Fast profiling of cocaine seizures by FTIR spectroscopy and GC-MS analysis of minor alkaloids and residual solvents

Maria Monfreda *, Francesco Varani, Fabrizio Cattaruzza, Simona Ciambrone, Alessandro Proposto

Central Directorate for Product Analysis and Chemical Laboratories, Italian Customs and Monopolies Agency, via M. Carucci 71, 00143 Rome, Italy

1. Introduction

Minor components can provide very useful information for cocaine profiling. They are alkaloids found in coca leaves at small amounts, depending on taxonomic factors; nevertheless, their traces are found in trafficked cocaine samples, depending on the efficiency of extraction and purification processes. Moreover, alkaloids may undergo chemical modifications due to some reactions that occur during these processes, so new impurities can be recovered [1,2]. Analyses of alkaloids for drug profiling are mainly carried out by GC-MS [3]; analytical methods have been widely studied and optimized [4]. Recently, nuclear magnetic resonance (NMR) was also applied for comparative analysis of cocaine seizures [5].

In addition to the minor alkaloids, residual solvents that can be occluded during the crystallization of cocaine in a salt form (HCl) are often used for comparative analysis in forensic investigations. Such solvents are analyzed through head space (HS) combined to gas chromatography: (solid phase microextraction) SPME-GC-MS [6] and static HS combined to GC-FID [7]. Chemometric tools could allow samples to be grouped according to the location where cocaine base was converted to cocaine hydrochloride; furthermore, an application of residual solvents analysis to distinguish the geographical origin of cocaine was also investigated [7]. There is, indeed, considerable international interest in the identification of the geographic origin of illicit drugs, especially for Customs authorities. For example, when the country of origin is different from that of Customs seizure, a need to increase controls of certain types of importations comes to light.

Recent studies move towards the trend of “fast profiling”. Some examples are the use of a single quadrupole mass spectrometer and an isotope ratio mass spectrometry (IRMS) as simultaneous GC detectors [8], the application of desorption electrospray ionization-mass spectrometry (DESI-MS), [9] and the possibility of feeding a common conventional GC-MS database with data coming from ultra-high-pressure liquid chromatography (UHPLC) [10]. Even if GC-MS currently remains the technique of choice to achieve drug profiling, mainly because it is generally present in most of analytical laboratories, UHPLC–MS/MS allows to avoid the derivatization step, could offer better chromatographic performance, and moves forward to the trend of “fast profiling.”

To the best of our knowledge, Fourier transform infrared (FTIR) spectroscopy has not been used in the field of drug profiling. However, this technique has been applied for testing some “fast methods”: an example is the characterization and discrimination between defective and non-defective coffee beans prior to roasting [11]. Some advantages of FTIR spectroscopy in drug profiling would be the shorter time of analysis compared to that required for the analysis of alkaloids and residual solvents by GC-MS, allowing a rapid assessment of possible heterogeneities in a big seizure due to different origins of the respective seizure portions. On the other hand, infrared spectra would be influenced by adulterants or diluents if contained in samples. For this reason, while a
drug profiling model based on FTIR spectroscopy should be developed analyzing the so-called "pure" samples (free from adulterants or diluents), the power of such a model for comparative analysis of seizures should be tested against the purity level of samples.

In this study, samples coming from large seizures of cocaine which took place in Italian Customs areas during 2011 and 2012 were analyzed. Minor alkaloids and residual solvents were analyzed and principal component analysis (PCA) was applied in order to highlight a natural grouping of samples according to their chemical similarity.

The aim of this study was to verify whether a “fast profiling” method based on infrared spectroscopy could be provided; it could be a useful tool to speed controls carried out by Customs authorities and to move them towards the sites where the seizures take place. First of all, only pure samples were analyzed; analytical data were processed by ANOVA and PCA in order to verify if results given by analyses of alkaloids and residual solvents could also be achieved by infrared spectroscopy. Performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose. The performances of KBr-FTIR and ATR-FTIR techniques were compared for this purpose.

2. Materials and methods

2.1. Samples

29 samples of cocaine, extracted from large seizures that occurred in Italian Customs areas during 2011 (14 samples) and 2012 (15 samples), were analyzed. The only available information concerning the boarding countries (Dominican Republic, Brazil, Argentina, Mexico, Colombia, Guatemala, and Benin) or, in some cases, the location where the seizures occurred (i.e. Gioia Tauro harbor). The total weight of seizures from which samples were extracted was 1015.90 kg; 988.42 kg of them were seized in 2012. According to Italian official data, the total amount of cocaine seized in Italy during 2012 is equal to 5323.84 kg: 3471.41 kg of them were seized in Italian Customs areas [12]. Therefore, the seizures of 2012 from which the samples analyzed in this work were extracted represented 18.6% of total cocaine seized in Italy during 2012 and 28.5% of cocaine seized in Customs areas during the same year.

