Journal of the American Society of Questioned Document Examiners, Inc.
Volume 22, Number 1, June 2019

Application of Multivariate Statistics (PCA and HCA) on ballpoint pen ink Infrared spectra for dating of forensic relevant documents

Carina Maria Bello de Carvalho [a, b, d], Moniquedos Reis[b], Rafael Scorsatto Ortiz[a, d], Marco Flores Ferrão[c], Renata Pereira Limberger[b, d]

[a]Rio Grande do Sul Technical and Scientifical Division, Brazilian Federal Police, Porto Alegre, RS, Brazil
[b]Department of Pharmacy, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil.
[c]Institute of Chemistry, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil.
[d] Brazilian National Institute of Forensic Science and Technology, 90160-030 Porto Alegre, RS, Brazil.

In the normal casework of Forensic Document Examiners (FDEs), a common type of forgery is the alteration of the date of creation of a document through the erasure and addition of another date with ballpoint pen of color similar to the original. The chronology of manuscripts in a document has been a concern within the FDE community. Recently, Infrared spectroscopy (IR) and multivariate statistical analysis have been used to differentiate pens, but this paper aimed to develop an IR method that could also date a manuscript by multivariate statistic. Ballpoint pen lines from different brands made in different kinds of papers (office white and recycled), with variable ages, were analyzed by Infrared Spectroscopy with Total Attenuated Reflectance Accessory. The spectra obtained were analyzed with multivariate statistical analysis, especially Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA). This study concludes that different kinds of paper do not influence the characterization of the pens by IR-ATR. The method was able to characterize the variables that were responsible for the discrimination of ballpoint pens from different brands/models and for differentiation of samples of different ages. The HCA and PCA were able to group pens with similar ages, and differentiate lines recently wrote from older lines. Within the loadings that are responsible for the differentiation, it was possible to identify absorption peaks from 2-phenoxyethanol, a solvent whose evaporation from the ink could indicate the line age. Further studies regarding other pen samples and multivariate statistical approaches could be conducted to confirm these findings.

Keywords: Forensic Document Examination; ballpoint pen; document dating; infrared spectroscopy; Attenuated Reflectance; ink; Multivariate Statistics;

Introduction

A common type of alteration in the date of creation of a document is the erasure of the original date and addition of another date with a ballpoint pen of color similar to original. The chronology of manuscripts in a document has been a concern within the Forensic Document Examination (FDE) community. It is reported that around 80% of questioned documents requiring analysis contain ballpoint pen ink[1].

The dyes used in blue and black ballpoint pens are basic dyes based on triarylmethane and acid dyes derived from diazo compounds or phthalocyanine. Both types are ionic in nature, with the basic and acid dyes generally containing iminium and sulfonate groups, respectively[2]. The dyes and pigments (organic and/or inorganic) make up about 25% of the formulation, while the solvent makes up about 50% by weight. The remainder is a variety of additives such as resins, viscosity ad-

Corresponding Author: e-mail address carina.cmbc@dpf.gov.br (Carina M. Bello de Carvalho)

justers, antioxidants, surfactants, softeners, and lubricants[3].

To determine each pen ink formulation in a way to differentiate and date the ink strokes, many destructive methods such as high performance liquid chromatography (HPLC) coupled to Diode Array or Ultraviolet Detection[4-8] or Mass Spectrometry detection[9,10] have been used, and more recently high resolution mass spectrometry (HMRS) like Orbitrap[11] and Time-of-Flight Second Ion Mass Spectrometry (TOF-SIMS)[12]. The destructive analysis starts by removal of a small section of the ink line followed by solvent extraction of the ink[13]. However, some documents are much important to allow cutting a piece of the paper, and here we recommend the option of non-destructive methods as Infrared Spectroscopy with Total Attenuated Reflectance Accessory (IR-ATR). Also, although many components of the ink may be collected following analysis by certain chemical techniques, the ink itself cannot be recovered in its original form.[14]

Non-destructive methods have been used to characterize and differentiate ballpoint pen inks. These include Diffuse Reflectance Infrared Fourier Transform (DRIFTS)[2, 14, 15]; IR-ATR[2, 16, 17] ; Micro-FTIR (Fourier Transform Infra Red) spectroscopy[18, 19] ; Raman spectroscopy[19] or Ultraviolet-Visible (UV-Vis) and Near-Infrared (NIR) chemical imaging methods[20, 21] , which are used to analyze the chemical composition of ballpoint pen inks.

Wang *,et al.*, Payne *,et al.,* Zieba Palus and Kunicki, differentiated the ballpoint pen inks with good precision using a pattern recognition system and a Discriminant Power factor, while Silva *,et al,* Thanasoulias *,et al,* Adam *,et al.* Senior *,et al,* and Kher *,et al* invested in multivariate statistics to differentiate the pens, obtaining good results[2, 13, 16, 19-24].

A chromatogram or a spectrum may be visualized as a pattern in multivariate space. Samples displaying similar features often cluster together and those displaying dissimilar features are located away from each other in multivariate space. By knowing this concept, the Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) of each ink mass spectrum can be used to differentiate the ballpoint pens[2].

Many studies of ballpoint pen ink characterization have been conducted with ballpoint pens from European manufacturers. However, the pens in the Brazilian market have different manufacturers and because of this, different formulations.

Microspectrometry[19] is a method that was applied to achieve differentiation between ballpoint pens. Some researchers tried to improve the discrimination power of microspectrometry on transferring a small area of inked fibers on a slide and smearing using an engraving tool or immersing in a mounting medium, but the type of paper influenced the results when the blue and black ink traces on white or brown paper were smeared on glass slides[25].

Some researchers are attempting to differentiate ballpoint pen inks by FTIR or UV/VIS of inks sampled direct from the pen cartridge or extracted from paper[14, 15, 22-24], and some other researchers published papers regarding FTIR of ballpoint ink pens analyzing them directly on paper[2, 12, 16, 18, 19]. From those, only Silva *,et al.* was concerned about the paper influence in FTIR-ATR spectra of the ink [16]. More recently, Lee *,et al.* observed that FTIR-ATR is a suitable method for determining intersecting lines sequence because the IR penetrates only a short distance into the ink sample, so that the spectra would contain less contribution from the paper background[12].

Other attempts to analyze nondestructively ballpoint ink on paper using a FTIR microscope were also unsuccessful, because strong absorption from the paper tended to mask any absorption contributed by the ink.[14, 25]With this past research in mind, the present study aimed to apply multivariate statistical methods to spectra data from FTIR-ATR of ballpoint pen strokes in different papers, in a way to clarify if the paper does influence the spectra of inks.

Besides of the differentiation of ballpoint inks, the dating of the manuscripts also is a challenge to forensic researchers. Until now, only Senior, *et al*. used Principal Component Analysis (PCA) of UV/VIS Spectroscopy with a dating objective[13]. However, they only did the dating on ballpoint ink collected directly from the cartridge and from reference dyes[13]. Silva*, et al* 2013, Feraru*, et al.*, Wang*, et al.* and more recently Sharma and Kumar identified in the ink spectra absorption peaks from 2- phenoxietanol and other solvents, which evaporation from the ink stroke could lead to dating of the manuscript16, 17, 22, 26

This study aimed to perform analysis of ink strokes on paper, of several models of blue and black ballpoint pens, by FTIR-ATR, to differentiate the pens by multivariate statistical (PCA and HCA). For the most representative pens in groups formed by PCA, this study follows the ageing of the pen strokes, acquiring IR-ATR spectra on several dates and performing multivariate statistic (PCA) to differentiate the pens by date. The contribution from the paper background for the FTIR-ATR spectra was also tested.

Materials and methods

Thirty-seven different brands/models of blue ballpoint pens and twenty-seven different brands/models of black ballpoint pens (Bic, Stabilo, Injex, Faber Castell, Pilot, Cis, Uni, BRW, Staedtler, Molin, Office, Mon Ami, Jocar, Tilibra, Pentel, Ita, Tris, Compactor, Masterprint, Paper Mate, NewPen) were applied as strokes, with a ruler, on white office paper (75 g/m^2) and on recycled office paper (75 g/m^2). The brands of pens analyzed were chosen because they are the most popular in Brazilian markets (Tables 01 and 02). Some models of pens had two individuals from different batches tested, to see if the method could group these individuals despite the manufacture batch.

The freshly applied strokes were analyzed on a Fourier Transform Infrared Spectrometer Nicolet 380 (Thermo), with an Attenuated Total Reflectance device, and the spectra obtained (three replicates) were analyzed through multivariate statistic Hierarchical Cluster Analysis (HCA) and Principal Component Analysis (PCA). The acquisition parameters of the spectra were 32 scans made from 4000 to 525 cm^{-1}, with a resolution of 4 cm^{-1}. The data were then processed with Chemostat software[27] and some representative brands/models (eight blue pens and six black pens), of similar formulation, were chosen to be analyzed, using the same method: recently applied, after 14 days (for blue pens only), 7 months, 1 year (for blue pens), 2 years and 2 years and 3 months (for blue pens) or 2 years and 6 months (for black pens). These time intervals were chosen based on the aged pen strokes collection already existing in our Laboratory[28]. Three replicates of spectra of each individual pen obtained in each date were analyzed together by multivariate statistics (performing PCA). During the ageing, the samples of the pen strokes collection were stored in darkness, in a controlled temperature of 20 ºC, and 30% of relative humidity.

