Rapid classification and quantification of cocaine in seized powders with ATR-FTIR and chemometrics

INTRODUCTION

Cocaine is the second most commonly seized drug of abuse entering Belgium, both for local consumption (street samples) and for distribution around Europe (trafficking samples). This alkaloid drug, originating mainly in South America and the Caribbean, enters the Belgian territory via the harbours and airports.[1] In 2014, the average amount of cocaine in samples entering Belgium was above 60 w% for 95% of the samples analysed in the National Institute of Criminalistics and Criminology (NICC). Cocaine street samples, however, are usually cut with adulterants that are pharmacologically active and/or inert cutting agents. The average amount of cocaine in street samples in 2014 was above 30 w% for 95% of the analysed samples and the most common adulterants were levamisole and phenacetin (NICC data). In Belgium, the seized cocaine is mainly in the hydrochloride salt form. The base form is only detected in less than 5 samples per year.[1]

The Belgian average figures are in agreement with the European data. The mean purity of cocaine in the European samples was above 33 w% in 75% of the samples in 2014[2] and the most frequent adulterants were levamisole, phenacetin and caffeine.[3]

For routine screening analysis of unknown drug seizures, colour tests are generally used by law enforcement agencies and forensic laboratories to quickly define the presence of a drug or a class of drugs. However, these colour tests present some major drawbacks such as lack of selectivity and difficulty in interpretation of the colour.[4–6]

In the NICC laboratory, Fourier Transform InfraRed (FTIR) spectrometry using Attenuated Total Reflectance (ATR) sampling (abbreviated as ATR-FTIR), is used as an initial screening, prior to performing laborious and expensive gas chromatography coupled with mass spectrometry (GC-MS) and gas chromatography with flame ionization detection (GC-FID) analyses for identification and quantification.

ATR-FTIR spectrometry requires minimal sample preparation (just homogenization of powders) resulting in a fast and low-cost method and is available in a portable device.[7] However, despite the fact that ATR-FTIR provides a molecular fingerprint of the compounds present in the sample, detection of cocaine in the samples is not always possible. A challenging issue when analysing cocaine samples is the presence of many different adulterants and/or cutting agents.

For identification of unknown compounds by ATR-FTIR, spectral libraries are used to match the spectra of known compounds to the unknown spectrum. A successful identification is crucially dependent on the available entries in the libraries. Moreover, while spectra of pure samples can be easily matched against a library, this is less feasible for mixtures containing multiple components. Even commercially available multicomponent search algorithms show limitations in their practical usefulness. Considering all these issues, the main limitation in the identification of cocaine with ATR-FTIR is the library matching, which is dependent on the cocaine concentration and the presence of adulterants and/or cutting agents and their concentration in the mixture. When the spectral contribution of the adulterant(s) and/or cutting agent(s) starts to dominate the spectra, cocaine is no longer detected. Finally, the result of the library matching always needs to be evaluated by a trained operator.

To overcome these limitations, ATR-FTIR can be combined with chemometric data processing, a powerful tool for extracting relevant information from high dimensional data (for example spectra) allowing both qualitative and quantitative analyses.[8–12]

The aim of this study was to optimize the routine screening of cocaine in unknown drug powders and to investigate the potential for quantitative application by using ATR-FTIR in combination with chemometric data processing techniques.

Representative street drug powders (n=482), seized during the period January 2013 to July 2015, and reference powders (n=33) were analysed with an ATR-FTIR spectrometer and also by means of GC-MS and GC-FID.

An advanced machine learning technique Support Vector Machines (SVM)[13,14] was compared with a commonly used chemometric technique, Partial Least Squares (PLS)[15,16]. Machine learning refers to computer algorithms that improve through experience.[9] The developed models were evaluated by determination of different qualitative and quantitative performance parameters.

As far as the authors are aware, the combination of ATR-FTIR and SVM has not yet been used to classify (n=515) and quantify cocaine (n=378) in such a large set of unknown drug powders.

EXPERIMENTAL

Seized samples

All street samples (drug powders) were seized in different regions of Belgium and analysed in the Drugs Laboratory of the NICC using three analytical techniques: GC-MS, GC-FID and ATR-FTIR. Prior to analysis, all samples were homogenized with a mortar and pestle.