Taking into account that the material under investigation could not be analyzed in the whole amount, it was necessary to carry out measurements on a smaller portion taken from the bulk. When sampling, the aim is to obtain samples, throughout the whole mass-reduction process, that are representative of the bulk material. Samples were taken from consistent populations; seizures related to non-consistent populations were divided into homogeneous subgroups. 10 increments of 1 g were taken from random positions in the material, combined and homogenized. From this primary sample, a 2 g amount was extracted and subjected again to homogenization (agate mortar and pestle were used) before each analysis was performed.

2.2. Alkaloids analysis (GC-MS)

Alkaloids were analyzed by GC-MS using an already published method [3]. Two portions for each sample were analyzed. GC was performed on a 30 m · 0.25 mm i.d. 0.25 μm film ZB-5 capillary column using an Agilent 6890 gas chromatograph connected to an Agilent 5979 mass selective detector.

2.3. Determination of residual solvents by HS-GC-MS

Residual solvents were analyzed by HS-GC-MS, making some modifications to an already published HS-GC-FID method [7]. Only relevant modifications are described in this section.

200 mg of sample was dissolved in 3 mL of bi-distilled ultra pure water saturated with 22% sodium sulphate. Two portions for each sample were analyzed.

A headspace injector Agilent 7694 E and an autosampler were used. GC was performed on a 30 m · 0.25 mm i.d. 0.25 μm film ZB-5 capillary column using an Agilent 6890 gas chromatograph connected to an Agilent 5979 mass selective detector.

The GC conditions were temperature set to 40 °C for 8 min, then ramped to 100 °C at 10 °C min⁻¹ and set to rise 200 °C at 10 °C min⁻¹ (held for 2 min), giving a total run time of 19.33 min. Helium gas carrier was held at a constant flow rate of 1 mL min⁻¹. Temperatures of transfer line, ion source, and quadrupole were set to 280 °C, 230 °C at 150 °C, respectively (acquisition in full scan mode; m/z 30–450 amu; sampling rate 2 scans/sec).

2.4. Quantification of cocaine by GC-FID

Quantification of cocaine was carried out with an internal method. For each homogenized sample, a test portion ranging from 10 to 150 mg was weighted and dissolved in a volumetric flask of 10 mL, with about 8 mL of a mixture of chloroform/methanol (30/70 v/v). The cocaine extraction from the sample was improved by putting the volumetric flask in an ultrasonic bath for 5 minutes; the solution was then thermostated at 20 °C and filled with chloroform/methanol to its final volume of 10 mL. 400 μL of such a solution (already homogenized and filtered) was put in a GC vial and 400 μL of internal standard solution (triacetone in chloroform at a concentration of 0.7 mg/mL) were added.

GC was performed on a 30 m · 0.25 mm i.d. 0.25 μm film ZB-5 capillary column using an Agilent 6890 gas chromatograph connected to a flame ionization detector (FID). The GC conditions used were as follows: injection volume 1 μL, split injection 40:0:1 at 280 °C; temperature set to 140 °C for 3 min, then ramped to 180 °C at 20 °C min⁻¹ (held for 1 min) and set to rise 300 °C at 10 °C min⁻¹ (held for 4 min) giving a total run time of 22 min. Helium gas carrier was held at a constant flow rate of 0.8 mL min⁻¹, while the detector was set at a temperature of 300 °C.

Cocaine quantification was done by a calibration curve method: a standard solution of 1 mg/mL expressed as cocaine base, was diluted with the mixture of chloroform/methanol (30/70 v/v), and internal standard (triacetone in chloroform at a concentration of 0.7 mg/mL) was added. Solutions at five concentration levels—0.05, 0.125, 0.25, 0.40, and 0.50 mg/mL—were therefore analyzed.

Taking into account the method repeatability and reproducibility and that each sample was analyzed in duplicate, uncertainties of measurement of 5.7% and 5.0% (expressed as percentage of cocaine in a sample) were estimated, respectively, for samples containing more than 70% or less than 70% of cocaine.

2.5. FTIR spectra

7 samples (labeled with nos. 2, 9, 20, 22, 23, 25, and 27) were found free from adulterants or diluents. 10 independent test portions for each of them were analyzed by KBr-FTIR and ATR-FTIR spectroscopy, respectively. “Cut” samples (3 independent test portions for each of them) were analyzed only by ATR-FTIR spectroscopy.