Additionally, to evaluate the influence of the support paper in the spectra of ballpoint pen ink line, FTIR-ATR spectra of different papers (Table 03) were obtained in six replicates, and these spectra were compared with the Principal Component Loadings obtained in the multivariate analysis of spectra from ballpoint pen lines made in each kind of paper. Freshly applied blue and black lines were made in Paper 14 (see Table 3), and there were analyzed blue lines with 2 years and 3 months of age made in paper 1 and in paper 2; and black lines with 2 years and 3 months of age made in paper 6 and in paper 8.

Another experiment was performed using the same brand/models of ballpoint pens. The ink was collected directly from the pen cartridge with a needle. The sample was subsequently analyzed in a Fourier Transform Infrared Spectrometer Nicolet iS50 (Thermo), with an Attenuated Total Reflectance device, installed at the Customer Experience Centre, in Thermo Fischer Scientific-São Paulo, Brazil. The spectra of all pens analyzed were then processed with Chemostat software[27] and the HCA results were compared with the HCA results obtained for the spectra of the same pens but wrote as lines in the papers, as described previously. This experiment is important to confirm if the power of discrimination of different ballpoint pens by FTIR-ATR using multivariate statistics is compatible with ink from the pen cartridge and ink lines on a paper substrate.

Table 1. Blue pens Analyzed.

Pen code	Brand/model
A1	Bic ecolutions round stic
A2	Bic fine point
A3	Bic cristal pocket
A4	Faber Castell trilux 032 Medium
A5	Faber Castell trilux 035 Fine
A6	Faber Castel EC-07
A7	Faber Castell Medium
A8	Compactor 07 fine
A9	Compactor TOP2000
A10	Pilot BPS 0.7 AZ
A11	Injex Pen new
A11.2	Injex Pen used
A12	BRW new
A12.1	BRW used
A13	Pilot Super Grip 1.0
A14	Tris Hit Blue
A15	Tris Exec 1.0
A16	CIS Silver Stick 1.0
A17	Pentel Star V 0.7– India

Table 1. (*Continued*).

Pen code	Brand/model
A19	Uni Laknok 1.4 (Japan)
A20	Office
A21	Stabilo Excel 828M
A22	Stabilo Tropikana (Germany)
A23	Paper Mate Kilometrica 100 1.0
A24	New Pen Clic
A25	Master Print
A26	Bic Cristal
A27	Unknown brand
A28	Compactor Economic
A29	Pilot Super Grip nova 1.6
A30	Bic Atlantis 1.6
A31	Tilibra super BP 1.0
A32	Uni SAS Fine (Japan)
A33	Uni Lakubo 1.0 (Japan)
A35	Stabilo bille 508 (Germany)
A36	Molin CR-45 (France)

Table 2. Black pens analysed.

Pen Code	Brand/model
P1	Bic cristal pocket
P2	Bic ECOLUTIONS round stic
P3	Bic Diamond
P4	Bic cristal
P5.1	Injex Pen used
P6	Faber Castel trilux 032 Medium
P7	Faber Castell fine point
P8	Pilot BPS 0.7 fine point
P9	Stabilo Excel 828 M
P10	BRW new
P10.2	BRW used
P11.1	Molin (France)
P12	Uni Laknok 0.7 fine point
P13	Ita

Table 2. (*Continued*).

Pen Code	Brand/model
P14	Cis Neotip 1.0
P15	Pentel Star V 0.7 – India
P16	Paper Mate Kilometrica 1.0
P17	Jocar Office
P18	Faber Castel trilux 035 fine
P19	Bic Atlantis 1.6
P20	Pilot Super Grip 1.6
P21	Tilibra Super BP 1.0
P22	Staedler Ball 432 M
P23	Molin 0.7 trion (France)
P25	Uni SAS Fine – Japan
P26	Stabilo M Bille 508 (Germany)
P27	Cis Speed 1.0

Table 3. Kind of papers analyzed.

Paper Code	Characteristics
Paper 1	White office paper 75 g/m^2 (Chamex®)
Paper 2	Recycled office paper 75 g/m^2 (Chamex®)
Paper 6	White office paper 75 g/m^2 (Premiatto®)
Paper 8	Recycled office paper 80 g/m^2 (Beller®)
Paper 14	White office paper 75 g/m^2 (One®)

Results and Discussion

In relation to the paper influence experiment, the results showed that the loadings of the ballpoint pens PC1, PC2, PC3, PC4 and so on, did not reproduce the absorption peaks of the paper spectra for all kinds of papers tested (Figure 1). These findings assure that the paper did not influence in the spectra acquisition for FTIR-ATR analysis, which concurs with the findings of Lee, *et al.*[12]. Silva, *et al.*[16] also tested three different papers to evaluate influence of paper on predictions for the Linear Discriminant Analysis test set, but they used a different approach to evaluate this influence[16].

In their study, Silva, *et al.*[16] tested three different variable selection procedures to perform the LDA. They used Successive Projections Algorithm (SPA), SW approach and Genetic Algorithm (GA) to define the variables for blue inks in paper 1. The PCA was performed after preprocessing and paper subtraction. Depending on the variable selection procedure used, the percent of correct classification results was different for the paper 1, paper 2 (both white paper) and paper 3 (recycled paper). However, for paper 1, the results obtained by Silva, *et al.*[16] were superior of those obtained by Kher, *et al.*2.

Once the paper showed no effect in the spectra of pen strokes, the spectra collected from the ballpoint pen strokes freshly applied in the paper 14 (White office paper 75 g/m^2 One®) were subjected to multivariate statistical analysis. Before performing PCA and HCA, the data was pre-treated with Savitzky-Golay transform and normalized by range 0-1. The HCA used the method of complete linkage to determine the Euclidian Distance of the samples.

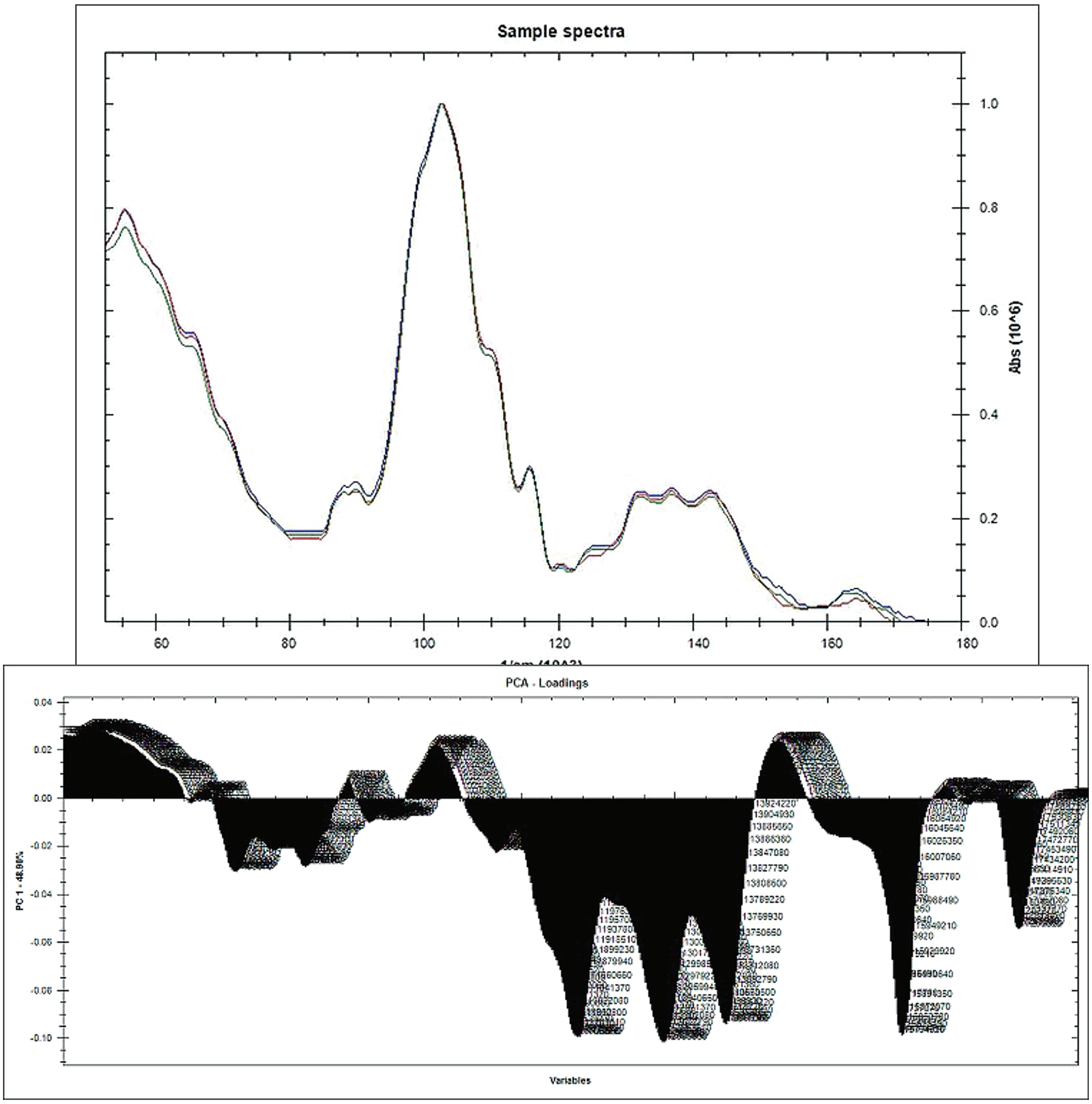

Figure 1. Illustrative image of paper influence experiment. Above: spectra of Paper 14. Below: PC1 loadings of blue pens.