Reference material

Cocaine hydrochloride (99.4%) was purchased from Lipomed (Arlesheim, Switzerland). Phenacetin (98.0%), diltiazem hydrochloride (p.a.), lidocaine hydrochloride monohydrate (p.a.), procaine hydrochloride (p.a.), hydroxyzine dihydrochloride (98.0%), benzocaine (p.a.), acetaminophen (99.0%), ephedrine hydrochloride (99.0%), creatine (p.a.), diphenhydramine hydrochloride (99.0%), phenylephrine hydrochloride (p.a.), ibuprofen (98.0%), antipyrine (98.0%), atropine (99.0%), ascorbic acid (99.0%) and myo-inositol (99.0%) were purchased from Sigma-Aldrich (St Louis, MO, USA). Benzoic acid (99.5%) and levamisole hydrochloride (99.0%) were purchased from Acros organics (Morris Plains, NJ, USA).

Caffeine (p.a.), boric acid (99.0%), glucose anhydrous (p.a.), maltose monohydrate (p.a.), starch (p.a.) were purchased from VWR International (West Chester, PA, USA) and D-sorbitol (99.0%) from Merck chemicals KGaA (Darmstadt, Germany). Dextromethorphan hydrobromide (p.a.) was purchased from Alfa Aesar (Karlsruhe, Germany).

GC-MS and GC-FID analysis

GC-MS and GC-FID analysis were used as reference methods for identification and quantification of cocaine powders. At the NICC Drugs laboratory, these methods are accredited by the ISO17025 standard[17]. The GC-FID method has an expanded uncertainty measurement of 6.4% (relative; with a coverage factor of 2) and a limit of quantification (LOQ) of 0.8%. Quality is assured with an in-house quality control (QC) sample and participation to proficiency tests.

Each homogenized sample was weighed (20 ± 5 mg) and dissolved in 10 mL internal standard solution (0.5 mg tribenzylamine (99+%, Alfa Aesar, Karlsruhe, Germany) per mL ethanol (Biosolve BV, Valkenswaard, The Netherlands). Then, 1 mL of this solution was transferred to a glass vial, sealed and subjected to chromatographic analysis.

GC-MS analysis (HP6890N-5973N, Agilent Technologies, Santa Clara, CA, USA) was performed for identification, based on comparison with in-house libraries (retention time and spectra). An Agilent DB5-MS column (15.0 m x 250 µm x 0.25 µm) was used with helium as carrier gas at constant pressure with retention time locking. The oven temperature was initially set as 100°C and then increased to 325°C. A volume of 1 µL was injected in the split mode with a split ratio of 40:1. The run time was 14.25 min. MSD Chemstation software (Agilent Technologies) was used for data retrieval.

GC-FID analysis was subsequently done for quantification purposes, using a HP6890N (Agilent Technologies) equipped with an autosampler 7683B Series. Separation was achieved on a HP 5 column (0.52 µm film thickness x 0.32 mm I.D. x 25 m; J & W Scientific, Agilent

Technologies). The helium carrier gas had a constant pressure with retention time locking. A 1µL sample was injected with a split ratio of 25:1 at 280°C. The oven temperature was initially programmed at 150°C (hold 1 min) and then increased to 300°C. Detector parameters were set at a temperature of 320°C, a makeup flow of nitrogen of 25 mL/min, a hydrogen flow of 30 mL/min and an air flow of 400 mL/min. The run time was 16.50 minutes. Chemstation software (Agilent Technologies) was used for data retrieval. Calibration was performed with 6 dilutions of a pure cocaine reference powder.

ATR-FTIR measurements

Spectra of reference materials and street samples were acquired using a portable FTIR spectrometer with a single reflection diamond crystal ATR accessory with pressure applicator (Mobile-IR, Bruker Corporation, Ettlingen, Germany). The pressure is approximately 80N according to Bruker's specifications. Prior to the measurement of the samples and every 10 measurements, a background measurement was collected with the empty ATR cell. The ATR-FTIR spectra were recorded from 4000 to 500 cm^{-1} with a resolution of 4 cm^{-1} (2440 data points). Each spectrum was an average of 24 scans. The measurements were obtained in absorbance mode. After each measurement, the ATR surface was thoroughly cleaned with water and isopropanol. Both an in-house spectral ATR-FTIR library, consisting of reference materials, and commercial libraries (Bruker; Merck; S.T. Japan, Tokyo, Japan) were used for library matching.

Description of spectral dataset

A dataset of 515 spectra was used to build and validate classification models for cocaine. Among these 515 spectra, there were 378 representative cocaine street samples seized during the period from January 2013 to July 2015 with concentrations ranging from 4 to 99 w% cocaine hydrochloride. These cocaine concentrations, obtained with GC-FID analysis, are not normally distributed with a mean concentration of 70 w% (SD ± 22) and a median

concentration of 76 w%. This distribution is representative for the cocaine samples annually analysed in the NICC.

