based on image processing combined with statistical reasoning and machine learning systems to identify different paper grades. A review of sorting methods for the paper industry can be found in Rahman et al.[6] However, available sorting systems either don't provide information about the composition of the analyzed samples or haven't been applied to determine the recovered fiber content in paper samples. This paper makes a contribution in this area since, as far as we know, it is the first attempt to automatically determine the recovered fiber content in paperboard samples.

There exist different analysis methods to identify paper products containing recovered fiber in their formulations. For example in Holik[9] it is described a system to determine the amount of damaged fibers. Other methods are based on the analysis of chemicals and products that remain in recovered paper fiber which inform of the content different from virgin fiber[10]. However these methods are time-consuming since they require sample preparation.

In this paper the recovered fiber content in different paper samples is determined by analyzing the spectral data provided by a mid-infrared spectrometer. Mid-infrared spectroscopy has been applied to analyze pulp composition and paper structure[1,11] and it is known to be very fast and non-destructive[12]. In this paper the Fourier transform mid-infrared (FTIR) spectrum of a given sample is further processed by applying multivariate feature extraction algorithms combined with classification and statistical regression methods. Unlike other approaches, the FTIR spectrum provides information about the paperboard sample compositions instead of the external physical appearance.

To determine the content of recovered fiber in a given paperboard sample, this paper applies two approaches. In the first one or classifier-based approach, unknown paperboard samples are classified into two groups, namely low and high recovered fiber content according to their composition by applying two feature reduction methods, namely principal component analysis (PCA) and canonical variate analysis (CVA) as well as the *k*-nearest neighbor (*k*NN) classifier. In the second approach or quantification-based approach, the content of recovered fiber is determined by applying a multivariate regression method, in this case the partial least squares (PLS) algorithm.

The proposed system for determining the recovered fiber content of an unknown sample has several appealing features including very fast response, it can be applied in situ, it doesn’t require the use of chemicals and reagents thus minimizing costs because both a chemical laboratory and a specialized technician are avoided. It is worth noting that recovered paperboard samples present a particularly varied diversity. Due to the wide range of compositions, i.e. the heterogeneity of the samples dealt with, this is a highly complex problem.

It is worth noting that the proposed quantification system, which is fast and easy-to-use, may be highly valuable for paperboard manufacturers since they need to check the quality of their incoming stock. It is also useful for packaging industries and especially for food packagers since they need to implement very strict quality controls to ensure that the content of recovered fiber is below a certain threshold value to avoid health related problems due to chemicals migration to foodstuffs.

II. MATERIALS AND METHODS

A. The analyzed samples

The recycled samples dealt with are composed of mixtures containing different proportions of raw pine mechanical pulp (virgin material) and pulp obtained from recycled newspapers, magazines (new samples returned to the printers) and grey paper scraps in a proportion of approximately 50/25/25. Note that the pulp samples were taken directly from the pulper. Next, appropriate dilutions and mixtures were done in the laboratory to obtain the different sample compositions, which afterwards were dried in a stove. By modifying the proportion of mechanical pulp depending on the desired quality, the material obtained can be used directly to manufacture the intermediate layer of the paperboard, which is analyzed in this work. The final product may include two other layers (they are not included in this work) which are composed of white recovered fibers (top side) and recovered paperboard (back side). When containing these two layers, the final product is designed as fully coated white lined chipboard with grey back and it is mainly applied for packing in the food industry, textiles, beverages, detergents and cleaning products among others.

It should be pointed out that the analyzed samples were prepared in two different time periods, therefore increasing the heterogeneity of the overall sample set since the incoming stock presents different origins and compositions. All samples were prepared in the facilities of Reno De Medici Ibérica.

A total amount of 31 paperboard samples were made by following the above mentioned manufacturing procedure. As explained, since the analyzed samples are made of recovered fiber with different proportions, this group of samples is highly heterogeneous. Therefore the automatic quantification of the recovered fiber content is a highly challenging problem.

The whole set of 31 samples was split into a training and a prediction set to evaluate the performance of the statistical models proposed in this paper[13]. The samples of the prediction set are different than those of the training set. Whereas the samples of the training set are required to calibrate the statistical classification and quantification models, the prediction set samples are used to predict the content of recovered fiber using different samples than those used in the calibration stage.

Table I shows the paperboard samples dealt with, their origin, and the set in which they are assigned.