Most of the variability in the region between 4000 and 2000 cm^{-1} is due to atmospheric constituents such as water vapor and carbon dioxide[2]. Typical cellulose bands corresponding to C-H and O-H stretching can be found between 3300 and 2900 cm^{-1} and print papers may contain spectral features by the presence of additives, such as calcium carbonate (around 1430, 875 and 712 cm^{-1})[14, 16].

Some authors that analyzed ink directly on paper substrates had taken the paper spectra and subtracted it from the ink stroke spectra, to avoid the effect of paper background[15-18]. In the present study, this subtraction of spectra was attempted; however, it was found that this process lead to a loss of important information from the spectra, once some absorption bands from paper coincide with absorption bands from ink dyes.

The most important region for ink analysis is from 1800 cm^{-1} and 650 cm^{-1}, so this study concentrated on this range. For the blue pens recently applied, the PC1 versus PC2 plot accounted for 64.04% data variance, and with the other PCs until PC5 accumulated 90.31% of the variance. Along with this, based on the Euclidian's distance of 295520186 (Figure 2), it is possible to see five

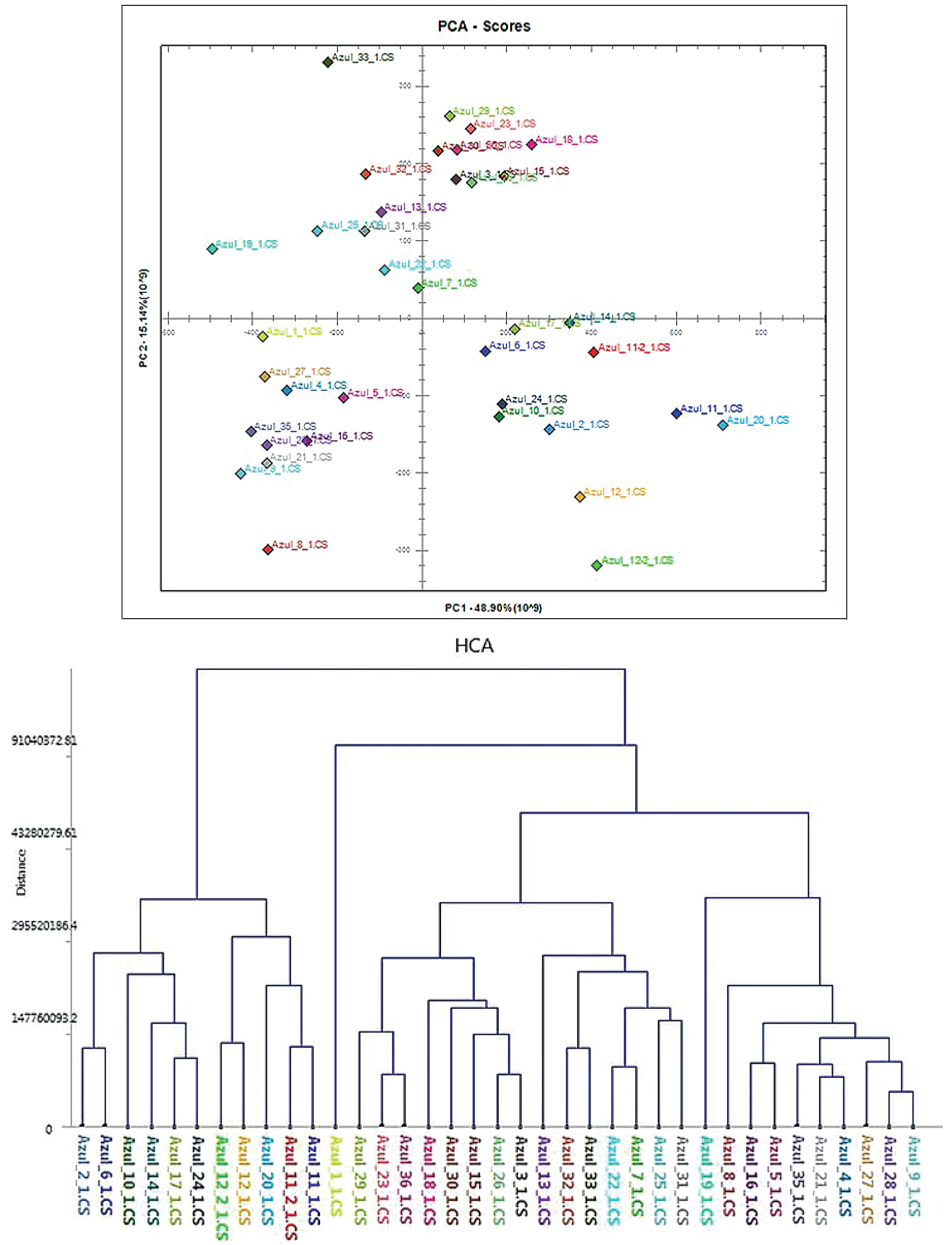

Figure 2. PC1 versus PC2 and HCA results for blue pens recently applied.

great hierarchical clusters and two differentiated samples (A19—Unilaknok and A1—Bic Ecolutions) in the HCA.

The HCA showed that the majority of the clusters were formed by pens from same brand and/or similar pictorial aspect of the ink (i.e. color, texture), so the method was able to differentiate the individuals but at the same time cluster the pens with great similarity (same brand/similar model).

For the black pens recently applied, the PC1 versus PC2 plot accounted for 71.41% data variance, and with the other PCs until PC5 accumulated 91.6% of the variability, a result very similar to the blue pens. Along with this, based on the Euclidian's distance of 46500158 (Figure 3), it is possible to see four great hierarchical clusters and two differentiated samples (P4—BIC Cristal and P13—Ita) in the HCA.

Similarly to the blue pens, the black pens HCA showed that the majority of the clusters were formed by pens from same brand and/or similar pictorial aspect of the ink (i.e. color, texture) and at the same was able to differentiate the individuals.

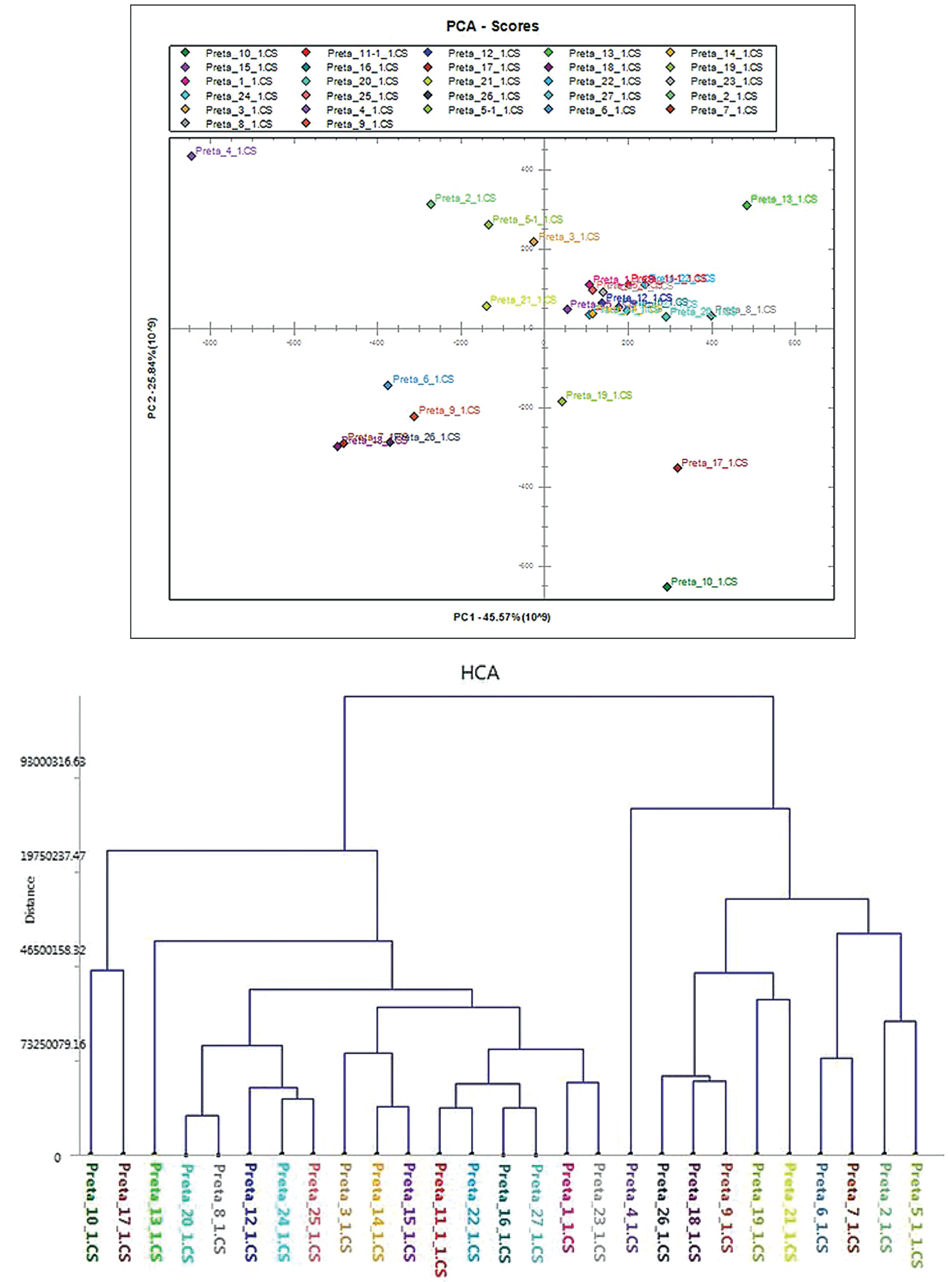

Figure 3. PC1 versus PC2 and HCA results for black pens recently applied. *Preta—means "black pen"

Looking at the results of the HCA obtained with the spectra of the same blue and black pens inks collected directly from the pen cartridge (figure 4), and analyzed in the FTIR- ATR Nicolet iS50 (Thermo), based on the Euclidian's distance of ~ 6.5 (Figure 4), the HCA showed that the grouping of pens remained the same, at least on the majority of the clusters formed. It can be observed, for example, that the pens A26 and A1 (both BIC brand), A32 and A33 (both Uni brand), P12, P24 and P25 (both Uni brand) clustered together in both methods.