KBr-FTIR spectra were recorded with a Nicolet 5700 instrument. A 2 mg aliquot of each powdered sample was mixed with 200 mg KBr in an agate mortar. Measurement range was 4000–400 cm⁻¹; 32 scans for sample were collected at a resolution of 2 cm⁻¹. The preparation of KBr pellets relied on a standardized procedure (i.e. constant sample/KBr weight ratio, grinding time, pressure, and pellet weight) to enhance reproducibility. Background was recorded each time before the measurement and subtracted automatically by the software.

ATR-FTIR spectra were recorded with a Nicolet 5700 instrument, equipped with a ZnSe crystal, between 4000 and 500 cm⁻¹, with 32
scans at a resolution of 2 cm$^{-1}$. Background was recorded each time before the measurement and subtracted automatically by the software.

2.6. Multivariate statistical analysis

2.6.1. Alkaloids

The following alkaloids were detected: ecgonine methyl ester, ecgonine, tropacocaine, cocaine, benzoylecgonine, norcocaine, cis-cinnamoylcocaine, trans-cinnamoylcocaine, and trimethoxycocaine.

Peak areas were normalized to the sum of cocaine and benzoylecgonine peak areas and multiplied by 10$^6$ in order to provide more significant digits for statistical software. As two independent portions of each sample were analyzed, the averaged normalized areas (except for cocaine and benzoylecgonine) were used as variables in multivariate statistical analysis.

PCA was applied to the set of 29 samples in order to find their correlations and to gain some useful information about how variable loadings could be associated to a potential separation between groups.

2.6.2. Residual solvents

More than 80 compounds were detected by HS-GC-MS; some of them were found only in one or two samples. Signals of compounds were normalized to the sum of all peak areas and multiplied by 10$^6$ in order to provide more significant digits for statistical software. Since two independent portions of each sample were analyzed, the averaged normalized areas were used as variables in multivariate statistical analysis.

To overcome the correlation between variables given by normalization (their sum is equal to 10$^6$), m-xylene was removed because it was found in samples always with ethylbenzene and p-xylene. Furthermore, m-xylene showed a strong correlation with ethylbenzene.

Variables were further reduced to 36 target solvents; they were processed by PCA in order to highlight a natural grouping of samples according to the fingerprint provided by residual solvents.

Information provided by alkaloids and residual solvents were finally put together, and a PCA was applied to all these variables.

2.6.3. KBr-FTIR and ATR-FTIR spectra

As seen before, adulterants and/or diluents, when incorporated in samples, would affect their infrared spectra; two models using KBr-FTIR and ATR-FTIR spectroscopy were therefore developed with the 7 samples found free from adulterants or diluents. Considering that 10 independent portions were analyzed for each sample, the two models were developed with 70 spectra.

The KBr-FTIR spectra were pre-processed with the baseline correction algorithm.

The ATR spectra were pre-processed by the function of ATR correction and the baseline correction algorithm.

Transmittances were exported using optimized combinations for spectral region, smooth and data spacing, as follows:

1) spectral region:
   a. 3040–2300 cm$^{-1}$ and 1800–400 cm$^{-1}$ for KBr-FTIR;
   b. 3040–2300 cm$^{-1}$ and 1750–500 cm$^{-1}$ for ATR-FTIR; (differences between the above spectral regions “a” and “b” can be explained as follows: ATR-FTIR acquisition ends at 500 cm$^{-1}$ depending on the transparency of the crystal (ZnSe) towards IR radiation, while KBr-FTIR acquisition ends at 400 cm$^{-1}$. Furthermore, the peak shape at 1730 cm$^{-1}$ is different between ATR and KBr spectrum: the region between 1750 and 1800 cm$^{-1}$ does not contain any signal of interest for ATR spectrum whereas it still includes a portion of the peak at 1730 cm$^{-1}$ for KBr spectrum).

2) smooth: 7 (13.499 cm$^{-1}$).
3) data spacing: 8 cm$^{-1}$.

The following statistical tools were applied:

- One-way ANOVA, for picking the most useful variables to provide the same groupings of samples already obtained from alkaloids and residual solvents analysis;
- PCA to the variables selected by ANOVA in order to display the groupings of samples and comparing them with results previously achieved by analyses of alkaloids and residual solvents.

The two models were then compared in order to select the most suitable for the evaluation of “cut” samples, using the following criteria:

- model ability of grouping samples according to results achieved by the previous analyses;
- repeatability and speed of analysis.