TABLE I
SAMPLES ANALYZED IN THIS WORK

Sample	**Recovered fiber content (%)**	**Training set**	**Prediction set**
1	0.0	x	
2	5.0	x	
3	10.0		x
4	15.0		x
5	20.0	x	
6	25.0		x
7	30.0	x	
8	35.0	x	
9	40.0	x	
10	45.0		x
11	50.0	x	
12	55.0	x	
13	60.0	x	
14	65.0		x
15	70.0	x	
16	75.0		x
17	80.0	x	
18	90.0	x	
19	95.0	x	
20	100.0	x	
21	0.0	x	
22	10.0	x	
23	20.0	x	
24	30.0		x
25	40.0	x	
26	50.0		x
27	60.0	x	
28	70.0		x
29	80.0	x	
30	90.0		x
31	100.0	x	

Fig. 1 shows two of the 31 samples analyzed in this paper. It is worth noting that all analyzed samples are light grey in color, so they cannot be screened by simple visual inspection, except those produced from 100% virgin fiber, which are yellowish in color. Whereas the samples with some content of recovered fiber are light grey in color (recovered fibers have an important content of light grey chemical pulp), the samples produced from 100% virgin fiber are composed of mechanical pulp of yellowish color due to the presence of lignin, which is not removed during the manufacturing process of the mechanical pulp. The samples with some content of recovered fiber are composed of physical mixtures of mechanical pulp and recovered fiber, thus acquiring the light grey color since this color always predominates over the yellow color.

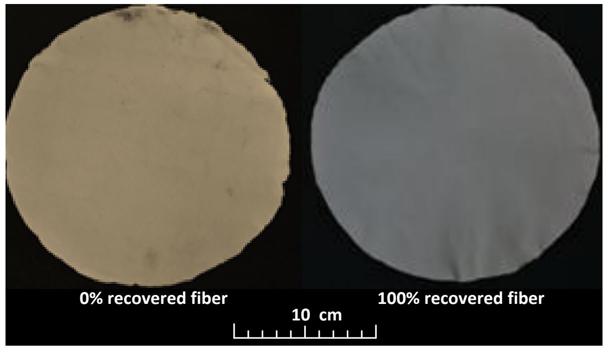

Fig. 1. Specimens 1 (left) and 20 (right) of the overall set of 31 samples studied.

B. *Spectral data acquisition*

To acquire the spectral data, a FTIR spectrometer model IR Spectrum One (S/N 57458) from PerkinElmer equipped with an attenuated total reflectance module (ATR) and a lithium tantalate ($LiTaO_3$) detector. The 45º ATR top-plate module has a clamping system to ensure an adequate contact between the solid sample and the single reflection diamond crystal.

The spectra of the raw paperboard samples were acquired at 25±1ºC by using an ATR cuvette over the wavenumber range 4000–650cm^{-1} by averaging four scans, with a resolution of 1 cm^{-1}. Three readings were done in different parts of each sample, which were averaged.

It is well known that by analyzing the ATR spectrum of a particular material, different types of components such as organic, inorganic and polymeric molecules among others may be identified. In the case of analyzing the ATR spectrum of a paperboard sample, most of the spectral bands are due to the cellulose[14].

In this paper the ATR spectra of 31 paperboard samples is acquired (one per sample), transformed to absorbance spectra and further analyzed by applying multivariate mathematical methods. The spectrum of each sample consists of 3351 data points (*x*,*y*), *x* being the wave-number and *y* the absorbance. This large amount of variables per sample combined with the inherent difficulty of the studied problem makes it is very difficult to determine the recovered fiber content of a given paperboard sample directly from the raw spectra data. Therefore, it is highly advisable to process this huge amount of spectral information by means of suitable multivariate statistical methods which are described in the following sections.

Fig. 2 shows the absorbance spectra of three paperboard samples with different content of recovered fiber, where it is possible to distinguish the characteristic bands of the cellulose (O-H, C-H and C-O-C). Fig. 2 also shows that the most marked differences among the three paperboard samples are found in the 1600-1500 cm^{-1} spectral band. The intensity of this spectral band can be associated to the presence of lignin in the samples (aromatic skeleton and C=O vibration modes). Regarding the analyzed samples, the intensity of this band decreases significantly when the recovered fiber content increases, so the intensity of this band is maximum for sample 1 (0% recovered fiber) and minimum for sample 20 (100% recovered fiber). However, the differences among samples with different recovered fiber content don’t seem to be linear by simple visual inspection, thus they need to be evaluated by means of suitable multivariate mathematical algorithms.