The pens that had two individuals of different batches tested (A11, A12) clustered together despite the manufacture batch, for both methods. Other examples of pens that clustered together for both approaches were the Pilot pens P8 and P20; P11, P14, P15, P22 and P27; A36, A30 and A29.

These results confirm that the FTIR-ATR can differentiate or group the pens the same way, even if they are collected from the bulk ink spectra or the pen line in a paper. Senior *,et al.* performed PCA analysis of spectra from 10 blue ballpoint

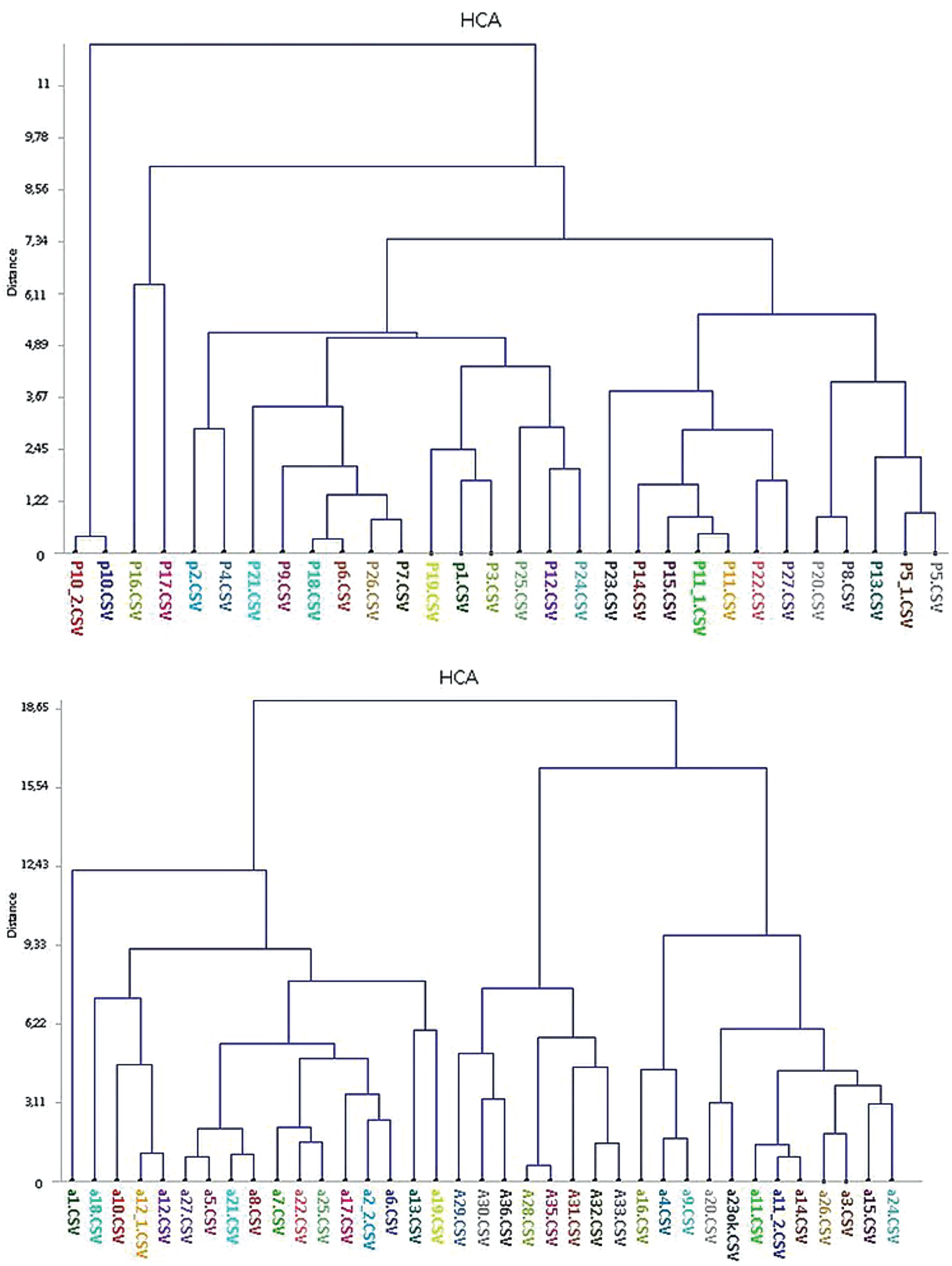

Figure 4. HCA results for black pens (P) (above) and blue pens (A)(below) sampled directly from the cartridge.

pen from different factories, and have encountered difficulties to discriminate all the pens[13]. Kher ,*et al.* achieved the complete separation of 17 pen pairs and 11 pen pairs partially resolved, using PCA analysis, and using Linear Discriminant Analysis approach, obtained a correct classification of 62.5% of the pen samples[2]. Silva ,*et al.* achieved 100% of correct classification for the test set, using LDA approach[16]. However, to perform LDA, the classes of samples have to be first determined, using a training set of samples.

From the hierarchical clusters formed in the HCA of freshly applied pen strokes, representative pens were chosen from each cluster, to perform the multivariate analysis to attempt to discriminate different ages, in a way to develop a suitable method to date a ballpoint pen ink in real forensic cases. The pens chosen were A26, A1, A12, A17, A21, A31, A19 and A32 for blue pens, and P1, P10, P6, P16, P15, P9, P21 and P25 for black pens. For blue ink pens, the PCA was able to resolve the samples from each age, with the first three PCs accounting for more than 95% of the variance.

Looking carefully at the PC loadings, it could be seen that the more recently samples were located in the 2D PC space where the absorption bands typical from solvents like 2-phenoxyethanol and benzyl alcohol are influencing in these loadings. The peaks at 695 cm^{-1}, 750 cm^{-1} (mono-substituted aromatic bending vibration), 1235 to 1294 cm^{-1} strong band (aromatic ethers Ar-O), 1456 to 1497 cm^{-1} (C=C aromatic vibration), 1025 to 1048 cm^{-1} (primary alcohol C-OH) and 1587 cm^{-1}, are characteristic of 2- phenoxyethanol and related solvents[17, 26]. In the PCA score plot of the blue pen samples that included samples "freshly applied", the 2D PCA score plot showed more the difference between fresh samples and the others, probably because the fresh ink have higher concentration of solvents. Removing the fresh samples from the set and repeating the PCA analysis, the 2D PCA score plot showed results that permitted to discriminate the samples "14 days", "7 months" and sometimes "1 year" from the other samples older (2 years; 2 years and 3 months). Figure 5 illustrates the results with "freshly applied" sample and removing "freshly applied" sample, for the pen A32 (UniSAS), and it can be seen that for A32pen, the one year sample did not differ from older ones.

Some brands of pens were differentiated well such as the "1 year" sample from the older "2 years" and "2 years and 3 months", while for other brands, it was not possible to differentiate the "1 year" sample from the oldest ones, and in the loadings responsible for grouping these older and "1 year" samples, the 2-phenoxyethanol absorption bands were not observed. This result was comparable with the results observed for samples analyzed in a recent study of our research group performed with the same samples (data not published yet). In that study, it was determined the 2-phenoxyethanol concentration in blue and black ballpoints aged for 4 months, 18 months, 27 months and 36 months[28], and some samples aged 18 months (1 year and a half) did not present a 2-phenoxyethanol concentration statistically different from the samples with older ages, in the same way that occurred in the IV-ATR study. Table 4 presents a summary of the results for blue pens.

Performing the HCA for some blue pens recently applied (A1, A12, A17, A21, A23, A26, A31 and A32) and comparing it with the HCA performed for the 14 days, 7 months and 2 years old pens, some changes in the clustering of the pens can be seen (Figure 6a, 6b and 6c). While fresh pens clustered together in two groups, after 14 days, the ageing the pens A12, A1 and A23 were affected more.. After 7 months, the clusters and the Euclidian distance between the pens increased, though some pens that were 14 days apart regrouped with 7 months.

