Identification of Natural Rubber Samples for High-Voltage Insulation Applications

Jordi-Roger Riba[1], Neudys González[1], Trini Canals[2], Rosa Cantero[2]
[1]Universitat Politècnica de Catalunya, 08222 Terrassa (Barcelona), Spain
[2]Universitat de Lleida, 08700 Igualada (Barcelona), Spain

***Abstract*—**Latex presents high variability due to inherent differences among varieties from different countries, producers or crop seasonality. Natural rubber formulations from natural latex, to be used in insulating materials intended for high-voltage applications, require a wide variety of compounding and multitude of industrial processes. These aspects make it very difficult ensuring the same dielectric properties of the final product. At manufacturing level, it is very important to apply strict control processes to ensure that the final product fulfills all quality specifications. In this paper, a promising approach was applied to automatically identify natural rubber samples with suitable dielectric behavior from those with unsuitable dielectric behavior. This approach is based on the study of FTIR spectral data by applying suitable multivariable methods, such as principal component analysis, canonical variate analysis and *k*-nearest neighbors. The accurate and fast results reported in this work prove the suitability and potential of the proposed approach.

Keywords—Natural rubber, multivariable methods, infrared spectroscopy, identification, chemometrics, high-voltage.

1. INTRODUCTION

Natural rubber (NR) is an elastomeric material derived from latex, which currently is widely applied because of its appealing properties once vulcanized, including high elasticity, low-temperature flexibility, fatigue and tearing resistance, building tack or low heat buildup, among others. These properties confer NR distinctive advantages over different types of synthetic rubbers in numerous applications [1].

Vulcanized NR is broadly used to manufacture a wide range of products, including tires, adhesives, protective thin films and coatings, joints, medium- and high-voltage gloves, surgical gloves, or as suspension elements in civil structures [2–4]. However, NR used in such applications require extensive compounding to achieve its commercial grade [5,6]. NR recipes are usually based on different formulations [7] and processing [8], which heavily determine the properties of the final product. The formulations include elements such as elastomeric polymers, fillers, plasticizers, antidegradants, processing and vulcanizing agents, or vulcanization accelerators and activators, among others [2]. These additives include polymeric and inorganic materials, being solid at room temperature [9].

Titanium dioxide (TiO_2) and zinc oxide (ZnO) are the most cost effective and used as inorganic reinforcing fillers added to unvulcanized latex [9]. ZnO improves the heat resistance of the vulcanized material while

acting as a crosslinking agent, since it activates the rate of sulfur cure, jointly with the use of accelerators [10,11]. TiO_2 is added to increase protection against UV radiation.

Insulating materials intended for high-voltage applications require enhanced dielectric properties [12], including improved voltage at breakdown, reduced leakage current or high reliability [13]. In particular, when such products are used in life-line maintenance works under live conditions, safety issues are given top priority. Live-line maintenance involve different practices which become essential to reduce electrical failure occurrence, while ensuring power system availability and reliability [14]. Therefore, by applying suitable safety practices and protections, including NR safety elements, workers' exposure against current paths can be effectively minimized and prevented. Improved safety levels require minimizing the amount of leakage current through the insulation [15]. To this end, the rubber industry is developing rubber composite materials with enhanced dielectric and mechanical properties. NR compounds are widely applied to manufacture protective barriers, including insulating gloves and sleeves, insulating boots, flexible coverings, or blankets, among others [16].

NR, once suitably compounded and cured, results in a thermostable material with excellent dielectric and mechanical properties. However, both the dielectric and mechanical properties of NR compounds are greatly influenced by the specific compounding formulation. Significant work is being done to improve the dielectric behavior of NR based protective elements.

The dielectric behavior of composite materials depends on the volume fraction, size and shape of the compounding elements or fillers, or processing methods, among others [17]. Polymeric composites with structural inhomogeneity present decreased dielectric performance [18]. Solid fillers already present within the polymeric matrix tend to distort the local electric field, this distortion depending on the size of such particles. The breakdown electric field strength in NR composites depends on the mobility and density of field-dependent charge carriers, and the trapping probability of such charge carriers. Depending on electromechanical and thermal conditions, dielectric breakdown can occur at reduced electric field strength [19].

With the same volume fraction of fillers, the local electric field strength tends to enhance with the size of the particles. Agglomerates inside the polymeric matrix generates conductive pathways, which increase porosity, leakage current through electronic and moving charges conduction, electrical conductivity [12] and dielectric loss [18,20]. Consequently, dielectric properties are boosted when homogenizing the microstructure [9,13,18,20,21]. Therefore, the addition of well dispersed particles of reduced size tends to improve the dielectric performance of the polymer composite [22]. This improvement owes to the greater superficial area of the small aggregates, and the slower dynamics of the polymer chains wrapping the small particles [23].

Insulating NR formulations intended for high-voltage applications have to withstand different standard

chemical, mechanical and electrical tests, the latter ones providing relevant information about the electrical behavior of the samples. However electrical tests are usually carried out once the product is fully manufactured, and thus, if the samples do not pass such tests, nothing can be done to remediate this situation. Therefore, simple and nondestructive intermediate tests providing insight about the electrical behavior of the samples, which can be performed during the manufacturing process, are highly appealing. By this way, any problem occurring during the manufacturing process can be solved.

Few instrumental methods are available to measure electrical properties of polymer samples. Contact methods cannot be applied due to the deformation induced by the probes [7]. Methods requiring the value of some physical properties such as the dielectric constant, the refractive index or the sound can lead to substantial errors, because of the need for accurate parameter estimation. BDS (broadband dielectric spectroscopy) has been applied to study molecular and collective fluctuations, phase transitions, or polarization and charge effects in semi-crystalline [24,25], amorphous [8,26], and elastomeric composites [27,28]. Because of the non-contact nature, optical techniques often present better performance [7]. Although there are analytical techniques to determine the degree of curing and the possible interactions of the compounding elements, these chemical methods are usually time-consuming and destructive. Therefore, infrared Fourier transform spectroscopy (FTIR) using the total attenuated reflection (ATR) configuration, can be very useful to identify possible changes of specific functional chemical groups in the NR samples [29]. The non-destructive nature of ATR-FTIR, makes it suitable to be applied in process control for industrial applications. FTIR is among the most applied infrared techniques, because it offers high resolution spectra [30]. FTIR has been applied in the rubber industry for determining the content of vulcanizing accelerators and antioxidants [31] or to identify structural changes when applying a mechanico-chemical devulcanization process [32]. In [33] it is proved that changes in the chemical structure of NR are reflected in the FTIR spectrum. The position and intensity of the characteristic bands in the FTIR spectrum is directly related to the nature and proportion of the functional groups in the compound. Thus, the FTIR spectrum is affected by any structural change in the sample [34], so a measure of these changes can provide information of the final properties of the manufactured material. In other works [35] this technique has been used to identify samples based on the degree of vulcanization. Having into account the possibilities offered by this technique, in this study FTIR is selected to identify manufactured samples with a suitable dielectric properties from those that do not meet the required specifications.

The proposed approach combines an instrumental measurement method, the ATR-FTIR spectrometer, and specific data processing to discriminate NR samples with suitable and unsuitable dielectric behavior for high-voltage applications. It is proved that this measurement system allows a fast, non-destructive and accurate discrimination of such samples. Appealing advantages of the approach proposed in this work

comprise no consumption of chemical products, no need of sample pre-treatment, laboratory grade facilities or a qualified laboratory technician.