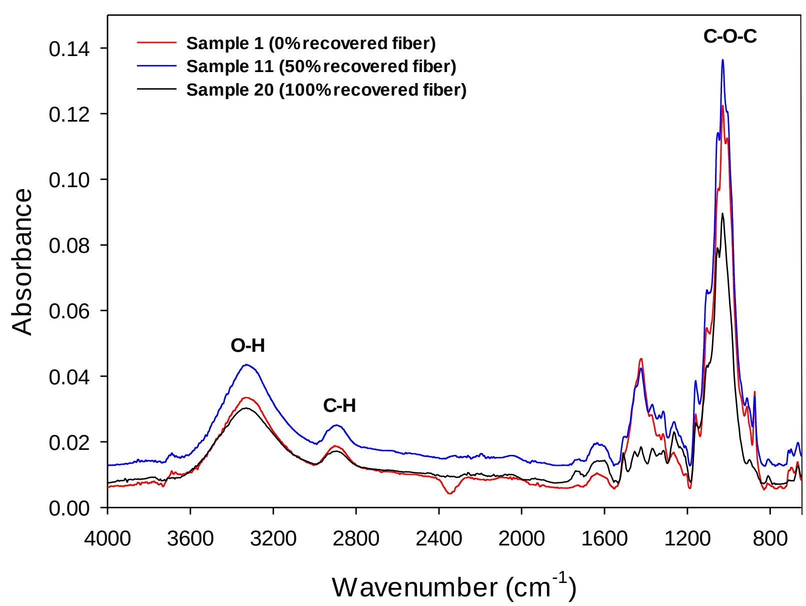

Fig. 2. Absorbance spectra of specimens 1, 11 and 20.

C. Feature Extraction Methods: PCA and CVA

It has already been explained that due to the large amount of measured or input variables included in the absorbance spectrum (3351) it is highly desirable to deal with statistical multivariate feature extraction methods since they allow concentrating the analytically significant information included in the measured variables in a smaller set of latent variables[12]. Feature extraction methods often remove most of the noise present in the measured variables. These new latent variables are usually calculated by applying mathematical combinations of the measured variables (absorbances at different wave-numbers).

Among the feature extraction algorithms, principal component analysis (PCA) is one of the most applied[15-17] although it is an unsupervised method. PCA combines linearly the measured variables, thus obtaining the latent variables or principal components (PCs), which are directed through orthogonal directions explaining the highest variance. PCA outputs the same number of PCs than original variables defined in the problem, although a reduced number of PCs are retained, those accounting for a sufficient portion of the total variance.

Supervised feature extraction methods are used to boost discrimination between classes[18]. Therefore, supervised feature extraction algorithms are preferred in classification problems. Unlike unsupervised methods, supervised methods use class labels to evaluate the latent variables performance. The class labels of the training samples are selected by a human expert.

Among the supervised feature extraction algorithms canonical variate analysis (CVA) highlights, since it is a multi-class method specifically designed to strengthen the differences between classes[19]. CVA calculates non-orthogonal latent variables called canonical variates (CVs). These latent variables are calculated by maximizing the differences between classes while minimizing samples dispersion within each class. CVA doesn’t work with data sets with a number of measured variables greater than the number of samples. This is the case of the problem analyzed in this paper, since there are 3351 variables and 31 samples. Therefore, to avoid this limitation, PCA is applied before CVA to reduce the number of variables dealt with, as shown in Fig. 3.

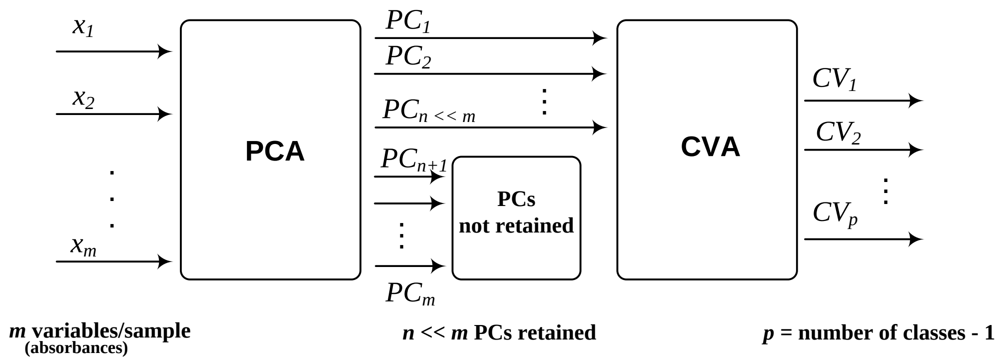

Fig. 3. Link between the PCA and CVA algorithms.