After two years, the A1 resulted in more dissimilarity to the others, and two great clusters were formed: cluster 1 (A26, A12 and A23 plus A1) and cluster 2 (A31, A21, A32, A17, A19). After two years, the pens became more similar, especially the ones made in Brazil. The inconstancy in HCA results through different ages could be due not just to the solvent evaporation, but to the ageing behavior of other ink components, like polymerization of resins and decomposition of triarylmethane dyes. Wang ,*et al.* observed a decrease in the absorption bands 1584, 1494, 1360, 1296, 1245, 1094 and 1055 cm^{-1} attributed to the changes in the epoxy resin and volatilization of solvents; the signal 1245 cm^{-1} almost disappeared after artificial ageing at 150 ºC for 24 min[22].

In relation to the black pens, the IV-ATR spectra were collected from ink strokes recently applied, with 7 months, with 2 years and with 2 years and 6 months. The PCA was able to resolve the samples from each age, with the first two or three PCs describing more than 98% of the variance. Illustration of the findings are shown in Figure 7a and 7b, with PCA and loadings for BRW pen (P10).

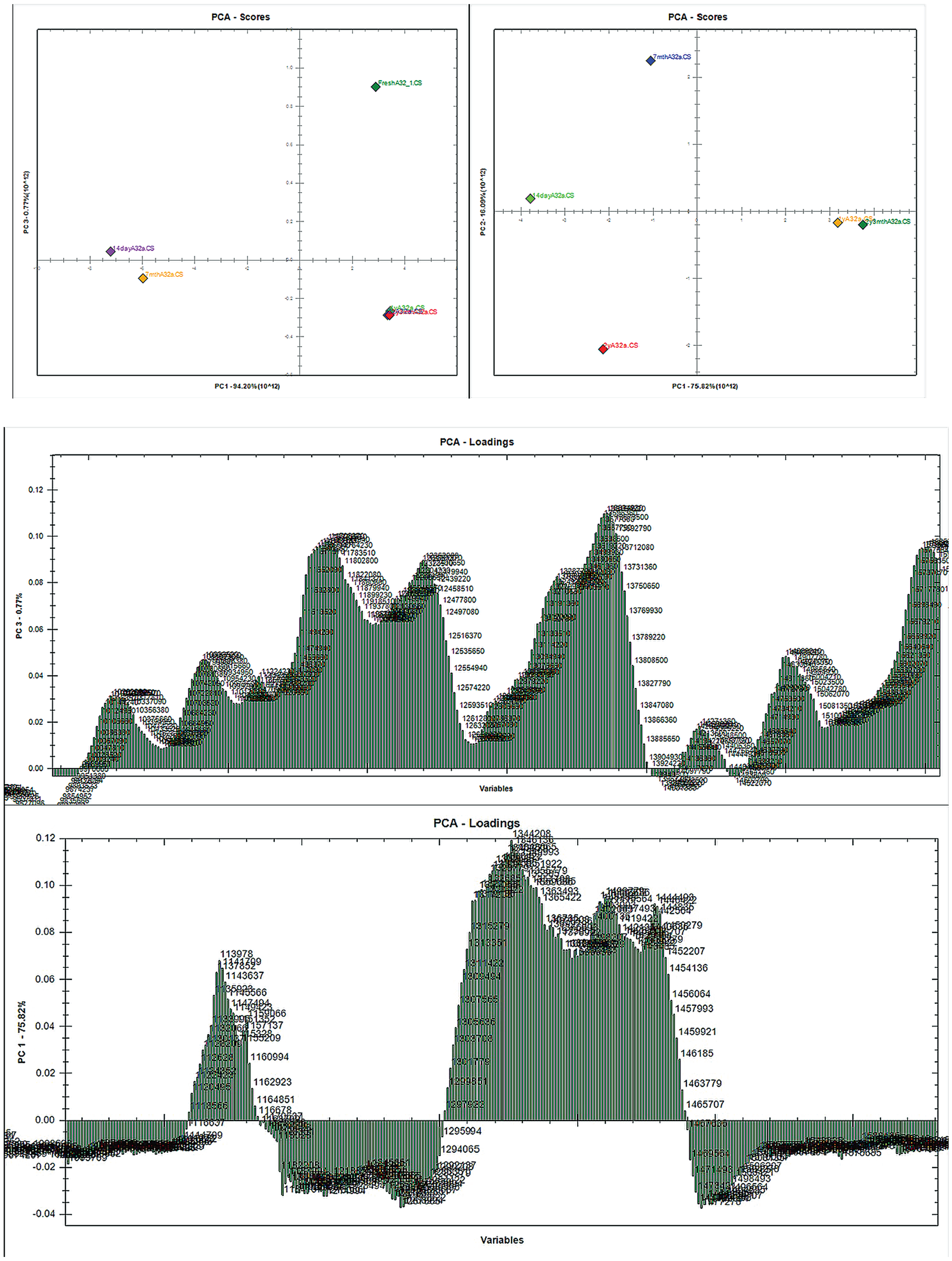

Figure 5. PCA results for A32 pen: left: PC1xPC3 with fresh sample; right: PC1xPC2 without fresh sample. Loadings for A32 pen: first figure, PC3 loadings with the fresh sample. Second figure, PC1 loadings without fresh sample.

Table 5 also presents a resume of the results for black pens. These were very consistent and similar for all samples, in which PC1xPC2 resolved fresh ink in PC1 positive, and the older samples in PC1 negative. The absorption peaks from solvents like 2-phenoxyethanol were more pronounced in PC1 positive and PC2 negative. The PC2xPC3 scores resolved samples with different ages, each one in a different PC quadrant, and PC3 did not influenced fresh samples. However, the solvent absorption peaks could not be attributed as loadings which accounted for the difference of

Table 4. Blue pens Analyzed by age.

Pen code	PCA Results With fresh samples	PCA Results Without fresh samples
A1	PC1 x PC2 Only resolved Fresh sample	PC1xPC2 resolved recent ink in PC1 negative Loadings of solvents in PC1 and PC3 negative
A12	PC2 x PC3 resolved fresh ink in PC3 positive Loadings of solvents in PC3 positive	PC1 x PC2 resolved earlier from older samples Recent ink in PC1 negative Loadings of solvents in PC1 negative
A17	PC2 x PC3 resolved Fresh ink in PC2 negative Loadings of solvents in PC2 positive	PC2 x PC3 resolved Recent ink in PC3 negative Loadings of solvents in PC3 negative One year sample differed from older ones
A19	PC1xPC2 resolved fresh ink in PC2 negative Loadings of solvents in PC2 negative	PC1 x PC2 resolved earlier from older samples Recent ink in PC1 negative Loadings of solvents in PC1 negative
A21	PC1 x PC2 resolved Fresh ink in PC1 positive Loadings of solvents in PC1 positive	PC1 x PC2 resolved all samples each one in a PC square Loadings of solvents in PC1 negative
A26	PC2xPC3 resolved well fresh ink from older ink Loadings of solvents in PC3 positive and PC2 negative	PC1xPC4 resolved 14days sample from de others Loadings of solvents in PC1 negative and PC4 positive
A31	PC1xPC2 resolved fresh and 14 days ink in PC2 negative Loadings of solvents in PC2 negative	PC1xPC4 resolved 14days sample from de others Loadings of solvents in PC1 negative and PC4 positive 1year sample did not differ from older samples
A32	PC1xPC3 resolved fresh ink in PC1 and PC3 positive and PC2 x PC3 resolved Fresh an 14 days ink Loadings in PC3 positive, 1year sample did not differ from older samples	PC1 x PC2 resolved earlier from older samples Recent ink in PC1 negative and PC2 positive 1year sample did not differ from older samples

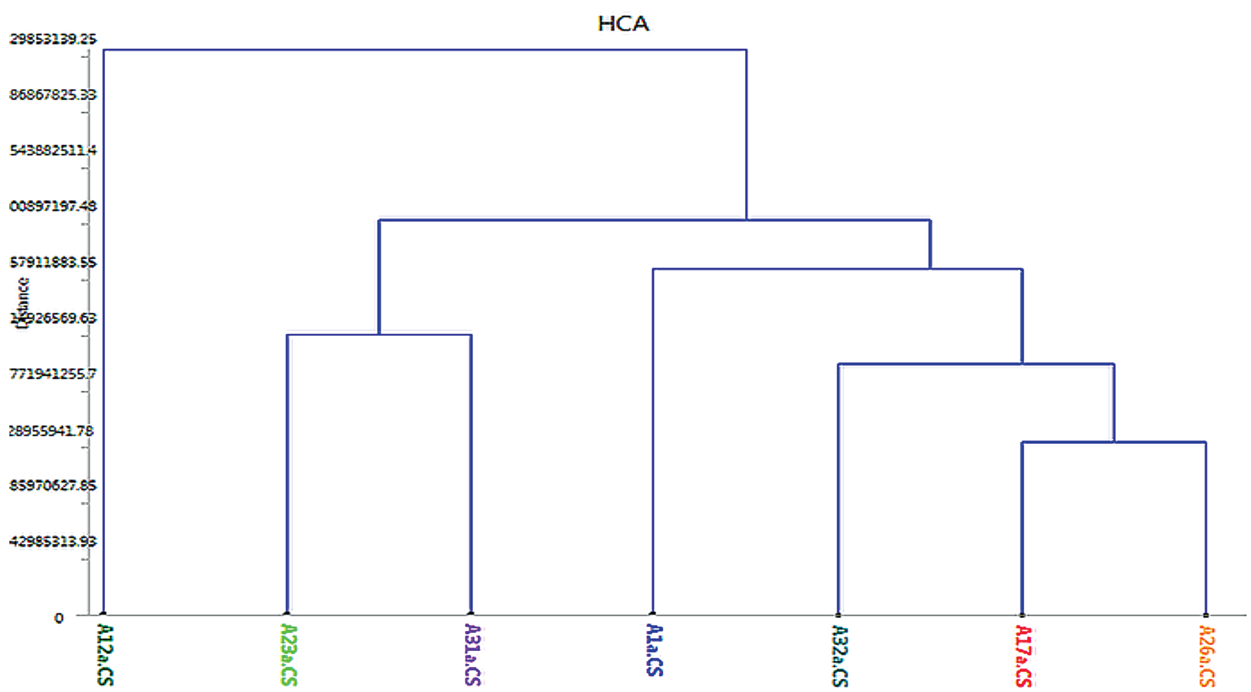

Figure 6a. HCA results for blue pens after 14 days.