“Cut” samples were assessed by linear discriminant analysis (LDA), a “supervised” chemometric tool, where a training set of samples belonging to known classes (pure samples in this case) is used to develop a mathematical model that correctly classifies the samples themselves. “Cut” specimens were processed as an external test set and the class assignment provided by LDA was compared with results achieved by the previous analyses and examined in relation to the cocaine content for each sample.

3. Results and discussion

3.1. Analysis of alkaloids

PCA was applied to the 29 samples and the alkaloids listed in Section 2.6.1. (except for cocaine, benzoylecgonine). The right origin of samples, usually ascribed to one of the three largest manufacturers in the world, Colombia, Peru, and Bolivia [13], is not known. Therefore, the aim of this study was to observe, by applying PCA, how samples are naturally grouped depending on some knowledge about the seizures where available. Samples were displayed highlighting the boarding countries (Dominican Republic: sample nos. 4, 6, 7, 10, 28, and 29; Brazil: sample nos. 5, 9, 12, and 17; Argentina: sample nos. 14, 15, and 16; Mexico: sample no. 18; Colombia: sample no. 13; Guatemala: sample no. 11; and Benin: sample no. 8) or the location where the seizures occurred (Gioia Tauro Harbour: sample nos. 19, 20, 21, 22, 23, 24, 25, 26, and 27).

It is the first three PCs that explain 77.9% of the total variance. From the biplot of PC2 versus PC1 (Fig. 1), it can be seen that samples seized in Gioia Tauro Harbour (except for sample no. 22) tend to cluster on negative values of PC1. Most of the samples shipped from Dominican Republic (except for sample no.10) are grouped in the same region, as well as samples embarked in Mexico and Guatemala. Moreover, sample nos. 2, 5, 9, 16, and 22 are grouped on positive values of PC1 and values of PC2 close to zero. On the basis of alkaloids analysis, it can be seen that sample no. 22 is very different from all other samples seized in Gioia Tauro Harbour. Sample nos. 8, 10, and 12 appear quite close to this cluster.

Regarding the loadings distribution, it can be noticed from the biplot of Fig. 1 that samples having positive values on PC1 can be differentiated due to their higher content of cis-cinnamoylcocaine and trans-cinnamoylcocaine (variables with high loadings on PC1) and lower amount of tropacocaine and trimethoxycocaine (variables with negative loadings on PC1). The opposite happens for samples with negative values on PC1.

Sample nos. 1, 15, and 17 are grouped on positive values of PC2 (their values of PC1 are close to zero) where the norcocaine loading is high. Such samples can be distinguished for their higher content of norcocaine.
Sample no. 13, shipped from Colombia, is not close to any of the clusters described above.

In the light of results coming from alkaloids analysis, a first hypothesis on the origin of some samples could be provided: it is reasonable to suppose a common origin for most of the samples seized in Gioia Tauro Harbour and those shipped from Guatemala, Mexico, and the Dominican Republic. These countries are near enough to each other, and the closest manufacturer is the State of Colombia. However, the sample coming from Colombia (assuming that the cargo shipped from Colombia was produced there) cannot be related to the group above mentioned, only on the basis of alkaloids profiling.

3.2. Residual solvents analysis

More than 80 compounds, strongly related to the extraction and purification of cocaine performed in clandestine laboratories, were detected by HS-GC-MS. However such a large number of chemical variables leads to underline the differences between samples rather than to highlight their natural grouping based on a similarity criterion. Therefore, a small number of solvents to be used as a target for the manufacturing process was selected; variables reduction was achieved as described below. Some compounds were found only in one or two samples: a first selection of variables was then performed by removing the solvents found in samples with a frequency less than or equal to 10% of them. Other solvents were taken out because of coelution problems in their chromatographic peak. A further selection was made by the evaluation of correlations among compounds with a very similar chemical structure (for instance l-ethyl-2-methylbenzene was removed since highly correlated to n-propylbenzene). Finally 36 target solvents (shown in Table 1) were picked out and normalized according to what is described in Section 2.6.2.

PCA was applied in order to highlight a natural grouping of samples depending on some knowledge about the seizures, as it has been done for the alkaloids (Section 2.4.).

The first three PCs explain 48.15% of the total variance. From the scores plots (Fig. 2(a) and (b)), it can be seen that some of the clusters highlighted by alkaloids analysis are found once more by processing residual solvents:

- sample nos. 2, 5, 9, 16, and 22 are grouped on positive values of PC1 and PC3 and negative values of PC2;
- samples seized in Gioia Tauro Harbour (except for no. 22) cluster on negative values of PC1; most of the samples coming from Dominican Republic and the sample shipped from Mexico (no. 18) can be found close to them. Differently from what is achieved through the analysis of alkaloids, even the sample shipped from Colombia (no.13) falls within this cluster.