D. The kNN classifier

After reducing the dimensionality of the problem by applying a feature reduction method (PCA + CVA in this paper), classification problems require the application of a suitable classifier. The *k* nearest neighbors (*k*NN) algorithm is one of the most used classifiers since it has several advantages, including simplicity and accurate results[1]. The *k*NN algorithm categorizes the studied sample in the class most voted by the *k* nearest neighbors of the training set, taking into account their weighted vote. For this purpose and by applying the Euclidean distance, the *k* nearest neighbors of a test sample are located. Once the *k* nearest neighbors have been identified, *k*NN searches the nearest neighbor, and assigns a score *k* to its class, a score $k - 1$ to the second nearest neighbor's class and so on until a unity score is reached. Finally, the studied sample is classified into the most voted class. It is recommended to select *k* within 3 and 5 [20]. The outputs of the *k*NN algorithm are within 0 and 1, representing the degree of membership of the studied sample to each class defined in the problem. A test sample is assigned to the class whose membership degree is higher than 0.5. For each input sample, *k*NN usually provides the same number of outputs in the range [0,1] as classes defined in the problem.

E. The PLS algorithm for quantification

This paper deals with the PLS regression algorithm since it has become a standard in chemical quantification problems to process spectral data[21], although it has been used in several other scientific areas such as medicine[22], or social sciences among others. The multiple linear regression (MLR) algorithm, i.e.

the natural extension of the univariate linear regression based on the classical least-squares, is well suited to deal with multivariate data sets. However, when the number of input variables is too large, the model tends to be over-fitted, i.e. fits the sampled data perfectly but doesn't predicts well new samples[21]. Often there are a few latent factors that account for most of the variance in the response variable. PLS is devoted to calculate this reduced number of underlying factors, thus improving the model of the responses. PLS-regression is particularly suitable when there are more measured variables or predictors than samples, as well as when the predictors are collinear. PLS finds a linear model by projecting the matrix X containing the measured variables and the vector containing the predicted or dependent variables y in new spaces. The PLS model tries to find the multidimensional direction in the X space to explain the maximum multidimensional variance in the y space.

Fig. 4 shows the mathematical methods applied in this work to determine the recovered fiber content of the analyzed samples.

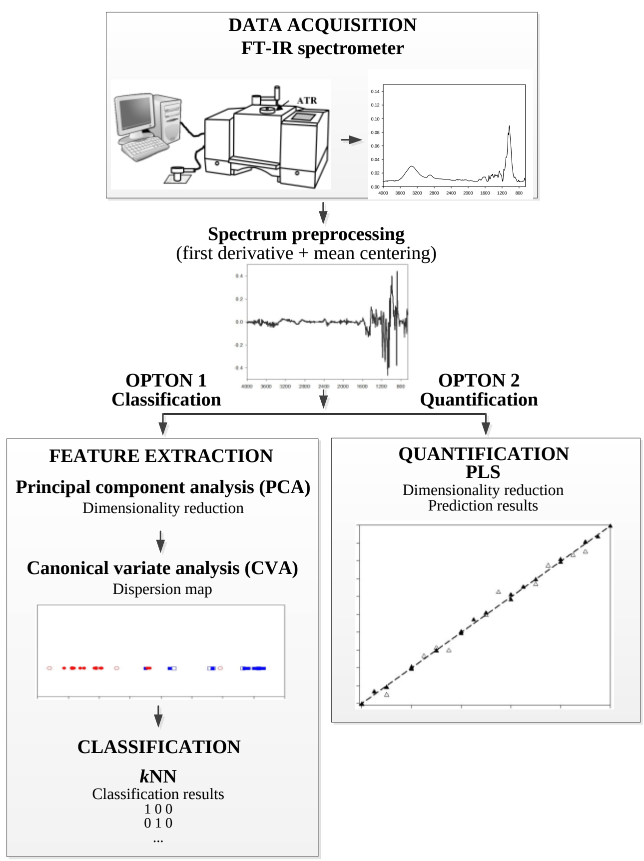

Fig. 4. The two approaches applied to determine the recovered fiber content of the analyzed samples.

All the multivariate statistical methods explained in this section have been programmed by the authors of this work using the Matlab® programming language.

III. RESULTS AND DISCUSSION

In this section the results attained with both analyzed approaches, i.e. the classification and quantification approaches are presented. All results are based on the analysis of the ATR spectra after suitable preprocessing, which includes baseline correction, smoothing, transformation to absorbance spectra, and analysis of the first and second derivatives with or without mean centering or unit variance scaling.