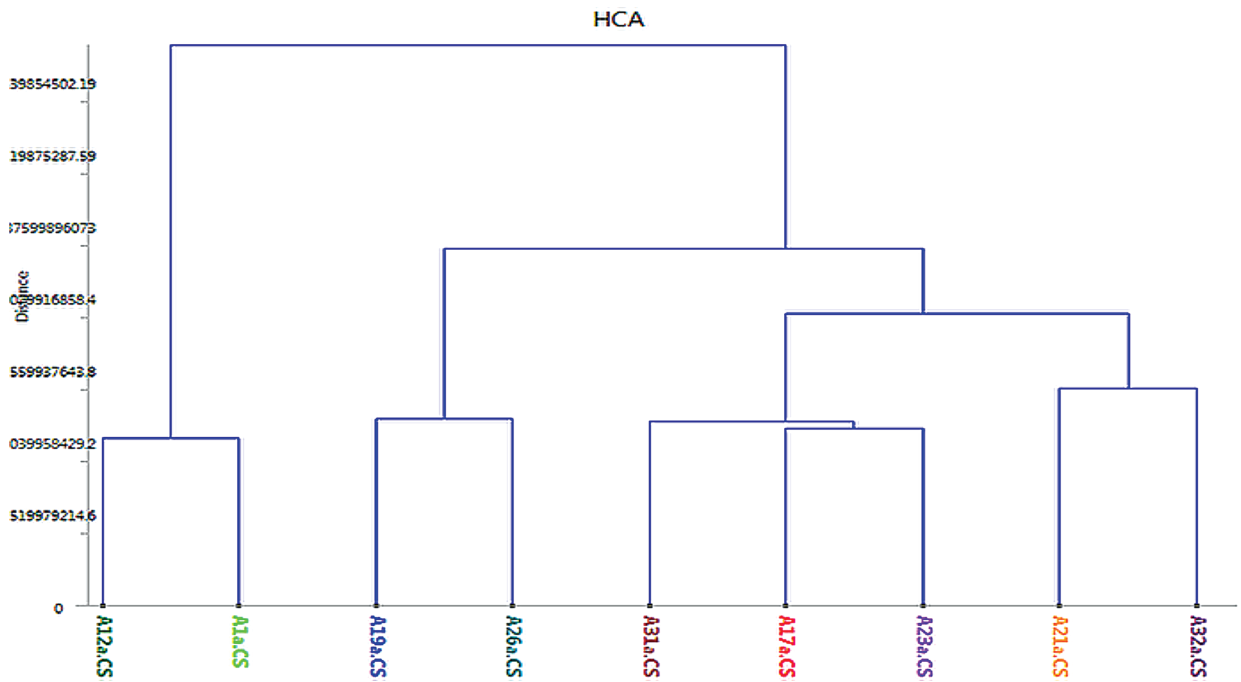

Figure 6b. HCA results for blue pens after 07 months.

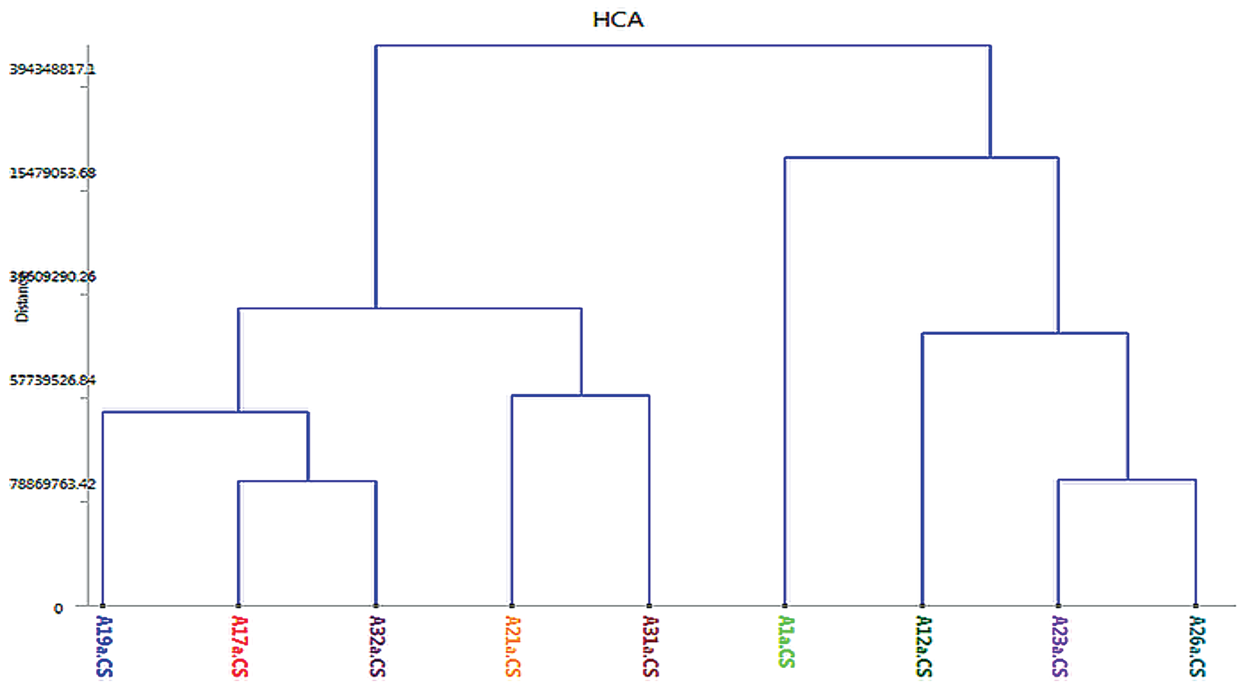

Figure 6c. HCA results for blue pens after 2 years.

the samples in PCA. Nevertheless, these results should be considered because other pen ink components such as resins and dyes that changes with time, could be the loadings that differentiate the samples in PCA.

Table 5. Black pens Analyzed by age

	PCA Results **With fresh samples**	**Other relevant PCs**
P1	PC1xPC2 resolved fresh ink in PC1 positive Loadings of solvents in PC1 positive	Without fresh samples: PC1xPC2 resolved recent ink in PC1 and PC2 positive Loadings of solvents in PC1 and PC2 positive
P9	PC1xPC2 resolved fresh ink in PC1 positive Loadings of solvents in PC1 positive	PC2 x PC3 resolved all samples each one in a PC square. These PCs did not influence in fresh sample
P10	PC1xPC2 resolved fresh ink in PC1 positive Loadings of solvents in PC1 positive	PC2 x PC3 resolved all samples each one in a PC square. PC3 did not influence in fresh and 7 months sample
P15	PC1 x PC2 resolved Fresh ink in PC1 positive Loadings of solvents in PC1 positive	PC2 x PC3 resolved all samples each one in a PC square. These PCs did not influence in fresh sample
P16	PC1 x PC2 resolved Fresh ink in PC1 positive Loadings of solvents in PC1 positive	PC2 x PC3 resolved all samples each one in a PC square. These PCs did not influence in fresh sample
P21	PC1 x PC2 resolved Fresh ink in PC1 positive Loadings of solvents in PC1 positive and PC2 negative	PC2 x PC3 resolved all samples each one in a PC square. PC3 did not influence in fresh sample Loadings of solvents in PC2 negative
P25	PC1 x PC2 resolved Fresh ink in PC1 positive Loadings of solvents in PC1 positive and PC2 negative	PC3 did not influence in fresh and 7 months samples Loadings of solvents in PC2 negative
P6	PC1xPC2 resolved fresh ink in PC1 positive Loadings of solvents in PC1 positive	NA

The only black pen that provided a good result removing the samples recently applied from the PCA was Bic Cristal (A1). For PC1 versus PC2 (Figure 8a) the 7 months ink in the positive quadrant of the PCs could be resolved, and the loadings from solvents could be seen in PC1 and PC2 positive (Figure 8b).

In the same way as was seen for the blue pens, performing the HCA for a few black pen ink strokes recently applied (P1, P9, P10, P15, P16, P20, P21, and P25) and comparing it with the HCA performed for the 7 months and 2 years old pens, some changes in the clustering of the pens can be seen. After 7 months, P10 and P25 differed more than the other pens. After two years, the ageing of the pens resulted in a HCA with two clusters, in which the pens made in Brazil tended to cluster together.

Conclusions

The analysis of ballpoint pen samples through IR-ATR was able to differentiate or group the samples, even for inks of very similar color, and even if the spectra was collected from the bulk ink or from the ink line in a paper. The spectra were submitted to Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA). This approach could confirm that the paper where the pen line is written did not influence the differentiation and grouping of the pens.