As reported in the technical literature, mathematical multivariable methods are often applied to FTIR spectral data, since this approach allows identifying minor differences between spectra [36]. The simultaneous application of PCA (principal components analysis) together with CVA (canonical variate analysis) has proved the accuracy and applicability of this approach in different areas, including testing of food authenticity, paper identification, or in the rubber industry among others [37–41]. This works is aimed to evaluate the applicability of PCA jointly with CVA to transform FTIR spectral data, in order to highlight structural changes in natural rubber composites used as insulating medium for high-voltage applications, which are related to the dielectric behaviour.

Since NR is derived from natural latex, which exhibits high variability between samples of different countries, producers and crops, its processing is very complex. In addition, the rich compounding required and the multitude of industrial processes applied under different ambient conditions, makes it difficult to ensure the same electrical and mechanical properties of the final product. Therefore, it is of paramount importance to apply strict control processes at industrial level, to ensure that the final product fulfills all specifications imposed by the international standards and costumers. This paper is focused in this direction, since it aims to identify vulcanized NR samples with suitable dielectric behavior (SDB) from those with unsuitable dielectric behavior (UDB), so that quality control methods can be simplified and expedited. In the case that samples of a particular production lot are identified as UDB, immediate corrective actions can be taken at production level, to minimize materials and economic losses. Therefore, due to the high variability of vulcanized NR samples, this results in a challenging problem, because of the heterogeneity and complexity of the analyzed data sets. It is worth noting that the fast and nondestructive approach proposed in this paper offers several advantages, since it does not require sample pretreatment or consumption of chemicals, thus being economically attractive because it avoids the need of laboratory grade facilities and a qualified laboratory technician.

2. EXPERIMENTAL

2.1 Materials

The natural latex from which the samples analyzed in this work were prepared, is a colloidal suspension, which was doubly centrifuged to rise the concentration of rubber particles. It was bought from diverse suppliers, mainly from Malaysia and Brazil, and was applied as received. Fillers and other ingredients added in the formulations were also of commercial grade. They were prepared by using commercial grade fillers and other ingredients and reagents. Titanium dioxide (TiO_2) and zinc oxide (ZnO) were acquired from Sigma Aldrich Spain. Reagent grade sulfur powder was used as vulcanizing agent, whereas ZDBC

(zinc dibuthyl dithiocarbamate) was applied as vulcanizing accelerator.

Typical formulations of the NR samples analyzed in this paper have expressed in % phr (mass parts per hundred of rubber) are as follows, ZnO (8.0% phr), TiO_2 (5.0% phr), antioxidant (8.0% phr), accelerator (6.0% phr), sulfur (6.0% phr), others fillers (7.0% phr) [9].

The manufacturing process of solid rubber films includes two main steps. Liquid NR was first masticated for 5 min and afterwards ZnO is added and mixed for 2 min. Then, the mixture is further mixed with stearic acid during 2 min. TiO_2 is then added into the mixing chamber after the pre-sonication and dispersion explained before. Next, the mixture is dumped out and ZDBC (zinc dibutyldithiocarbamate) is added and mixed during 2 min. Then, sulfur is incorporated and further mixed for 2 min. Next, the emulsion is stirred during 30 min before dipping tube tests into the liquid suspension. Then, the tubular shaped specimens are introduced in an oven to pre-vulcanize for 3h at 76ºC, thus obtaining solid specimens. Finally, specimens are washed in a distilled water bath, dried, and then cured for 24 h before passing the electrical test.

The samples analyzed in this work present a high variability, which depends upon several factors, including latex origin, environmental conditions, or the conditions of several manufacturing processes, among others. Other important factors during the manufacturing stage, are the washing and vulcanization processes. The latex origin is an important factor, since NR could present a higher content of minerals or non-rubber products, such as proteins, which also affect the electrical properties. During washing, all compounds which do not react effectively during the pre-vulcanization process must be eliminated, otherwise the final properties can be greatly affected. Environmental conditions greatly impact the final properties of NR samples. NR is also sensitive to high temperature, since agglomerates can be generated within the material, these heterogeneities greatly affecting the electrical properties. Sample thickness also plays an important role on the final dielectric properties, since thicker samples tend to present more agglomerates and inhomogeneities.

Samples were collected samples at different seasons of the year, they include different types of washing processes and a broad range of thickness, comprised between 0.75 and 2.9 mm. Since samples were collected during one full year, the incoming latex presented diverse origins. This is a challenging problem due to the high variability of the samples analyzed.

The electrical test, which is schematized in Fig 1, consists in introducing the tubular latex samples into osmotized water, and measuring the leakage current between the high-voltage electrode, which is placed inside the tubular specimen, and the ground electrode. The specimen passes the test if the leakage current is below a threshold value.

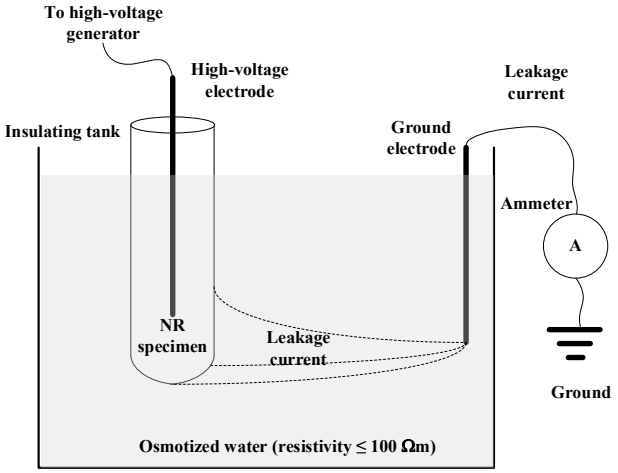

Fig. 1. Electrical test applied to the analyzed NR specimens to measure the leakage current. The specimens that pass the electrical are classified as SDB (suitable dielectric behavior), the remaining ones being classified as UDB (unsuitable dielectric behavior).

2.2 Infrared spectroscopy

Transmittance spectra of the NR samples were acquired by using a PerkinElmer Spectrum One (S/N 57458) FTIR spectrometer, which included an attenuated total transmittance (ATR) accessory (S/N P0DL01101418). Spectral data with a resolution of 1 cm^{-1} were acquired within the interval 4000–650 cm^{-1}, at 25±1ºC. Spectra were obtained by averaging 4 scans per sample.

3. MULTIVARIATE DATA ANALYSIS METHODS

As operated in this study, raw the ATR-FTIR instrument provides 3351 wavenumbers per spectrum (4000–650 cm^{-1} with a resolution of 1 cm^{-1}). This is a large amount of data, so it becomes imperative to apply efficient and very fast multivariate methods to solve this problem. To this end, the problem dimensionality is reduced, thus concentrating the information that is analytically relevant into a reduced set of inferred variables, while removing the random noise included in the raw spectral data. These new variables are the latent variables and are calculated by means of weighted combinations of the original ones [36,42]. The methods intended for this purpose are called feature extraction methods. Methods intended for feature extraction are divided into supervised and unsupervised methods. Supervised methods are more suitable to deal with classification problems, since they use class labels for the calibration set, which are based on the criterion of an expert user to guide the classification process [40].

To calibrate the algorithms it is essential to split the total number of samples into a calibration and a prediction set. The calibration set of samples is used to calibrate the mathematical multivariate methods, the prediction set being used to test the performance of the mathematical model over samples which were not included during the calibration stage.