All results shown in this section are based on the 31 paperboard samples, which recovered fiber content is known, since they were expressly prepared for this research work in the Reno De Medici Ibérica facilities. These samples were split into two groups, i.e. the training and the prediction sets. Whereas the training set contains 21 samples, the prediction set includes the remaining 10 samples. Therefore the prediction set contains approximately one-third of the total set of samples.

The whole absorbance spectrum (4000-650 cm^{-1}) for the 31 paper samples provided a data matrix with 31 rows and 3351 columns, from which a first-derivative matrix of 31×3341 components was obtained as well as a 31×3331 second-derivative matrix by applying the Savitzky–Golay algorithm, which are shown in Fig. 5. It is worth noting that prior to calculating the derivatives, spectra were preprocessed by applying the baseline correction and smoothing operations. However, a prospective analysis showed more accurate results for both the classification and quantification approaches when dealing with the first derivative of the spectra with mean centering, so all results presented in this paper are based on this preprocessing method.

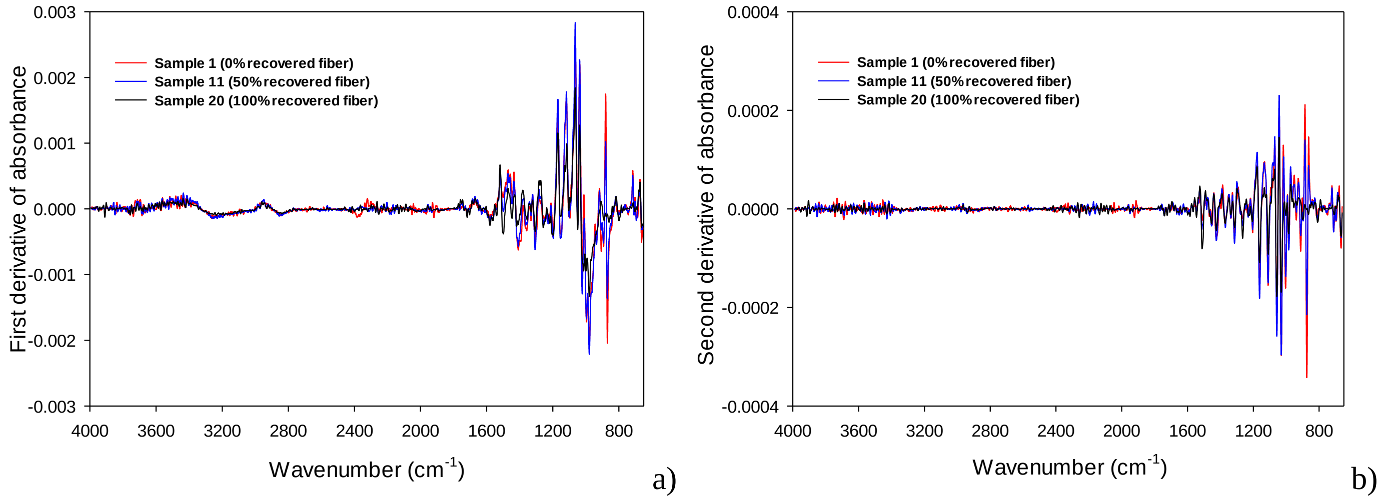

Fig. 5. First and second derivatives of the absorbance spectra of specimens 1, 11 and 20 obtained by applying the applying the Savitzky–Golay algorithm.

Paper industry market often demands paperboard products with either high or low recycled fiber content, which depends on the specific application of the final product. In these cases, a screening tool such as the one developed in the next subsection, based on PCA + CVA may be suitable. However, when a quantification of the recovered fiber content is required, the former method is not suitable. When applying the quantification approach based on the PLS algorithm, it is mandatory to prepare a calibration set of paperboard samples (the recovered fiber content of each sample must be known accurately) containing all the interval of recovered fiber content. Therefore, this strategy requires a more complex and accurate preparation of the pattern samples in the whole interval of concentrations dealt with.

A. Classification approach (PCA + CVA + kNN)

In some cases, manufacturers need a fast method to determine if the incoming paper samples contain or not contain a high percentage of recycled fiber. This is the case, for example, of the packaging industry, where the use of recovered fiber in paperboard formulations used as packaging materials to be in contact with foodstuffs is of special concern. Therefore, in such applications it is highly desirable to dispose of a fast and nondestructive screening tool for discriminating between incoming samples with high and low content of recovered fibers.