The PCA revealed the variables that were responsible for the variation and/or similarity of the samples. Performing PCA for ballpoint pen line IR-ATR spectra collected from samples with different ages, this approach was able to differentiate pen lines freshly applied, with 14 days and 7

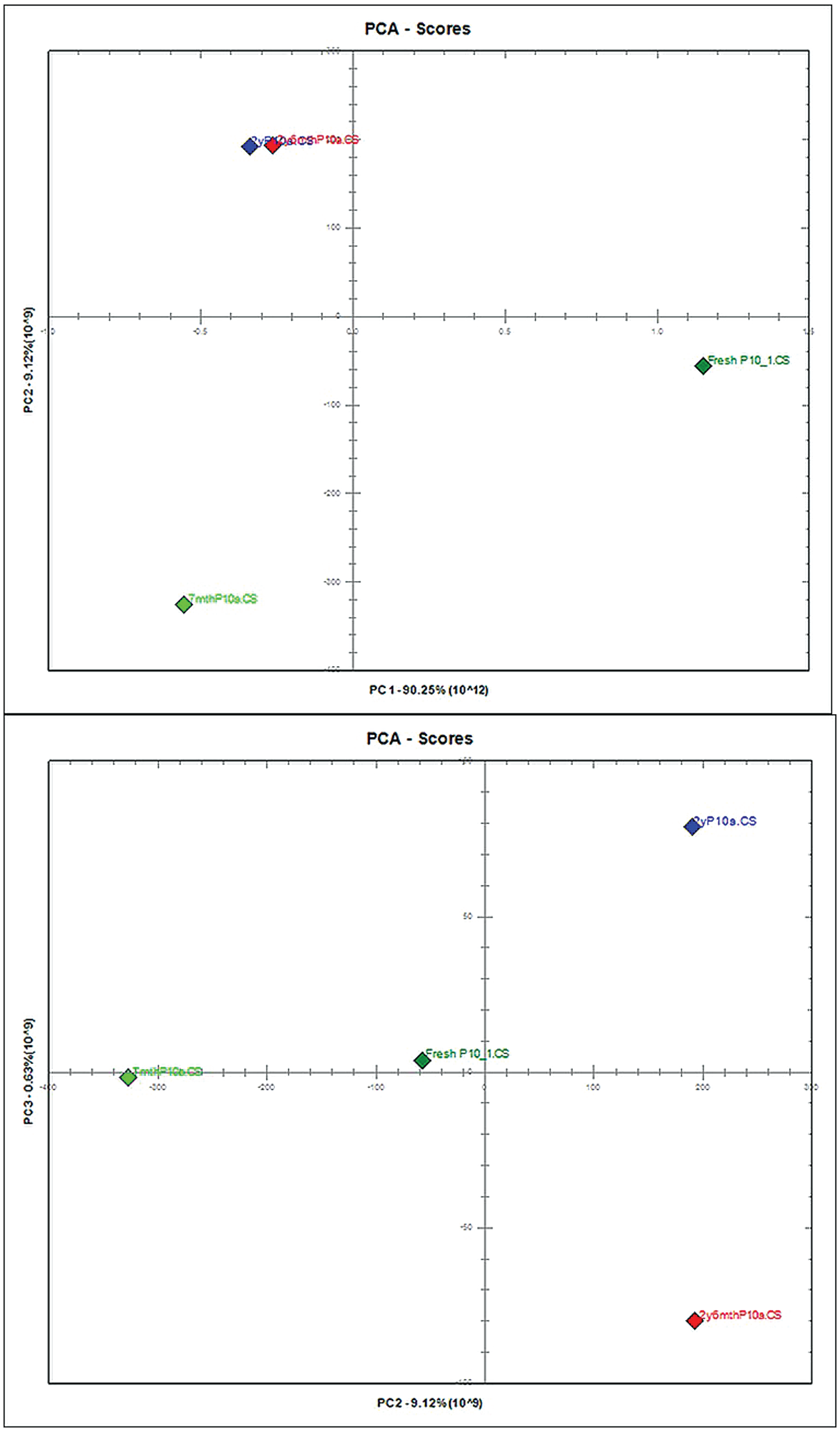

Figure 7a. PCA results for black pen BRW (P10): over: PC1xPC2; down: PC2xPC3.

months from pen lines older than 7 months. The PCA loadings showed that the samples recently applied until 7 months had in common the absorption peaks from solvents like 2-phenoxyethanol, which are the formulation components that evaporate through time.

The HCA results for blue and black pens, showed different clusters of brands of pens, for

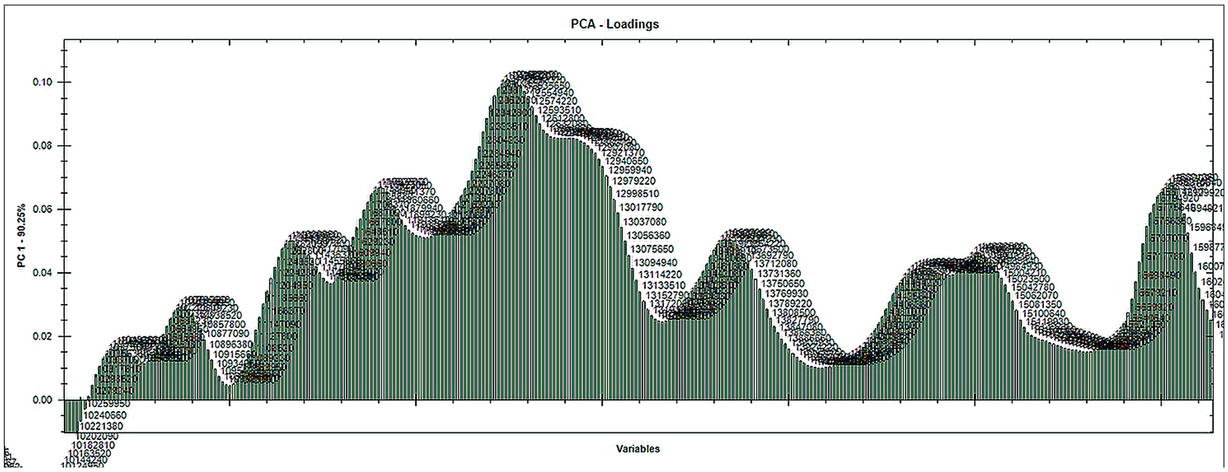

Figure 7b. Loadings of PC1 for black pen BRW (P10).

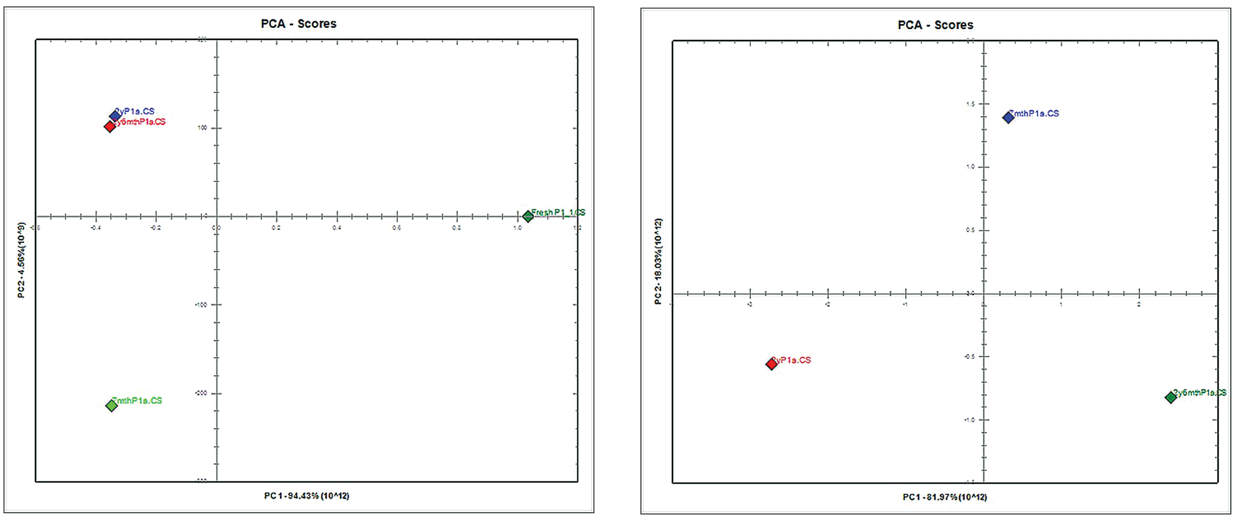

Figure 8a. PC1 x PC2 for Bic Crystal (P1): left: with fresh sample; right: removing fresh sample.