Table 1  
Target variables picked out for multivariate statistical analysis of residual solvents.

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Identified solvent name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ethyl methyl ether</td>
</tr>
<tr>
<td>2</td>
<td>Ethanol</td>
</tr>
<tr>
<td>3</td>
<td>Acetone</td>
</tr>
<tr>
<td>4</td>
<td>2-propanol</td>
</tr>
<tr>
<td>5</td>
<td>2-butanone</td>
</tr>
<tr>
<td>6</td>
<td>Ethyl hexanoate</td>
</tr>
<tr>
<td>7</td>
<td>Methyl acetate</td>
</tr>
<tr>
<td>8</td>
<td>Ethyl acetate</td>
</tr>
<tr>
<td>9</td>
<td>Isoamyl acetate</td>
</tr>
<tr>
<td>10</td>
<td>N-propyl acetate</td>
</tr>
<tr>
<td>11</td>
<td>Isobutyl acetate</td>
</tr>
<tr>
<td>12</td>
<td>Ethyl chloride</td>
</tr>
<tr>
<td>13</td>
<td>Methylene chloride</td>
</tr>
<tr>
<td>14</td>
<td>N-propyl chloride</td>
</tr>
<tr>
<td>15</td>
<td>Pentane</td>
</tr>
<tr>
<td>16</td>
<td>2,2-dimethylbutane</td>
</tr>
<tr>
<td>17</td>
<td>3-methylpentane</td>
</tr>
<tr>
<td>18</td>
<td>Hexane</td>
</tr>
<tr>
<td>19</td>
<td>Methyl cyclopentane</td>
</tr>
<tr>
<td>20</td>
<td>3-methylxane</td>
</tr>
<tr>
<td>21</td>
<td>Heptane</td>
</tr>
<tr>
<td>22</td>
<td>Methyl cyclohexane</td>
</tr>
<tr>
<td>23</td>
<td>Dimethyl cyclopentane</td>
</tr>
<tr>
<td>24</td>
<td>Nonane</td>
</tr>
<tr>
<td>25</td>
<td>Decane</td>
</tr>
<tr>
<td>26</td>
<td>Tetradecane</td>
</tr>
<tr>
<td>27</td>
<td>Benzene</td>
</tr>
<tr>
<td>28</td>
<td>Toluene</td>
</tr>
<tr>
<td>29</td>
<td>Ethylbenzene</td>
</tr>
<tr>
<td>30</td>
<td>p-xylene</td>
</tr>
<tr>
<td>31</td>
<td>n-propylbenzene</td>
</tr>
<tr>
<td>32</td>
<td>1-ethyl-4-methyl benzene</td>
</tr>
<tr>
<td>33</td>
<td>1,2,3-trimethylbenzene</td>
</tr>
<tr>
<td>34</td>
<td>1,2,4-trimethylbenzene</td>
</tr>
<tr>
<td>35</td>
<td>Benzaldehyde</td>
</tr>
<tr>
<td>36</td>
<td>Benzoic acid methyl ester</td>
</tr>
</tbody>
</table>
It can be assumed, therefore, that samples belonging to this group are compatible with an origin from Colombia which is also, among the producers of cocaine, the nearest country to Mexico and Dominican Republic.

Looking at Fig. 2(a) and (b), it can be noticed that the first group (sample nos. 2, 5, 9, 16, and 22) has a certain degree of similarity with sample nos. 1, 8, 12, 14, 15, and 17: they all have positive values on PC1. Therefore, an affinity was found among a seizure from Gioia Tauro Harbour (sample no. 22) and samples shipped from Brazil (nos. 5, 9, 12, and 17) and Argentina (nos. 14, 15, and 16) as well as with the sample that followed the African route (no. 8).

The loadings distribution on PC1 and PC2 is shown in Fig. 2(c) from which it can be seen that the derivatives of benzene tend to positive values on PC1 (except for toluene), while the aliphatic hydrocarbons have negative loadings on PC1 (except for tetradecane). Furthermore, halogenated solvents have a significant loading on negative values of PC2. Ethyl chloride and n-propyl chloride also have a negative loading on PC1. Acetone and ethanol have a positive loading on PC1 and a negative loading on PC2. Further remarks are presented in Section 2.6, where the combination of the analysis of alkaloids and residual solvents is discussed.