Under the classification approach, paperboard samples were split into two groups, namely low and high recovered fiber content, as shown in Table II. Therefore, this approach classifies unknown incoming paperboard samples within one of these two classes by applying the feature extraction methods PCA + CVA in combination with the *k*NN classifier.

TABLE II
GROUPS CONSIDERED UNDER THE CLASSIFICATION APPROACH

Classes	**Recovered fiber content (%)**	**Class labels**
Low	0 – 50	1
High	51 – 100	2

As detailed in Section II, the PCA is applied before the CVA algorithm. Therefore it is mandatory to select a reduced number of PCs arising from the PCA. Although there is not any standard method to select the appropriate number of PCs, in this paper those explaining at least the 97 % of the overall variance were retained. Fig. 6 shows that this condition is accomplished when retaining the first 10 PCs. Afterwards the CVA algorithm was applied to the 10 retained PCs.

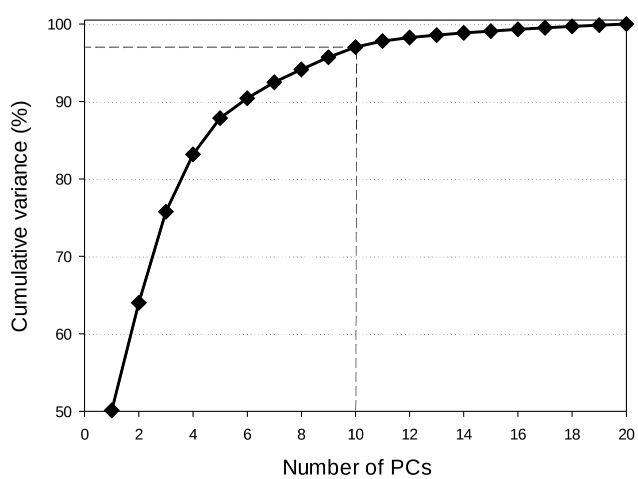

Fig. 6. Cumulative variance as a function of the number of retained PCs in the training data set when considering the overall 4000-650 cm^{-1} spectral interval.

Fig. 7 shows the results of the CVA algorithm for this two-class problem. It is worth noting that the number of CVs provided by the CVA is the number of classes minus one, i.e. in a two-class problem only one CV is calculated by the CVA algorithm. Fig. 7 shows both the training and prediction sets plotted in the one-dimensional space defined by the only CV arising from the CVA algorithm. Note that to achieve good classification results it is highly desirable that samples in classes 1 and 2 are as far apart as possible.

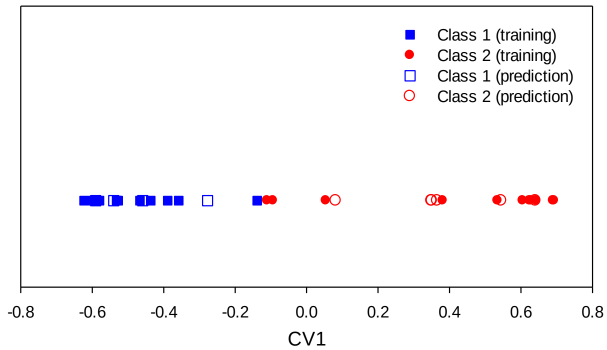

Fig. 7. Training and prediction samples plotted in the space defined by the only CV arising from the PCA (10 PCs, mean centering) + CVA algorithms.

Finally, the *k*NN (*k* = 3,4,5) classifier was applied to the data outputted by the PCA + CVA algorithms. Classification results attained by this method are summarized in Table III, which show that in all cases, i.e with *k* = 3, 4 or 5 neighbors, all samples were correctly classified according to their recovered fiber content.

TABLE III
PCA + CVA + *k*NN RESULTS SUMMARY

Feature extraction methods	***k*NN classifier**	**Prediction set success rate**
Spectral interval: 4000-650 cm^{-1}		
	k = 3	10/10 (100.0%)
PCA (10 PCs) + CVA	*k* = 4	10/10 (100.0%)
	k = 5	10/10 (100.0%)

B. Quantification approach (PLS)

Since the content of recovered fibers in the incoming stock has a profound impact on the quality and final properties of the manufactured paperboard products, in many applications it is highly desirable to dispose of a fast and nondestructive tool to determine the approximated content of recovered fibers.