Although other multivariate techniques are well suited to identify samples from the spectral information,

previous works [35,36,40] have observed that the combination PCA (principal components analysis) + CVA (canonical variate analysis) + *k*NN (*k* nearest neighbor) provides simplicity, a very fast and accurate response (in the order of some milliseconds), thus being a reliable and feasible option for this purpose. CVA is a supervised feature extraction method intended for dimensionality reduction [43,44] in multi-class problems. CVA takes the samples described by the original variables (transmittance at each wavenumber) and determines the directions in space boosting the differences between the analyzed data groups [45,46]. CVA calculates the non-orthogonal latent variables, known as canonical variates (CVs). There are as many CVs as the number of classes minus one. CVA needs data sets with more samples than original variables, this being its main weakness. Since ATR-FTIR spectra have 3351 variables, it is unfeasible to deal with data sets with a higher number of samples. Consequently, the number of variables must be decreased previous to the application of CVA. This dimensionality reduction is carried out by means of the unsupervised PCA algorithm [47,48] before application of CVA. PCA converts the original set of correlated variables, into a reduced number of orthogonal uncorrelated PCs (principal components), which are obtained by linearly combing the original variables. PCA outputs as many PCs as original variables in the spectra. PCA ranks the PCs from highest to lowest variance [49], and only retains the first PCs that explain a suitable amount of the total variance, so the remaining ones are disregarded to avoid overfitting [50]. References [45] and [46,51] detail the mathematical background of CVA and PCA, respectively. Once reduced the dimensionality of the problem by means of the use of the PCA + CVA algorithms, the classification process is performed by applying the supervised *k*NN classifier algorithm. *k*NN assigns unknown incoming samples to their pertinence group according to the majority-voting rule. To this end, for each analyzed prediction sample, the *k*NN algorithm locates the *k* nearest neighbors within the calibration sample set in the space defined by the CVs. Once located, *k*NN gives *k* votes to the nearest neighbor's class and *k* - 1 votes to the class of the second nearest neighbor. This process is repeated until giving one vote to the class of the *k*-th nearest neighbor. Finally, *k*NN assigns the unknown incoming sample to the most voted class [52]. Although the technical literature has explored different values of the parameter *k*, values of *k* in the range 3-6 are often applied [38,53,54]. *k*NN generates as many outputs as classes defined in the problem, which are normalized in the 0-1 range, indicating the membership level to each class of the incoming samples.

It is worth noting that the FDA algorithm, which is in essence a combination of the CVA algorithm applied in this paper jointly with the application of a discriminant function, which is useful for discrimination and assumes identical *a priori* probability for each class, can also be applied to solve this problem [43]. However, the *k*NN algorithm is applied in this paper instead of the discriminant function, since it is one of the most effective classifiers [41,55].

A summary of the classification strategy proposed in this paper to classify unknown incoming NR samples is shown in Fig. 2.

It is noted that to validate the results provided by the PCA + CVA + *k*NN approach, all NR samples were electrically tested according to the requirements of the IEC 60903 international standard [56]. These leakage current tests are routinely applied for quality control purposes in the production line.

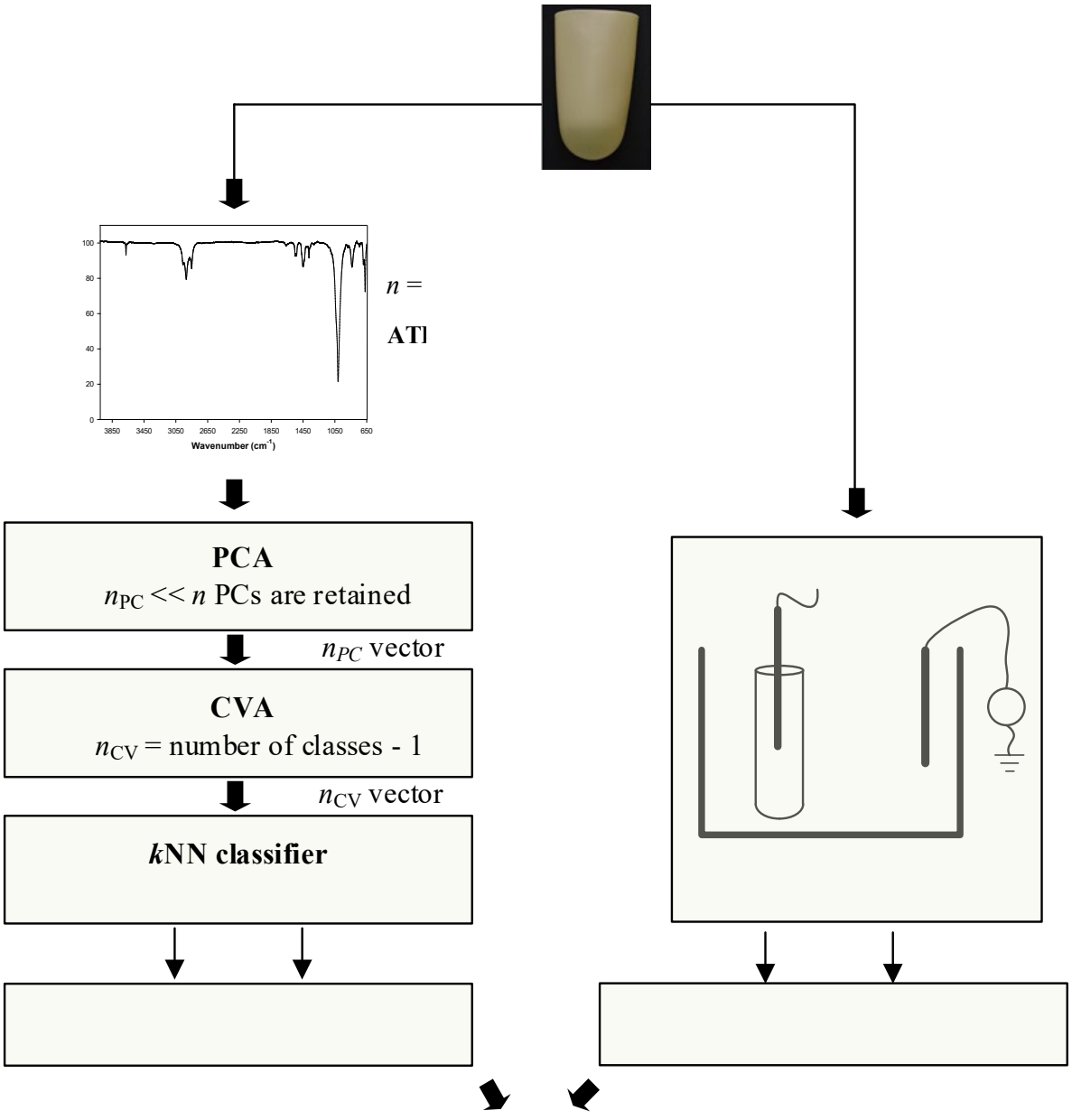

Fig. 2. Process applied to classify an unknown incoming sample from the information of the ATR-FTIR spectrum. Electrical tests were applied to all samples to determine their electrical behavior according to the IEC 60903 international standard [56].

4. RESULTS AND DISCUSSION

The samples studied were analyzed by means of ATR-FTIR spectroscopy. Fig. 3 shows the ATR-FTIR

spectra of two samples of the two analyzed groups, that is, two samples of the SDB group and two more samples of the UDB group.