Finally, sample nos. 11 and 29, while having a greater similarity with the second grouping (it can be seen in Fig. 2(a) and (b)), show the following features:
- sample no. 11, due to the highest values on PC2 and PC3, tends to separate from the second group: it can be distinguished for more significant traces of hydrocarbons, especially aliphatic. Taking into account that aliphatic hydrocarbons are not naturally present in coca leaves, they represent a marker of the laboratory of illegal production;
- sample no. 29, due to its high negative value on PC1, can be recognized for the high content of ethyl chloride and n-propyl chloride: these solvents were seldom found in the other samples and at less significant amounts.

### 3.3. PCA applied to a data set made up by alkaloids and residual solvents

Analyses of alkaloids and residual solvents were put together and processed by PCA in order to gather all the information provided apart by these parameters (see Sections 2.4 and 2.5). The first three PCs
explain 48.03% of the total variance. The scores-plots are shown in Fig. 3(a) and (b): the two groupings outlined on PC1 by the previous statistical analyses are confirmed. The first cluster is divided into two subgroups on the PC2 axis, while the difference tends to shrink on PC3. Regarding to the loadings distributions (Fig. 3(c)), the following observations can be done:

- the first grouping (sample nos. 2, 5, 9, 16, 22, 1, 8, 12, 14, 15, and 17) can be distinguished for a significant content of cis-cinnamoylcocaine, trans-cinnamoylcocaine, and derivatives of benzene. Within this group some differences are found:
  - samples nos. 2, 5, 9, 16, 22 show a greater content of acetone, ethanol, and methylene chloride; taking into account that acetone was often found at significative amounts in samples originating from Bolivia [7], samples of this subgroup might be associated to Bolivia as origin country;
  - samples nos 1, 8, 12, 14, 15, and 17 show a higher content of norcocaine (an impurity due to the use of oxidants during the clandestine production; it is not naturally found in coca leaves), 2-propanol, and isopropyl acetate;

- the second cluster (sample nos. 3, 4, 6, 7, 10, 11, 13, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, and 29) stands out for a significant content of tropacocaine, trimethoxyccocaine, and eggonine methyl ester. As for residual solvents these samples can be distinguished for aliphatic hydrocarbons, toluene, isobutyl acetate, n-propyl acetate, 2-butanol, and ethylhexanoate. At this regard, it is known that samples originating from Colombia are often found to contain significative amount of exane (here detected among aliphatic hydrocarbons), n-propyl acetate, and toluene [7]. This grouping of samples might be associated therefore to Colombia as origin country.

It is eventually observed that the total weight of exhibits seized during 2012 from which samples associated to Colombian origin were taken and analyzed is equal to 909.72 kg which corresponds to 92% by weight of seizures of 2012 analyzed for the present work (988.42 kg); whereas the total weight of the seizures from which samples associated to Bolivian origin were extracted is equal to 62.7 kg, corresponding to 6% of the seizures of 2012 analyzed for the present work.
3.4. KBr-FTIR and ATR-FTIR analyses.

3.4.1. Development of models based on KBr-FTIR and ATR-FTIR spectroscopy

A drug profiling model based on FTIR spectroscopy should be developed analyzing samples free from adulterants. Considering that only 7 exhibits (nos. 2, 9, 20, 22, 23, 25, and 27, coming from large seizures) were available for this study, 10 independent portions for each sample were analyzed and processed, simulating an analysis of 10 samples belonging to the same batch. Moreover, this approach allowed a first assessment on the repeatability associated with FTIR spectroscopy in drug profiling issues.

In relation to results achieved by applying PCA to alkaloids and residual solvents, specimens were supposed to be grouped as follows:

- samples nos. 2, 9 and 22, forming the class 1 (of 30 spectra);
- samples nos. 20, 23, 25 and 27, forming the class 2 (of 40 spectra).

Processed variables were the transmittances corresponding to each wave number for the spectral regions specified in Section 2.6.3. First of all, a one-way ANOVA was performed in order to compare, for each variable, the variance within any category with the one between categories; the more significant variables, forming the spectral regions highlighted in Fig. 4, were then picked out for PCA.

As for KBr-FTIR, PCA was applied to the spectral regions selected by ANOVA and highlighted in Fig. 4(a), corresponding to 103 variables. The first two PCs explain 89.89% of the total variance; their scores plot is shown in Fig. 5(a). With regard to ATR-FTIR, PCA was applied to the spectral regions selected by ANOVA and highlighted in Fig. 4(b), corresponding to 133 variables. Through the evaluation of Fisher weights (calculated, for each couple of classes and for each PC, through the ratio of the between-category variance and the within-category variance), the most significant components for the purpose of grouping samples according to their class were found to be PC1 (which explains...
86.41% of the total variance) and PC4 (accounting 0.94% of total variance). The scores plot of PC1 versus PC4 is shown in Fig. 5(b).