In these samples the following adulterants and cutting agents were identified with ATR-FTIR and/or GC-MS analysis: acetaminophen, benzocaine, boric acid, caffeine, diltiazem, hydroxyzine, levamisole, lidocaine, phenacetin, procaine and sugars.

Besides the cocaine street samples, also 137 spectra of samples without cocaine were included to build the classification model: 14 reference materials of adulterants and cutting agents present in the seized cocaine samples (as described above), 19 reference materials of white powders (cellulose, inositol, maltose, fructose, glucose, creatine, griseofulvin, ascorbic acid, ephedrine, diphenhydramine, atropine, benzoic acid, acetylsalicylic acid, antipyrine, ibuprofen, bupivacaine, phenylephrine, paracetamol and dextromethorphan) and 104 seized samples without cocaine were included. The seized street samples consisted of: adulterants (benzocaine, paracetamol, caffeine and phenacetin), heroin, amphetamine, methamphetamine, 3,4-methylenedioxy-methamphetamine, sugars (sorbitol and lactose), other drugs (ibuprofen, paracetamol, sildenafil, methadone, bromazepam and ketamine), new psychoactive substances (methylone and 25I-NBOMe) and diluting agents (starch, sodium hydroxide, milk powder, washing powder and calcium carbonate).

To build and validate quantification models for cocaine, only the samples with cocaine (n=378) described above, were included.

Chemometric analysis

Chemometric software was used to develop and validate classification and quantification models for cocaine. The models were built using the full ATR-FTIR spectrum in order to take the information of the adulterants and/or cutting agents and other impurities (such as natural alkaloids) into account. These compounds can have multiple bands with local correlations in the spectrum and these bands can overlap each other.[10] This outlines the need of

chemometric data processing. Figure 1 shows the ATR-FTIR spectra in the fingerprint region (1800-500 cm^{-1}) of a pure cocaine reference (a) and three cocaine street samples (b-d) with a (b) high (80 w%), (c) medium (67 w%) and (d) low (12 w%) cocaine content. The main vibrational bands of cocaine hydrochloride are marked in Figure 1 (black dashed lines).

The ATR-FTIR spectra were imported as Opus file format (Bruker) in the chemometric software (PLS Toolbox, Version 8.1, Eigenvector Research, Inc., Manson, WA, USA) and were then pre-processed using standard normal variate (SNV) transformation. SNV removes scatter effects by centering and scaling each individual spectrum based on the following formula (equation 1)[18]:

$$x'_{i,j} = \frac{x_{i,j} - \mu_i}{\sigma_i} \quad \text{(eq. 1)}$$

where $x'_{i,j}$ is the corrected value, $x_{i,j}$ the original absorbance, μ_i the mean of the i^{th} spectrum and σ_i the standard deviation of the i^{th} spectrum, j varying from 1 to k, the number of variables (k=2440).

SVM with linear kernel and PLS analysis were performed with the PLS Toolbox. SVM is a supervised machine learning algorithm that can be used for classification (called SVM discriminant analysis, SVM-DA) and quantification purposes (called SVM regression, SVMR). SVM uses kernel-functions (e.g. linear, polynomial, sigmoid, etc) to project data into a high-dimensional feature space. For a more detailed explanation of SVM theory, many references are available on kernel methods and SVM.[11,13,14,19–21]

PLS is a very well-known regression technique that has been widely described in literature and can be applied for classification (called PLS discriminant analysis, PLS-DA) and/or regression analysis of spectral data (called PLS regression, PLSR).[15,16,21–25]

More details about the mathematics behind the chemometric techniques used in this study can be found in literature.[11,13–16,19,21–25]

Model evaluation

To build and evaluate the constructed models, double cross-validation was performed.[26] For classification, the dataset (n=515 spectra) was randomly divided into five subsets by PLS toolbox. Four subsets (n=412 calibration spectra, 80% of the samples) were used to construct a model and the remaining subset (n=103 test spectra, 20% of the samples) was used to evaluate the model. This process is repeated five times until all samples are tested. Additionally, within the calibration dataset, a cross-validation (with five subsets and one iteration) was carried out to optimize the model, i.e. to find the optimal number of latent variables (LVs) for PLS, the optimal number of support vectors (SVs) and the best combination of the 2 parameters cost and epsilon for SVM. Cost [0 - ∞] represents the penalty associated with errors larger than epsilon. Increasing cost value causes closer fitting to the calibration. In the training of the regression function there is no penalty associated with points which are predicted within distance epsilon from the actual value. Decreasing epsilon forces closer fitting to the calibration data. The predictions of each classification model were compared to the results of the reference method GC-MS by calculating the following qualitative performance parameters: sensitivity, specificity and efficiency. Formulas and additional details about the performance parameters are described in the study of López et al.[27]