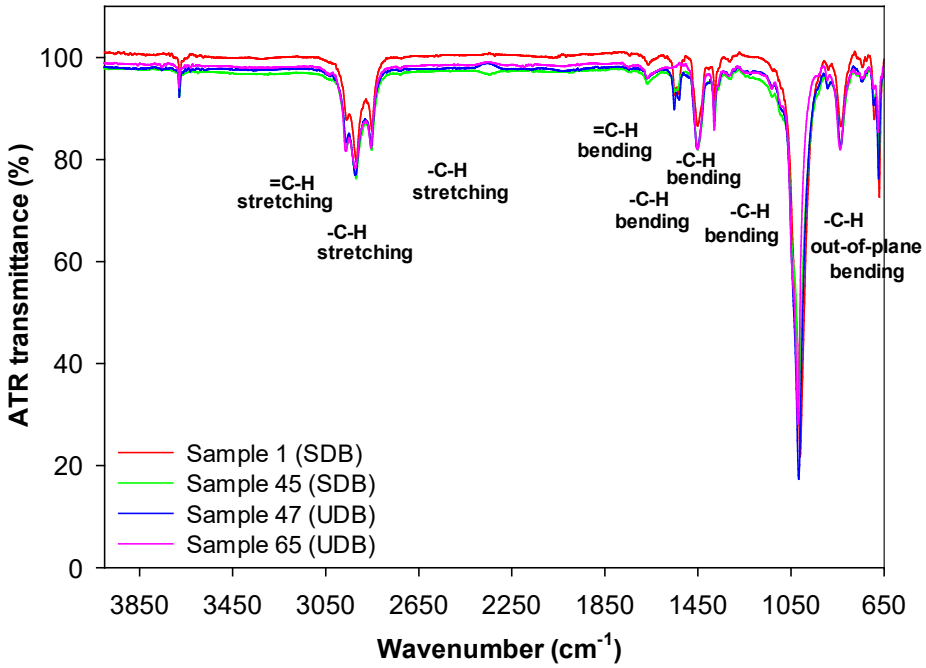

Fig. 3. ATR-FTIR transmittance spectra of some samples analyzed in this work.

Table 1 summarizes the characteristic bands of the NR samples and assigns the corresponding functional groups.

Table 1. Functional groups and wavenumbers of the ATR-FTIR spectra of the insulating NR samples.

Assignments	Wavenumbers, (cm^{-1})	Nature of vibration
-C-H groups	2916, 2846	stretching
	1541, 1449, 1378	bending
	833	out-of-plane bending vibrations
=C-H groups	2957	stretching
	1656	bending

Vulcanized NR include -C-H and =C-H bonds. The more vulcanized the NR is, the double C=C bonds tend to decrease. The band around 1656 cm^{-1} characterizes the double bonds C=C. Since it is very small in all spectra, it indicates that all analyzed samples are vulcanized.

From a direct visual analysis of the ATR-FTIR spectra of the two groups of NR samples, no differences can be observed in the bands of the functional groups indicated in Table 1. Therefore, to perform a classification of incoming unknown samples into one of the two groups for a suitable control and supervision of the NR insulating samples, it requires to analyze the spectral information by means of appropriate multivariate mathematical methods. To this end, the spectra of all samples are sequentially transformed by applying PCA, CVA and finally, the *k*NN algorithm. By this way each unknown incoming sample of the prediction set is assigned to their pertinence classes according to their composition.

The two groups or classes of samples analyzed in this work have different dielectric performance because of the different chemical structure. It is assumed that these subtle differences are revealed by their ATR-FTIR spectra.

This work deals with 78 NR samples, 45 of which exhibit suitable dielectric behavior, whereas the remaining ones show inappropriate or unsuitable dielectric behavior. Therefore, the 78 samples were divided into two groups or classes, that is, suitable (45 samples) and unsuitable (33 samples). Next, the calibration and prediction sets were generated, which include roughly 50% of the NR samples each, which were randomly selected. Table 2 shows the number of samples in each set.

Table 2. NR samples dealt with in this work.

Data set	Number of samples
Calibration set	40
Prediction set	38
SDB (total)	45
UDB (total)	33
SDB (Calibration set)	23
UDB (Calibration set)	17
SDB (Prediction set)	22
UDB (Prediction set)	16

SDB: suitable dielectric behaviour
UDB: unsuitable dielectric behaviour

The initial data matrix containing the raw spectral data has 78 rows (samples) and 3351 columns (wavenumbers interval). So the matrix describing the calibration set has 40x3351 data points, whereas the matrix defining the prediction set includes 38x3351 elements. The data matrixes described above were transformed into their first- and second-derivatives, since this procedure is habitually applied in chemometrics [36,57]. Next subsections describe the results attained by means of the initial raw spectral data matrixes and those from the first and second derivatives.

Experimental current leakage tests according to the setup shown in Fig. 1 were carried out in order to classify the 78 NR samples as SDB or UDB. Table 3 summarizes the results obtained.

Table 3. Results of the leakage current tests to classify the NR samples as SDB or UDB

Maximum thickness (mm)		**0.5 mm**	**1 mm**	**2.3 mm**	**2.9 mm**	**3.6 mm**	**Total samples**
Data set							
SDB (Calibration)	NoS	2	4	4	7	6	23
	Current (mA)	8.1 - 11.9	4.2 - 9	7.6 - 9	8.7 – 11.2	8.2 – 11.5	
SDB (Prediction)	NoS	2	5	5	4	6	22
	Current (mA)	4.1 - 11.9	7.5 - 10	7.5 - 10	8.2 – 10.1	9.3 - 12	
UDB (Calibration)	NoS	4	3	0	2	8	17
	Current (mA)	14.6 - 25	21 - 22	-	34.9 – 35.0	30 - 49	
UDB (Prediction)	NoS	2	3	0	5	6	16
	Current (mA)	15 - 17	18 - 25	-	23 – 40.5	29 - 47	
Current limit (mA) [56]		**12**	**12**	**16**	**18**	**22**	

NoS = number of samples

4.1 Results attained from the raw spectral data

This section classifies the prediction set samples from the raw data of the ATR-FTIR spectra. Firstly, the mathematical classification model is calibrated by means of the calibration set. Since it includes 40 samples, the associated data matrix has 40x3351 data points. Next this matrix is processed via PCA, resulting in a matrix of 40x3351 elements (3351 PCs). After being ranked in descending order with respect to the variance explained, only the 18 first PCs explaining 99.5% of the total variance were considered, thus resulting in a matrix of 40x18 data points, thus allowing to apply the CVA algorithm. Finally, the CVA algorithm was applied to the former matrix, resulting in a matrix of only 40x1 components. It is noted the great reduction of the dimensionality of this problem.

Fig. 4 displays the cumulative variance of the first principal components outputted by the PCA algorithm, as well as the location of the samples of the calibration and prediction sets in the space defined by the only canonical variate arising from the CVA algorithm.