Separation between the two classes was achieved with both models based on KBr-FTIR and ATR-FTIR spectroscopy. Results look rather equivalent, except that with KBr-FTIR, the two groupings were observed on the first two PCs (explaining a variance of 89.89%), while in ATR-FTIR, the two classes are displayed on PC1 and PC4 (accounting a variance of 87.35%).

The two models can also be compared regarding their repeatability. Considering that all the IR spectra acquired for this study perfectly overlapped and spectral data were processed in a multivariate way, a first assessment on the repeatability could be done by the evaluation of samples dispersion along their principal components.

A view of the two models where samples are divided in 7 classes (corresponding to the 7 batches they belonged) is shown in Fig. 6. Regarding KBr-FTIR (Fig. 6(a)), it can be seen that the batches of the sample nos. 2, 9, and 22 (forming the class 1) are quite overlapped. The same situation arises for the batches of the sample nos. 20, 23, 25, and 27 (forming the class 2). On the other hand, in ATR-FTIR (Fig. 6(b)), all the batches (except no. 23) show boundaries rather narrow and less overlapped between each other. The simultaneous analysis of Figs. 5 and 6 allows, for each model, a graphical comparison between the within-category variance (that is the variance within classes 1 and 2) and the within-batch variance (linked to the repeatability, considering that the 10 test portions are taken from a well-homogenized sample, according to Section 2.1). It can be seen that in KBr-FTIR, the boundaries of the batches tend to overlap with the boundaries of their category; therefore, the within-batch variance does not look significantly lower than the within-category one. The opposite happens for ATR-FTIR, where the within-batch variance looks significantly lower than the within-category one. Such a graphical comparison suggests for the model based on ATR-FTIR spectroscopy a better repeatability than it is for the KBr-FTIR spectroscopy one.

Fig. 5. (a) Scores plot of PC1 versus PC2 for the model based on KBr-FTIR spectroscopy. (b) Scores plot of PC1 versus PC4 for the model based on ATR-FTIR spectroscopy. ▲ Sample nos. 2, 9, and 22 (class 1). ▼ Sample nos. 20, 23, 25, and 27 (class 2).
The within-category and the within-batch variances for PC1 have been calculated for each model and compared by a one-tailed F-test at 0.05 significance level, obtaining the following results:

- KBr-FTIR: the within-category variance is not significantly higher than the within-batch variance for each of the 7 samples; therefore, the analytical repeatability is not significantly lower than the variability that could be found between samples belonging to the same class (i.e. having the same origin) but taken from different seizures.

- ATR-FTIR: the within-category variance is significantly higher than the within-batch variance for all the 7 samples except no. 23 (according to results achieved by the previous graphical evaluation).

This comparison suggests that the model based on ATR spectroscopy shows a better performance, in terms of repeatability, than the one based on KBr spectroscopy.

As regards the speed of analysis of the two techniques, it is noted that ATR-FTIR requires only the time associated with the instrumental reading, not involving any sample preparation. On the other hand, KBr-FTIR involves the weight of the sample and KBr, their homogenization in mortar and compression to obtain a pellet. The better repeatability and the higher speed of analysis given by ATR-FTIR with respect to KBr-FTIR spectroscopy allow definitely to choose the first model for developing a “fast profiling” method.

3.4.2. Application of the model based on ATR-FTIR spectroscopy to the analysis of “cut” samples

The FTIR method was developed using PCA, an exploratory tool which offers a general overview of the subject in question, showing the relationship that exists among objects as well as between objects and variables. Indeed, PCA allows to study and understand such systems, helping the human eye to see in two or three dimension systems that otherwise would necessarily have to be seen in more than three dimensions.

Fig. 6. Scores plots of models based on FTIR spectroscopy where samples are divided in 7 classes, corresponding to the 7 batches they belonged. (a) KBr-FTIR. (b) ATR-FTIR.
dimensions in order to be studied; regarding the FTIR methods presented in this paper, they involved a hundred of variables. PCA allows data to maintain their original structure, making only an orthogonal rotation of variables, which helps to simplify the visualization of all the information already contained in the data. However, the assessment of “cut” samples requires a supervised chemometric tool such as LDA, where, first of all, the 70 spectra of pure samples were used to develop a mathematical model that correctly classifies the samples themselves.

Because discriminant analysis provides reliable results if the ratio between the number of samples and variables is greater than 3 [14], the variables processed were the first ten principal components instead of the original transmittances. The cross validation was performed with 10 cancellation groups, obtaining a classification and a prediction ability both of 97.56%.