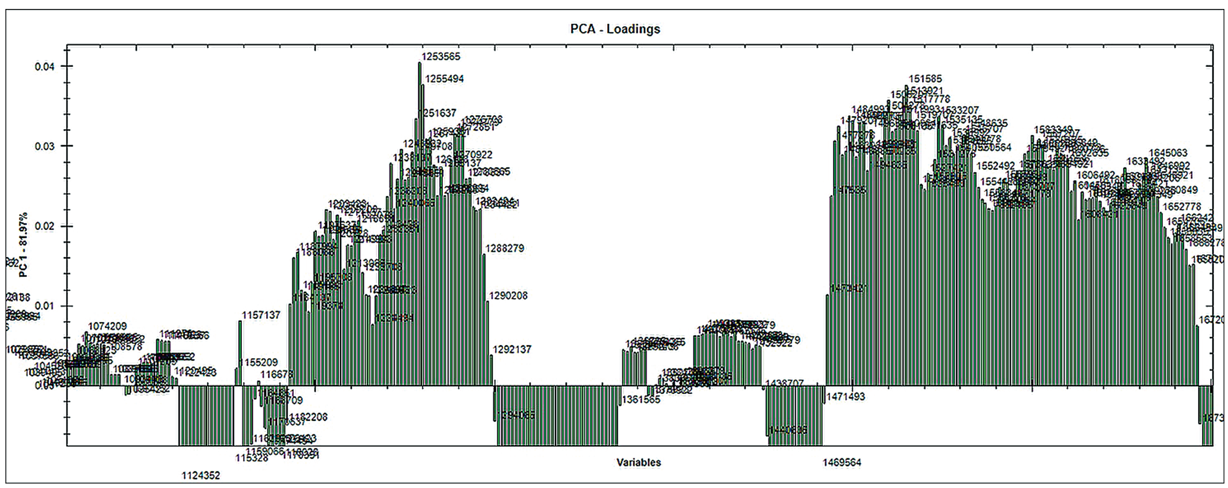

Figure 8b. PC1 Loadings for Bic Crystal pen, removing the freshly applied sample.

spectra collected from different ages. Ballpoint pens made in Brazil, which are different from each other when recently applied, tended to group together after two years, becoming more similar.

Some blue pen lines with one year of age did not differ from older pens in the PCA. At the same time, other one year and seven months pen lines were divergent from older ones, but the loadings did not show absorption peaks of the solvents. This result could be explained by other pen ink components as resins and dyes, that changes with time, and could be the loadings that differentiate the samples in PCA.

Additional research should be performed with real cases of pen manuscripts naturally aged, and with experimental samples followed from more than two years, to confirm this behavior and the applicability of this IR-ATR method to ballpoint pen dating. Other models of Multivariate Analysis like Linear Discriminant Analysis could be applied forward to classify pen inks by date.

Acknowledgements

The authors wish to thank the Scientific Section of Brazilian Federal Police—Rio Grande do Sul Bureau and the Brazilian National Institute of Forensic Science and Technology.

References

1. Andrasko, J. HPLC analysis of ballpoint pen inks stored at different light conditions. *J Forensic Sci,* 46(1), (2001) 21-30.

2. Kher, A.; Green E.V.; Mulholland M.I. Forensic classification of ballpoint pen inks using high performance liquid chromatography and infrared spectroscopy with principal components analysis and linear discriminant analysis *Vibrational Spectroscopy* 40, (2006) 270-277.

3. Denman J.A.; Skinner, W.M.; Kirkbride, K.P.; Kempson, I.M. Organic and inorganic discrimination of ballpoint pen inks by ToF-SIMS and multivariate statistics *Applied Surface Science,* 256, (2010) 2155-2163.

4. Andrasko, J. Changes in composition of ballpoint pen inks on aging in darkness *J Forensic Sci,* 47(2), (2001) 324-327

5. Liu, Y.; Yu, J., Xie, M.X.; Liu, Y.; Jing, T.T. Classification and dating of black gel pen ink by ion-pairing high-performance liquid chromatography *J. Chromatogr. A.* 1135 (2006) 57-64.

6. Liu, Y., Yu, J., Xie, M.X.; Chen, Y.; Jiang, G.Y.; Gao, Y. Studies on the degradation of blue gel pen dyes by ion-pairing high-performance liquid chromatography and electrospray tandem mass spectrometry *J. Chromatogr. A.* 1125(1) (2006) 95-103.

7. Wang J., Luo G., Sun S, Wang Z, Wang Y. Systematic analysis of bulk blue ballpoint pen ink by FTIR spectrometry *J Forensic Sci* 46(5) (2001) 1093–1097.

8. Lee, L.C., Shandu, K. T. S., Razi, N. S. M., Ishak, A. A., Osman, K. Forensic analysis of blue ballpoint pen inks using ultraviolet-visible spectrometer and ultra-performance liquid chromatograph *Malaysian Journal of Analytical Sciences*, 19(2) (2015) 397 - 401

9. Confortin, D. Brustolon, M.; Franco, L.; Neevel, H.; Bommel, M.R. Crystal violet: study of the photo-fading of an early synthetic dye in aqueous solution and on paper with HPLC-PDA, LCMS and FORS. *Journal of Physics: Conference Series* 231 (2010) 012011.

10. Freidenfelds, V., Mekss, P. The possibilities of determination of the relative age of ballpoint pen ink entries by high performance liquid chromatography *Latvian Journal of Chemistry,* 3 (2012) 242–248.

11. Bello de Carvalho, C. M., Ortiz, R. S., Dos Reis, M., Zamboni, A.; Limberger, R.P., Ferrão, M.F., Vaz, B. G. Characterization and Differentiation of Ballpoint Pen Ink Strokes on Paper Using Orbitrap Mass Spectrometry and Multivariate Statistic *Forensic Sci Add Res* 2(2) (2018) FSAR.000537

12. 12 Lee, J., Nam, YS., Min, J., Lee, K.B., Lee, Y. TOF-SIMS Analysis of Red Color Inks of Writing and Printing Tools on Questioned Documents. *Forensic Sci* 61(3) (2016) DOI 10.1111/1556-4029.13047.

13. Senior, S., Hamed, E., Masoud, M., Shehata, E. Characterization and dating of blue ballpoint pen inks using principal component analysis of UV-Vis absorption spectra, IR spectroscopy, and HPTLC *J Forensic Sci* 57(4) (2012) 1087–1093.

14. Harris, J. A. Preliminary Report on the Nondestructive Examination of Ballpoint Pen Ink on Questioned Documents by FT-IR Spectroscopy *Canadian Society of Forensic Science Journal*, 24 (1) (2016)5-21.

15. Causin V., Casamassima R., Marega C., Maida P., Schiavone S., Marigo A., Villari A. The discrimination potential of ultraviolet-visible spectrophotometry,

thin-layer chromatography, and Fourier transform infrared spectroscopy for the forensic analysis of black and blue ballpoint inks. *J Forensic Sci* 53(6) (2008) 1468–1473.

16. Silva, C.S., Borba, F.S.L., Pimentel, M.F., Pontes, M.J.C., Honorato, R.S., Pasquini, C. Classification of blue pen ink using infrared spectroscopy and linear discriminant analysis *Microchemical Journal,* 109 , (2013) 122-127.

17. Sharma, V., Kumar, R. Fourier transform infrared spectroscopy and high-performance thin layer chromatography for characterization and multivariate discrimination of blue ballpoint pen ink for forensic applications *Vibrational Spectroscopy* 92 (2017) 96-104.

18. Nam, Y.S., Park, J.S., Lee, Y., Lee., K. Application of Micro-Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy to Ink Examination in Signatures Written with Ballpoint Pen on Questioned Documents. *J Forensic Sci,* 59 (3) (2014) 800-805.

19. Zieba-Palus, J., Kunicki, M. Application of the micro-FTIR spectroscopy, Raman spectroscopy and XRF method examination of inks *Forensic Sci Int*, 158, (2006), 164-172.

20. 20 Payne, G., Wallace, C., Reedy, B., Lennard, C., Schuler, R., Exline, D., Roux, C. Visible and near-infrared chemical imaging methods for the analysis of selected forensic samples *Talanta*, 67, (2005), 334-344.

21. Silva, C.S., Pimentel, M.F., Honorato, R.S., Pasquini, C., Prats-Montalbán, J.M., Ferrer, A. Near infrared hyperspectral imaging for forensic analysis of document forgery *Analyst,* 139 (2014) 5176-5184.

22. Wang, J., Luo, G., Sun, S., Wang, Z., Wang, Y Systematic analysis of bulk blue ballpoint pen ink by FTIR spectrometry. *J Forensic Sci,* 46 (5) (2001) 1093-1097.

23. Thanasoulias, N.C., Parisis, N.A., Evmiridis, N. P. Multivariate chemometrics for the forensic discrimination of blue ball-point pen inks based on their Vis spectra. *Forensic Sci Int*, 138, (2003), 75- 84.

24. Adam, C. D., Sherrat, S. L., Zholobenko, V. L. Classification and individualisation of black ballpoint pen inks using principal component analysis of UV–vis absorption spectra *Forensic Sci Int*, 174, (2008), 16-25.

25. Chen, H.S., Meng H.H., Cheng, K.C., A survey of methods used for the identification and characterization of inks. *Forensic Science Journal* 1, (2012) 1-14.

26. Feraru, D.L., Meghea, A., Badea, N. Forensic discrimination of ballpoint pen inks based on correlation of data obtained by optical and spectral methods. *Rev. Chim.* 64, (1), (2013) 74-80.

27. Helfer, G.A., Bock, F., Marder, L., Furtado, J. C., Costa, A.B., Ferrão, M.F. Chemostat: exploratory multivariate data analysis software. *Quimica Nova,* 38 (4), (2015) 575-579.

28. Carvalho, C.M.B., Ortiz, R. S., Limberger, R. P. Manuscripts from Different Brands of Ballpoint Pens – GC/MS Ageing Profile in Brazilian Tropical Weather. *Curr. Chrom.,* 6, (2019) 1-15 (*in press*).