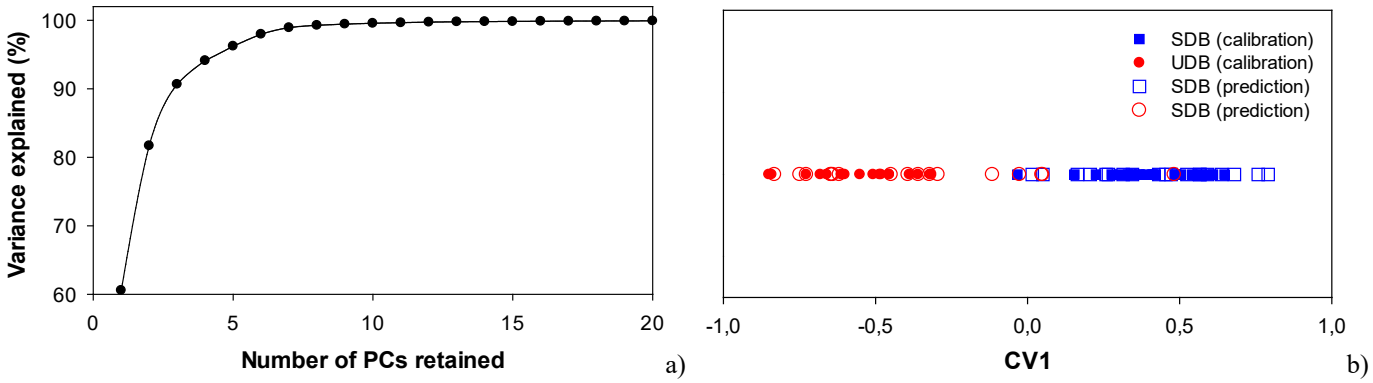

Fig. 4. Classification results attained from the raw spectral data by applying the PCA+CVA algorithms. a) Cumulative variance explained against the number of principal components retained. b) Samples of the calibration and prediction sets (40 and 38 samples, respectively) plotted in the space defined by the only CV after application of the PCA when retaining the first 18 PCs.

Finally, the *k*NN classifier is applied with $k = 3$, 4, 5, and 6, the results being summarized in Table 4.

Table 4. Raw spectral data. Classification results when applying PCA (18 PCs) + CVA (1 CV) + *k*NN to the samples of the prediction set.

k	Correctly classified samples	Success rate (prediction samples)
3	34/38	89.5%
4	35/38	92.1%
5	35/38	92.1%
6	35/38	92.1%

Results shown in Table 4, show a moderately high classification accuracy when applying consecutively PCA (20 PCs) + CVA + *k*NN, since around 90% of the samples in the prediction set are correctly classified.

4.2 Results attained from the first derivative of the spectral data

In this section the samples of the prediction set are classified by taking into account the first derivative of

the spectral data. It is calculated by applying the Savitzky–Golay algorithm, which considers five right-sided and five left-sided spectral data points in each spectrum. Therefore, once the spectral matrix with 40x3351 elements defining the calibration set has been derived, it results in a 40x3341 data matrix with the first derivative of the spectra. It is used to calibrate the classification model and it is transferred to the PCA algorithm. The first 20 PCs are retained, since they explain 99.5% of the variance. Finally, the CVA algorithm is applied.

Fig. 5 shows the cumulative variance of the first principal components generated by the PCA algorithm, and the distribution of the samples of the calibration and prediction sets in the space defined by the canonical variate generated by the CVA algorithm.

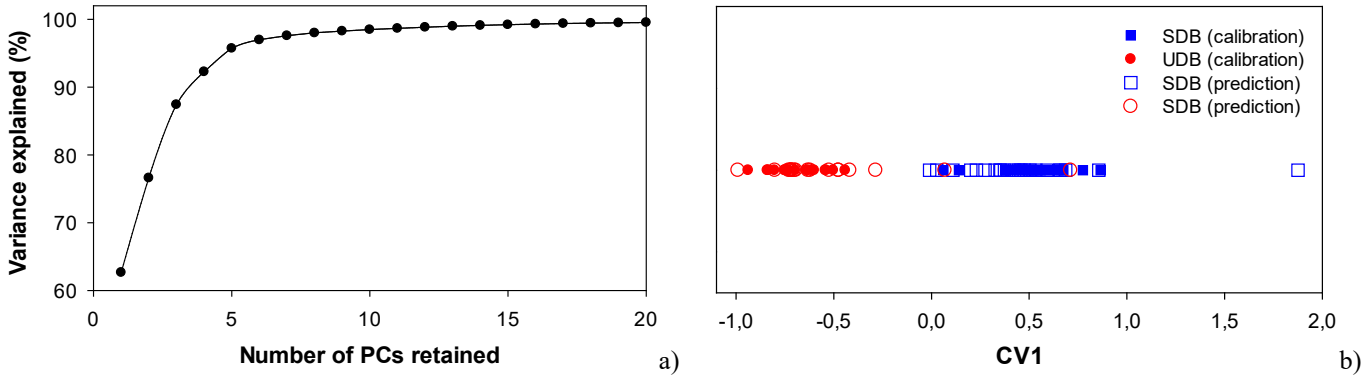

Fig. 5. Classification results attained from the first derivative of the spectral data by applying the PCA+CVA algorithms. a) Cumulative variance explained against the number of principal components retained. b) Samples of the calibration and prediction sets (40 and 38 samples, respectively) plotted in the space defined by the only CV after application of the PCA when retaining the first 20 PCs.

Next, the *k*NN classifier is applied, taking into account k = 3, 4, 5, and 6. These results are summarized in Table 5.

Table 5. First derivative of the spectral data. Classification results when applying PCA (20 PCs) + CVA (1 CV) + *k*NN to the samples of the prediction set.

k	Correctly classified samples	Success rate (prediction samples)
3	36/38	94.7%
4	36/38	94.7%
5	36/38	94.7%
6	36/38	94.7%

Results summarized in Table 5, show a high classification accuracy when applying consecutively PCA (20 PCs) + CVA + *k*NN, since 94.7% of the samples in the prediction set are correctly classified with k = 3, 4, 5 and 6.

4.3 Results attained from the second derivative of the spectral data

The last test was carried out taking into account the second derivative of the spectra, which was obtained by means of the Savitzky–Golay algorithm, considering ten right-sided and ten left-sided data points. After this transformation, the data matrix of the calibration set includes 40x3331 elements, which is the input

data of the PCA algorithm. In this case the first 20 PCs are considered, since they explain 99.5% of the variance, so although PCA calculates 3331 PCs, only the first 20 are retained. Finally, the CVA algorithm is applied.

Fig. 6 illustrates the cumulative variance of the first principal components generated by the PCA algorithm, and the samples distribution of the calibration and prediction sets in the space defined by the canonical variate arising from the CVA algorithm.

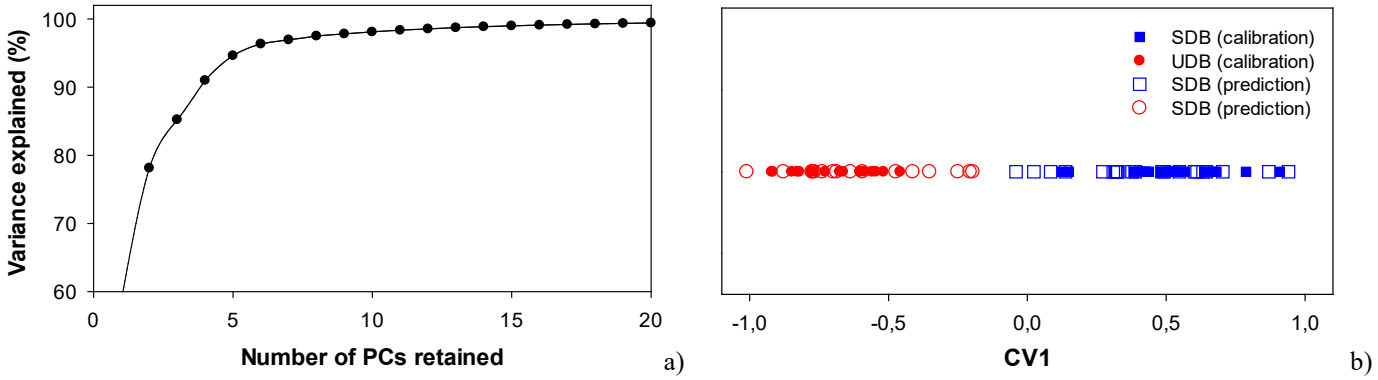

Fig. 6. Classification results attained from the second derivative of the spectral data by applying the PCA+CVA algorithms. a) Cumulative variance explained against the number of principal components retained. b) Samples of the calibration and prediction sets (40 and 38 samples, respectively) plotted in the space defined by the only CV after application of the PCA when retaining the first 20 PCs.