Such a prediction ability is effective for “pure” samples (each of the 10 cancellation groups above mentioned was excluded, one at a time, from the training set of 70 samples). However, this method should be tested with the remaining 22 “cut” samples. They were analyzed in triplicate by ATR-FTIR, and the transmittances were averaged. Each sample was described by the first ten principal components computed by adding one sample at a time to the data set of the 70 pure samples. “Cut” samples were finally processed by LDA as an external test set: they were assigned, by LDA, to the class which they showed the greater similarity to.

A synthesis of classification results is reported in Table 2, where for each sample is indicated:

- the class that it is supposed to belong, based on the results achieved by applying PCA to alkaloids and residual solvents;
- the class assignment provided by LDA;
- the percentage of cocaine detected according to what is described in Section 2.4;
- the diluents and/or adulterants detected by GC-MS analysis of alkaloids.

From Table 2, it can be noticed that the comparative analysis through ATR-FTIR provided results rather consistent with those obtained by alkaloids and residual solvents. However, results given by ATR-FTIR method are influenced by the presence of diluents and/or adulterants in the sample. Indeed, it can be remarked that:

- for samples containing only one substance, prediction errors occur for percentages of cocaine lower than 50% (i.e. sample no. 13);
- if more than one substance was detected, prediction errors occur for percentages of cocaine lower than 60% (i.e. sample nos. 17, 7, and 8).

4. Conclusions

Analyses of alkaloids and residual solvents, processed with chemometric tools and combined with both knowledge coming from scientific literature and information about the boarding countries, allowed to make a hypothesis about the geographical origin of samples. Assuming the representativeness of the samples analyzed, it can be argued that for the seizures of 2012 examined in this work (corresponding to 18.6% of total seizures made in Italy and 28.5% of cocaine seized only in the Customs areas):

- 92% by weight might be associated to Colombian origin;
- 6% by weight might be associated to Bolivian origin;
- the remaining 2% shows elements of similarity, although not totally identical, with the samples of Bolivian origin.

Eventually, the ATR-FTIR method for comparative analyses of cocaine seizures proved to be very fast and effective. For the model development (a kind of learning phase), pure samples belonging to known classes need to be analyzed. LDA provided, for pure samples, a classification and a prediction ability both of 97.56%. The new method can be usefully employed for the comparative analysis of large seizures because although they are sometimes added with adulterants such as levamisole, the concentration of cocaine is often still high to enable the method to work. Indeed, for samples added with only one substance, prediction errors start to occur for percentages of cocaine lower than 50%. Indeed, it is not negligible that infrared spectroscopy could be a very useful technique for a rapid assessment on some heterogeneities in a big seizure due to different origins of the respective seizure portions. Compared with other fast methods such as UHPLC-MS/MS, infrared spectroscopy

<table>
<thead>
<tr>
<th>Sample identification</th>
<th>Class supposed, based on the results of PCA (applied to alkaloids and residual solvents)</th>
<th>Class provided by LDA (applied to ATR-FTIR data)</th>
<th>% of cocaine (GC-FID analyses)</th>
<th>Diluents and/or adulterants detected by GC-MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>63.7</td>
<td>Levamisole</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>47.6</td>
<td>Levamisole</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td>66.6</td>
<td>Caffeine, hydroxyzine</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>71.26</td>
<td>Saccharides, levamisole</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td>68.0</td>
<td>Caffeine, levamisole, hydroxyzine</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
<td>49.5</td>
<td>Phencetin, levamisole</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
<td>25.9</td>
<td>Phencetin, levamisole, saccharides</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>66.3</td>
<td>Caffeine, levamisole</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>2</td>
<td>65.2</td>
<td>Levamisole</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>74.5</td>
<td>Mannitol, levamisole</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>1</td>
<td>45.9</td>
<td>Levamisole</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>1</td>
<td>76.5</td>
<td>Levamisole</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>73.1</td>
<td>Levamisole</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1</td>
<td>72.3</td>
<td>Mannitol</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>1</td>
<td>58.8</td>
<td>Levamisole, glutamate</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>2</td>
<td>55.2</td>
<td>Mannitol, levamisole, glucose</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>2</td>
<td>69.5</td>
<td>Levamisole</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>2</td>
<td>70.5</td>
<td>Levamisole</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>2</td>
<td>68.2</td>
<td>Levamisole</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>2</td>
<td>49.9</td>
<td>Levamisole, glucose</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>2</td>
<td>63.3</td>
<td>Levamisole</td>
</tr>
</tbody>
</table>
offers the advantage to be cheaper despite its limitation due to the interference of adulterants or diluents on the spectra; furthermore, FTIR analyses can be easily carried out where the seizures take place.

References