The second approach to determine the recovered fiber content of the analyzed paperboard samples is based on the PLS regression algorithm. Similarly as in the case of the PCA algorithm, it is required to select the appropriate number of latent variables to avoid over fitting the prediction model. To this end, the mean squared error of cross-validation (MSECV) of the calibration sample set was calculated as a function of the number of PLS components retained, which is shown in Fig. 8. Note that the MSECV was calculated as,

$$\mathrm{MSECV} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{1}$$

$\hat{y}_i$ and y_i being the PLS prediction of the *i*-th sample and the reference value in p.u., respectively, and *n* the number of samples evaluated. The y_i values are known a priori since the samples were prepared expressly for this work.

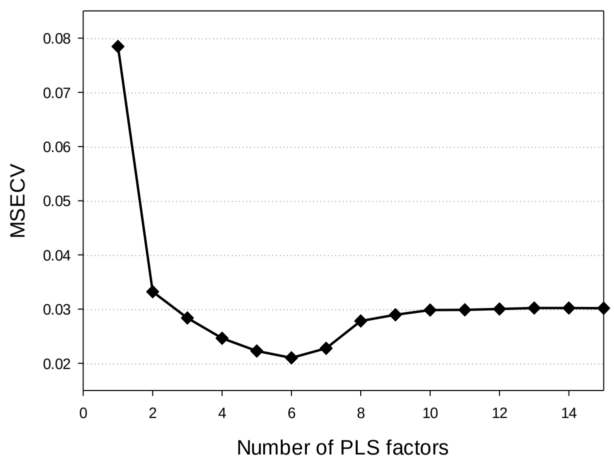

Fig. 8. Leave-one-out cross-validation MSECV of the training data as a function of the number of PLS factors retained PCs set when considering the overall 4000-650 cm^{-1} spectral interval.

After analyzing the values surrounding the minimum of the MSECV, the first 7 PLS components were retained.

To evaluate the accuracy of the PLS results, the root mean square error of calibration (RMSEC) or prediction (RMSEP) has been calculated as follows,

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2 / n} \tag{2}$$

Table IV shows a summary of the results attained by applying this method.

TABLE IV

PREDICTION DATA SET. PLS RESULTS WITH 7-FACTORS

Sample number	**Recovered fiber content (y_i, %)**	**PLS output ($\hat{y}_i$, %)**
3	10	13.14
4	15	14.80
6	25	24.31
10	45	37.87
14	65	72.43
17	75	71.00
25	30	29.68
27	50	49.53
29	70	65.12
30	90	86.23
RMSEP		**4.1**

Fig. 9 plots the recovered fiber content predicted by the PLS model in front of the results provided by the paperboard manufacturer for both the calibration and the prediction data sets. Results from Fig. 9 show a strong correlation (high R^2 values) between the manufacturer data and the results predicted by the PLS algorithm.

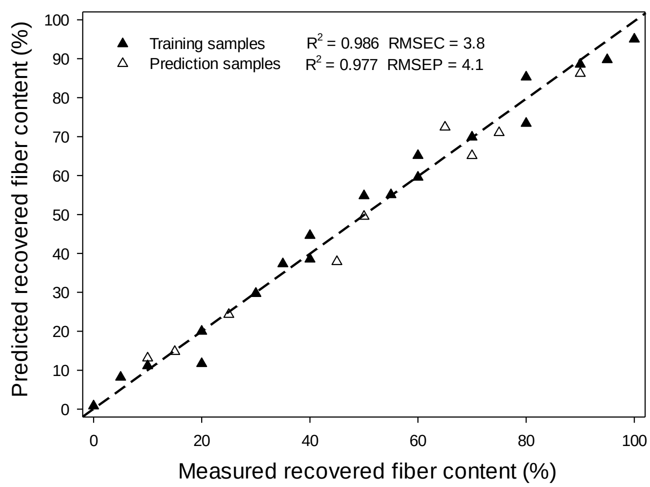

Fig. 9. Generalized correlation for calibration and prediction sets, respectively, when applying 7-factors PLS with mean-centered preprocessing.

IV. CONCLUSION

In this paper the recovered fiber content of paperboard samples has been determined from the ATR spectral information by applying two approaches. Under the first approach, unknown incoming paperboard samples have been classified into two classes, namely low and high recovered fiber content by applying the feature extraction methods PCA + CVA in combination with the *k*NN classifier. Under the second approach, the recovered fiber content of unknown paperboard samples is estimated by means of the PLS algorithm, achieving a root mean square error of prediction of 4.1. Therefore, these promising results show the potential of this method to determine the recovered fiber content in paperboard samples.

Appealing features of the proposed system include high speed, easy-to-use while avoiding the need of laboratory grade facilities. Therefore it may be very useful for paperboard and packaging industries since it allows a very fast and nondestructive testing of the incoming stock quality.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Miquel Figuera from Reno de Medici Ibérica for his helpful support during the preparation of this work.