Finally, the *k*NN classifier was applied, with $k = 3$, 4, 5, and 6. The results obtained are shown in Table 6.

Table 6. Second derivative of the spectral data. Classification results when applying PCA (20 PCs) + CVA (1 CV) + *k*NN to the samples of the prediction set.

k	Correctly classified samples	Success rate (prediction samples)
3	38/38	100%
4	38/38	100%
5	38/38	100%
6	38/38	100%

As shown in Table 6, classification results obtained by applying consecutively PCA (20 PCs) + CVA + *k*NN, results in 100% correctly classified samples with $k = 3$, 4, 5 and 6.

4.4 Results summary

This subsection summarizes all results obtained in this paper. For each incoming sample, the *k*NN classifier generates two normalized output values, which corresponds to one output per class. These values are within the interval [0,1]. The output values quantify the degree of membership of the samples to each class. When the output value of an unknown incoming sample is higher than 0.5, it is assumed that this sample belongs to the class corresponding to this output. Output values less than or equal to 0.5 indicate no pertinence of the incoming sample to the class considered. Therefore, extreme values closer to 1 or 0 indicate almost total confidence in the classification result. However, output values denoting membership levels around 0.5, present high uncertainty. To assess the accuracy of the results obtained, the predictive residual error sum of squares (PRESS) is often used as an indicator, which is defined as,

$$\text{PRESS} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (1)$$

y_i [0,1] being the output value of the *k*NN algorithm for the *i*-th prediction sample, and $\hat{y}_i$ is the actual value of y_i. The actual values of $\hat{y}_i$, which are either 1 or 0 are settled by an expert, so they are already known. According to (1), small values of PRESS indicate high accuracy of the classification results.

Table 7 shows the classification results of the samples in the prediction set, obtained by applying the PCA + CVA + *k*NN approach.

Table 7. Summary of results attained with the PCA + CVA + *k*NN approach

Type of data	*k*NN classifier	Classification success rate	PRESS
Raw data (18 PCs)	$k = 3$	34/38 (89.5%)	3.25
	$k = 4$	35/38 (92.1%)	3.17
	$k = 5$	35/38 (92.1%)	3.20
	$k = 6$	35/38 (92.1%)	3.20
First derivative (20 PCs)	$k = 3$	36/38 (94.7%)	2.00
	$k = 4$	36/38 (94.7%)	2.00
	$k = 5$	36/38 (94.7%)	2.00
	$k = 6$	36/38 (94.7%)	2.00
Second derivative (20 PCs)	$k = 3$	38/38 (100%)	0.17
	$k = 4$	38/38 (100%)	0.24
	$k = 5$	38/38 (100%)	0.24
	$k = 6$	38/38 (100%)	0.24

Results summarized in Table 7 clearly indicate that the best results are obtained with the second derivative of the ATR-FTIR spectra. Furthermore, when dealing with the raw spectra and the first derivative of the spectral data, the classification success rate to identify unknown incoming samples of the prediction set is between 89.5% and 94.7%. Results presented in Table 7 also show that in order to speed up the calculations, with $k = 4$ neighbors it is enough, since there are no further improvements with $k = 5$ or 6.

5 CONCLUSION

Natural rubber products applied as insulating medium for high-voltage applications require strict quality controls to ensure suitable dielectric behavior. However, the final properties of such products highly depend on latex origin, compounding and processing conditions, among others. Therefore, it is of paramount importance to develop strict quality control methods to ensure that the final product fulfills all specifications. To this end, this work has suggested a noninvasive, very fast, easy-to-apply and accurate approach to identify vulcanized NR samples with suitable dielectric behavior from those with unsuitable behavior. It is highlighted that this approach does not need any sample treatment, or the use of chemicals or reagents, thus being compatible for continuous quality control of the production process of insulating NR goods. The proposed approach is based on a multivariate chemometric treatment of ATR-FTIR spectral data by means of the PCA+CVA+*k*NN sequence to identify vulcanized NR samples with suitable dielectric

behavior from those with unsuitable dielectric behavior, thus allowing to simplify and expedite quality control methods. Experimental results reported in this paper have shown the suitability and applicability of this approach since the success rate in classifying unknown incoming NR samples can be as high as 100%. The results presented in this paper prove the potential of the proposed approach to be applied in the NR industry to control industrial processes.

REFERENCES

[1] N. González, M. dels À. Custal, G.N. Tomara, G.C. Psarras, J.-R. Riba, E. Armelin, Dielectric response of vulcanized natural rubber containing BaTiO3 filler: The role of particle functionalization, Eur. Polym. J. 97 (2017) 57–67. doi:10.1016/J.EURPOLYMJ.2017.10.001.

[2] N. González, M. del À. Custal, S. Lalaouna, J.-R. Riba, E. Armelin, Improvement of dielectric properties of natural rubber by adding perovskite nanoparticles, Eur. Polym. J. 75 (2016) 210–222. doi:10.1016/j.eurpolymj.2015.12.023.

[3] N. Wang, L. Mi, Y. Wu, J. Zhang, Q. Fang, Double-layered co-microencapsulated ammonium polyphosphate and mesoporous MCM-41 in intumescent flame-retardant natural rubber composites, J. Therm. Anal. Calorim. 115 (2014) 1173–1181. doi:10.1007/s10973-013-3404-9.

[4] M. Mariano, N. El Kissi, A. Dufresne, Cellulose nanocrystal reinforced oxidized natural rubber nanocomposites, Carbohydr. Polym. 137 (2016) 174–183. doi:10.1016/J.CARBPOL.2015.10.027.

[5] X. Wu, T.F. Lin, Z.H. Tang, B.C. Guo, G.S. Huang, Natural rubber/graphene oxide composites: Effect of sheet size on mechanical properties and straininduced crystallization behavior, Express Polym. Lett. 9 (2015) 672–685. doi:10.3144/expresspolymlett.2015.63.

[6] C.P. Davi, L.F.M.D. Galdino, P. Borelli, O.N. Oliveira, M. Ferreira, Natural rubber latex LbL films: Characterization and growth of fibroblasts, J. Appl. Polym. Sci. 125 (2012) 2137–2147. doi:10.1002/app.36309.

[7] F. Carpi, I. Anderson, S. Bauer, G. Frediani, G. Gallone, M. Gei, C. Graaf, C. Jean-Mistral, W. Kaal, G. Kofod, M. Kollosche, R. Kornbluh, B. Lassen, M. Matysek, S. Michel, S. Nowak, B. O'Brien, Q. Pei, R. Pelrine, B. Rechenbach, S. Rosset, H. Shea, Standards for dielectric elastomer transducers, Smart Mater. Struct. 24 (2015) 105025. doi:10.1088/0964-1726/24/10/105025.

[8] A. Patsidis, G.C. Psarras, Dielectric behavior and functionality of polymer matrix-ceramic BaTiO3 composites, EXPRESS Polym. Lett. 2 (2008) 718–726. doi:10.3144/expresspolymlett.2008.85.

[9] N. González, M. del À. Custal, D. Rodríguez, J.-R. Riba, E. Armelin, N. González, M. del À. Custal, D. Rodríguez, J.-R. Riba, E. Armelin, Influence of ZnO and TiO2 Particle Sizes in the Mechanical and Dielectric Properties of Vulcanized Rubber, Mater. Res. 20 (2017) 1082–1091. doi:10.1590/1980-5373-mr-2017-0178.

[10] C.A. Harper, Handbook of plastics technologies : the complete guide to properties and performance, McGraw-Hill, 2006. https://www.accessengineeringlibrary.com/browse/handbook-of-plastics-technologies-the-complete-guide-to-properties-and-performance (accessed April 26, 2018).

[11] B. Seentrakoon, B. Junhasavasdikul, W. Chavasiri, Enhanced UV-protection and antibacterial properties of natural rubber/rutile-TiO2 nanocomposites, Polym. Degrad. Stab. 98 (2013) 566–578. doi:10.1016/J.POLYMDEGRADSTAB.2012.11.018.

[12] N.A.M. Jamail, M.A.M. Piah;, N.A. Muhamad;, Z. Salam;, N.F. Kasri;, R.A. Zainir;, Q.E. Kamarudin;, Effect of Nanofillers on the Polarization and Depolarization Current Characteristics of New LLDPE-NR Compound for High Voltage Application, Adv. Mater. Sci. Eng. 2014 (2014) 1–7. doi:http://dx.doi.org/10.1155/2014/416420.

[13] N. Gonzalez, J.-R. Riba, M. Dels Angels Custal, E. Armelin, Improvement of insulation effectiveness of natural rubber by adding hydroxyl-functionalized barium titanate nanoparticles, IEEE Trans. Dielectr. Electr. Insul. 24 (2017) 2881–2889. doi:10.1109/TDEI.2017.006176.

[14] G. Cain, Rubber-insulating goods, in: 2011 IEEE PES 12th Int. Conf. Transm. Distrib. Constr. Oper. Live-Line Maint., IEEE, Providence, Rhode Island, USA, 2011: pp. 1–8. doi:10.1109/TDCLLM.2011.6042231.

[15] M. Amin, S. Amin, M. Ali, Monitoring of leakage current for composite insulators and electrical devices, Rev.Adv.Mater.Sci. 21 (2009) 75–89.

[16] S. Shunmugam, Live maintenance in high voltage substations: Malaysia's experience, in: 2014 11th Int. Conf. Live Maint., IEEE, Zagreb, Croatia, 2014: pp. 1–4. doi:10.1109/ICOLIM.2014.6934331.

[17] A. Ladhar, M. Arous, H. Kaddami, M. Raihane, A. Kallel, M.P.F. Graça, L.C. Costa, AC and DC electrical conductivity in natural rubber/nanofibrillated cellulose nanocomposites, J. Mol. Liq. 209 (2015) 272–279. doi:10.1016/j.molliq.2015.04.020.

[18] A. Choudhury, Preparation, characterization and dielectric properties of polyetherimide nanocomposites containing surface-functionalized BaTiO3 nanoparticles, Polym. Int. 61 (2012) 696–702. doi:10.1002/pi.4181.

[19] P. Tewari, Interfacial effects in oxide-polymer laminar composite thin film dielectrics for capacitor applications, The

Pennsylvania State University, 2009. https://etda.libraries.psu.edu/catalog/9714 (accessed June 25, 2016).
[20] S. Yu et al., Improving the dielectric properties of poly(vinylidene fluoride) composites by using poly(vinyl pyrrolidone)-encapsulated polyaniline nanorods, J. Mater. Chem. C. 4 (2016) 1504–1510. doi:10.1039/C5TC04026D.
[21] S.A. Paniagua, Y. Kim, K. Henry, R. Kumar, J.W. Perry, S.R. Marder, Surface-Initiated Polymerization from Barium Titanate Nanoparticles for Hybrid Dielectric Capacitors, ACS Appl. Mater. Interfaces. 6 (2014) 3477–3482. doi:10.1021/am4056276.
[22] D. Ma, R.W. Siegel, J.-I. Hong, L.S. Schadler, E. Mårtensson, C. Önneby, Influence of Nanoparticle Surfaces on the Electrical Breakdown Strength of Nanoparticle-Filled Low-Density Polyethylene, J. Mater. Res. 19 (2004) 857–863. doi:10.1557/jmr.2004.19.3.857.
[23] S. Javadi, M. Sadroddini, M. Razzaghi-Kashani, P.N.B. Reis, A.A. Balado, Interfacial effects on dielectric properties of ethylene propylene rubber–titania nano- and micro-composites, J. Polym. Res. 22 (2015) 162. doi:10.1007/s10965-015-0805-4.
[24] S. Dalle Vacche, F. Oliveira, Y. Leterrier, V. Michaud, D. Damjanovic, J.-A.E. Månson, Effect of silane coupling agent on the morphology, structure, and properties of poly(vinylidene fluoride–trifluoroethylene)/BaTiO3 composites, J. Mater. Sci. 49 (2014) 4552–4564. doi:10.1007/s10853-014-8155-x.
[25] Z.-M. Dang, H.-Y. Wang, Y.-H. Zhang, J.-Q. Qi, Morphology and Dielectric Property of Homogenous BaTiO3/PVDF Nanocomposites Prepared via the Natural Adsorption Action of Nanosized BaTiO3, Macromol. Rapid Commun. 26 (2005) 1185–1189. doi:10.1002/marc.200500137.
[26] I.A. Asimakopoulos, G.C. Psarras, L. Zoumpoulakis, Barium titanate/polyester resin nanocomposites: Development, structure-properties relationship and energy storage capability, EXPRESS Polym. Lett. 8 (2014) 692–707. doi:10.3144/expresspolymlett.2014.72.
[27] L. Jiang, A. Betts, D. Kennedy, S. Jerrams, Improving the electromechanical performance of dielectric elastomers using silicone rubber and dopamine coated barium titanate, Mater. Des. 85 (2015) 733–742. doi:10.1016/J.MATDES.2015.07.075.
[28] S. Salaeh, G. Boiteux, P. Cassagnau, C. Nakason, Flexible 0-3 Ceramic-Polymer Composites of Barium Titanate and Epoxidized Natural Rubber, Int. J. Appl. Ceram. Technol. 12 (2015) 106–115. doi:10.1111/ijac.12129.
[29] S. Rolere, S. Liengprayoon, L. Vaysse, J. Sainte-Beuve, F. Bonfils, Investigating natural rubber composition with Fourier Transform Infrared (FT-IR) spectroscopy: A rapid and non-destructive method to determine both protein and lipid contents simultaneously, Polym. Test. 43 (2015) 83–93. http://linkinghub.elsevier.com/retrieve/pii/S0142941815000525 (accessed March 22, 2017).
[30] Jin Huang, R. Gosangi, R. Gutierrez-Osuna, Active Concentration-Independent Chemical Identification With a Tunable Infrared Sensor, IEEE Sens. J. 12 (2012) 3135–3142. doi:10.1109/JSEN.2012.2212186.
[31] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, E. Bertran, Determination of accelerators and antioxidants in vulcanized rubber by fourier transform infrared spectrophotometry, Anal. Chim. Acta. 353 (1997) 351–358. doi:10.1016/S0003-2670(97)87797-5.
[32] G.K. Jana, C.K. Das, Recycling natural rubber vulcanizates through mechanochemical devulcanization, Macromol. Res. 13 (2013) 30–38. doi:10.1007/BF03219012.
[33] C. Tzoganakis, Q. Zhang, Devulcanization of Recycled Tire Rubber Using Supercritical Carbon Dioxide | Society of Plastics Engineers, in: Glob. Plast. Environ. Conf., 2004: pp. 1–11. http://www.4spe.org/Resources/resource.aspx?ItemNumber=7898 (accessed November 6, 2015).
[34] X. Hou, S. Lv, Z. Chen, F. Xiao, Applications of Fourier transform infrared spectroscopy technologies on asphalt materials, Measurement. 121 (2018) 304–316. doi:10.1016/J.MEASUREMENT.2018.03.001.
[35] J.-R. Riba, T. Canals, R. Cantero, Supervision of Ethylene Propylene Diene M-Class (EPDM) Rubber Vulcanization and Recovery Processes Using Attenuated Total Reflection Fourier Transform Infrared (ATR FT-IR) Spectroscopy and Multivariate Analysis, Appl. Spectrosc. 71 (2017). doi:10.1177/0003702816653131.
[36] J.R. Riba, J. Cailloux, R. Cantero, T. Canals, M.L. Maspoch, Multivariable methods applied to FTIR: A powerful technique to highlight architectural changes in poly(lactic acid), Polym. Test. 65 (2018). doi:10.1016/j.polymertesting.2017.12.003.
[37] L.E. Rodriguez-Saona, M.M. Giusti, M. Shotts, 4 – Advances in Infrared Spectroscopy for Food Authenticity Testing, in: Adv. Food Authent. Test., 2016: pp. 71–116. doi:10.1016/B978-0-08-100220-9.00004-7.
[38] J.-R. Riba, T. Canals, R. Cantero, Supervision of Ethylene Propylene Diene M-Class (EPDM) Rubber Vulcanization and Recovery Processes Using Attenuated Total Reflection Fourier Transform Infrared (ATR FT-IR) Spectroscopy and Multivariate Analysis, Appl. Spectrosc. 71 (2017) 141–151. doi:10.1177/0003702816653131.
[39] J.-R. Riba, T. Canals, R. Cantero, H. Iturriaga, Potential of infrared spectroscopy in combination with extended canonical variate analysis for identifying different paper types, Meas. Sci. Technol. 22 (2011) 025601. doi:10.1088/0957-0233/22/2/025601.
[40] J.-R. Riba, T. Canals, R. Cantero, Recovered Paperboard Samples Identification by Means of Mid-Infrared Sensors, IEEE Sens. J. 13 (2013) 2763–2770. doi:10.1109/JSEN.2013.2257943.
[41] J.-R. Riba, T. Canals, R. Gómez, Comparative Study of Multivariate Methods to Identify Paper Finishes Using Infrared Spectroscopy, IEEE Trans. Instrum. Meas. 61 (2012) 1029–1036. doi:10.1109/TIM.2011.2173048.