REFERENCES

1. J.-R. Riba, T. Canals, R. Cantero. "Recovered Paperboard Samples Identification by Means of Mid-Infrared Sensors". IEEE Sensors J. 2013. 13(7): 2763-2770.
2. R Miranda, M. C. Monte, A. Blanco. "Analysis of the quality of the recovered paper from commingled collection systems". Resour. Conserv. Recy. 2013. 72: 60– 66.
3. European Recovered Paper Council (ERPC), Paper recycling. Monitoring report 2012. European Declaration on Paper Recycling 2011 – 2015, pp. 1-8, 2012.
4. U.S. Environmental Protection Agency. "Wastes - Resource Conservation - Comprehensive Procurement Guidelines". 2014. http://www.epa.gov/waste/conserve/tools/cpg/products/define.htm [accessed May 28 2014].
5. M.O. Rahman et al. "An Efficient Paper Grade Identification Method for Automatic Recyclable Waste Paper Sorting". Eur. J. Sci. Res. 2009. 25(1): 96-103.
6. M.O. Rahman, A. Hussain, E Scavino, H Basri, M.A. Hannan. "Intelligent computer vision system for segregating recyclable waste papers". Expert Syst. Appl. 2011. 38: 10398–10407.
7. H. Würtz, N O. Breum. "Exposure to Microorganisms During Manual Sorting Of Recyclable Paper Of Different Quality". Ann. Agric. Environ. Med. 1997. 4: 129–135.
8. P. Tatzer, M. Wolf, T. Panner. "Industrial application for inline material sorting using hyperspectral imaging in the NIR range". Real-Time Imaging. 2005. 11: 99–107.
9. H. Holik. Handbook of paper and board. Weinheim, Germany: Wiley-VCH, 2006. 1st ed.
10. B.L. Browning. Analysis of paper. Dekker, New York: United States, 1977. 2nd ed.
11. J. Pan, K. L. Nguyen. "Development of the photoacoustic rapid-scan FT-IR-based method for measurement of ink concentration on printed paper". Anal. Chem.2007. 79(6): 2259–2265.
12. S.-S. Kim et al. "Potential and Challenges for Mid-Infrared Sensors in Breath Diagnostics". IEEE Sensors J. 2010. 10(1): 145-158.
13. J.-R. Riba, T. Canals, R. Cantero. "Comparative Study of Multivariate Methods to Identify Paper Finishes Using Infrared Spectroscopy". IEEE Trans. Instrum. Meas., vol. 61, no. 4, pp. 1029-1036, March 2012. DOI: 10.1109/TIM.2011.2173048.

14. B. Stuart, Infrared Spectroscopy: Fundamentals and Applications. Wiley: Chichester, U.K, 2004.
15. J.-R. Riba. “Paper samples identification by means of IR spectrometry and multivariate methods”. Afinidad. 2012. LXIX(558): 87-94.
16. L. Lenhardt, I. Zeković, T. Dramićanin, M. D. Dramićanin, and R. Bro. “Determination of the Botanical Origin of Honey by Front-Face Synchronous Fluorescence Spectroscopy”. 2014. 68(5): 557-563.
17. S. Mert and M. Çulha. “Surface-Enhanced Raman Scattering-Based Detection of Cancerous Renal Cells”. Appl. Spectrosc. 2014. vol. 68(6): 617-624.
18. J.M. Leiva-Murillo, R. Santiago-Mozos, F. Pérez-Cruz, A. Artés-Rodríguez. “Comparison of Supervised Feature Extraction Methods for Multispectral Images”. Paper presented at Learning 2002, Madrid, Sapin.
19. L. Nørgaard, R. Bro, F. Westad, and S.B. Engelsen. “A modification of canonical variates analysis to handle highly collinear multivariate data,” J. Chemometr.. 2006. 20(8–10): 425–435.
20. L. A. Berrueta, R. M. Alonso-Salces, and K. Héberger. “Supervised pattern recognition in food analysis”. J. Chromatogr. A. 2007. 1158(1/2): 196–214.
21. R. Rosipal, N. Krämer. “Overview and recent advances in partial least squares”. Paper presented at the Iinternational conference on Subspace, Latent Structure and Feature Selection SLSFS 2005.
22. J. Xue et al. “Noninvasive Measurement of Glucose in Artificial Plasma with Near-Infrared and Raman Spectroscopy”. Appl. Spectrosc. 2014. vol. 68(4): 428-433.