[42] J.-C.J.-C. Urresty, R. Atashkhooei, J.-R.J.-R. Riba, L. Romeral, S. Royo, No Title, IEEE Trans. Ind. Electron. 60 (2013) 3454–3461. doi:10.1109/TIE.2012.2213565.
[43] B. Jiang, X. Zhu, D. Huang, J.A. Paulson, R.D. Braatz, A combined canonical variate analysis and Fisher discriminant analysis (CVA–FDA) approach for fault diagnosis, Comput. Chem. Eng. 77 (2015) 1–9. doi:10.1016/J.COMPCHEMENG.2015.03.001.
[44] Q. Lu, B. Jiang, R.B. Gopaluni, P.D. Loewen, R.D. Braatz, Locality preserving discriminative canonical variate analysis for fault diagnosis, Comput. Chem. Eng. 117 (2018) 309–319. doi:10.1016/J.COMPCHEMENG.2018.06.017.
[45] L. Nørgaard, R. Bro, F. Westad, S.B. Engelsen, A modification of canonical variates analysis to handle highly collinear multivariate data, J. Chemom. 20 (2006) 425–435. doi:10.1002/cem.1017.
[46] R. A. Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis, 6th editio, Prentice-Hall, Englewood Cliffs, NJ (USA), 2007.
[47] W. Lai, X. Zeng, J. He, Y. Deng, Aesthetic defect characterization of a polymeric polarizer via structured light illumination, Polym. Test. 53 (2016) 51–57. doi:10.1016/j.polymertesting.2016.05.011.
[48] D. Song, J. Gao, X. Li, L. Lu, Evaluation of aging behavior of polypropylene in natural environment by principal component analysis, Polym. Test. 33 (2014) 131–137. doi:10.1016/j.polymertesting.2013.11.014.
[49] M. Werteker, S. Huber, S. Kuchling, B. Rossmann, M. Schreiner, Differentiation of milk by fatty acid spectra and principal component analysis, Measurement. 98 (2017) 311–320. doi:10.1016/J.MEASUREMENT.2016.10.059.
[50] N. Bhattacharyya, R. Bandyopadhyay, M. Bhuyan, B. Tudu, D. Ghosh, A. Jana, Electronic Nose for Black Tea Classification and Correlation of Measurements With “Tea Taster” Marks, IEEE Trans. Instrum. Meas. 57 (2008) 1313–1321. doi:10.1109/TIM.2008.917189.
[51] A. Sánchez-Fernández, G.I. Sainz-Palmero, J.M. Benítez, M.J. Fuente, Linguistic OWA and two time-windows based fault identification in wide plants, Comput. Chem. Eng. 115 (2018) 412–430. doi:10.1016/J.COMPCHEMENG.2018.04.020.
[52] D. Ha, U. Ahmed, H. Pyun, C.-J. Lee, K.H. Baek, C. Han, Multi-mode operation of principal component analysis with k-nearest neighbor algorithm to monitor compressors for liquefied natural gas mixed refrigerant processes, Comput. Chem. Eng. 106 (2017) 96–105. doi:10.1016/J.COMPCHEMENG.2017.05.029.
[53] L.A. Berrueta, R.M. Alonso-Salces, K. Héberger, Supervised pattern recognition in food analysis, J. Chromatogr. A. 1158 (2007) 196–214. doi:10.1016/j.chroma.2007.05.024.
[54] J.-R. Riba Ruiz, T. Canals Parello, R. Cantero Gomez, Identification of NR and EPDM Samples by Means of Thermogravimetric Analysis and Multivariate Methods, IEEE Sens. J. 16 (2016) 7705–7712. doi:10.1109/JSEN.2016.2603172.
[55] M. Hammad, A. Maher, K. Wang, F. Jiang, M. Amrani, Detection of abnormal heart conditions based on characteristics of ECG signals, Measurement. 125 (2018) 634–644. doi:10.1016/J.MEASUREMENT.2018.05.033.
[56] IEC, IEC 60903:2014. Live working - Electrical insulating gloves, (2014) 117. https://webstore.iec.ch/publication/3871 (accessed September 4, 2018).
[57] T. Canals, J.R. Riba, R. Cantero, J. Cansino, D. Domingo, H. Iturriaga, Characterization of paper finishes by use of infrared spectroscopy in combination with canonical variate analysis, Talanta. 77 (2008) 751–757. doi:10.1016/j.talanta.2008.07.059.