To build and evaluate the quantification models, the same double cross-validation strategy was applied using subsets only consisting of samples with cocaine (n=378 spectra). The root mean squared error of calibration (RMSEC), cross-validation (RMSECV) and prediction (RMSEP), the coefficients of determination (R^2calibration, R^2cross-validation and R^2validation) and the bias were calculated.

The root mean square errors (RMSEC, RMSECV and RMSEP) were calculated as follows (equation 2):

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{y_i - \hat{y}_i^2}{n}} \quad \text{(eq. 2)}$$

where y_i is the reference concentration of the sample i, $\hat{y}_i$ is the cocaine concentration predicted by the model and n is the number of samples.

The repeatability and the intermediate precision (expressed as standard deviation (SD)) were evaluated by performing replicate analyses of a pure cocaine reference powder (99.4%) and two in-house QC samples (69.0% and 34.2%). These samples were measured 8 times on the same day (repeatability) and on 8 different days (intermediate precision).

The in-house QC sample of 69.0 w% is a cocaine street sample consisting of 69.0 w% cocaine and 24.0 w% levamisole. The other in-house QC sample of 34.2 w% is a cocaine street sample consisting of 34.2 w% cocaine, 30.0 w% phenacetin, 7.4 w% levamisole and benzocaine (not dosed).

The reproducibility was evaluated by predicting spectra of proficiency samples from 2014 to 2015 (UNODC, LGC and ENFSI). Results were evaluated by calculating a z-score using the predicted value and the mean cocaine concentration and SD of the participating laboratories. Outliers were detected by Grubbs' test and are not included in mean and SD.

Formulas and additional details about the performance parameters are described in the ASTM standard.[28]

RESULTS

Classification of cocaine

In a first part of this study, we investigated the use of ATR-FTIR spectra with chemometric techniques to detect cocaine in unknown drug powders. An advanced machine learning technique SVM-DA was compared with a commonly used chemometric technique PLS-DA. Table 1 summarizes the SVM-DA and PLS-DA cross-validation and validation results. In this contingency table, the number of correctly predicted and misclassified samples obtained with

each calibration set (cross-validation) and each test set (validation) are presented. It can be concluded that the number of misclassified samples for cross-validation and validation was lower with SVM-DA (2 to 6 for cross-validation and 0 to 3 for validation) in comparison with PLS-DA (9 to 18 misclassifications for cross-validation and 2 to 6 for validation) (Table 1). Based on the predictions of the test spectra: the number of true positives, true negatives, false positives, false negatives, sensitivity, specificity and efficiency were calculated (Table 2). Among the 19 misclassifications obtained with PLS-DA, 9 false positives and 10 false negatives were observed. The false positive samples consisted of reference spectra without cocaine (benzocaine, diphenhydramine, boric acid, mannitol, lidocaine, diltiazem and levamisole) and seized samples without cocaine (boric acid and phenacetin). The false negative samples were cocaine street samples with concentrations below 36 w%. Based on the misclassifications, the sensitivity of PLS-DA over the entire validation dataset was 97%, the specificity and efficiency were 94 and 96% respectively (Table 2). Figure 2 shows the probabilities for each sample obtained with PLS-DA. Samples with cocaine are presented as black squares and samples without cocaine as grey rhombs. The false negatives are represented by red circles and the false positives by green triangles. As shown in this figure, there is an overlap of non-cocaine samples (class 0) and cocaine samples (class 1). It is observed that false positives, such as boric acid, had probabilities above 0.999 and were wrongly classified as 'cocaine detected'. Among the false negatives (classified as 'cocaine not detected'), probabilities below 0.02 were observed.

With SVM-DA only one false positive, e.g. the pure reference mannitol, was observed. SVM-DA also resulted in 5 false negatives, which were street samples with a cocaine content below 33 w%. Based on these validation results we can conclude that SVM-DA has a high sensitivity, specificity and efficiency of 99% (Table 2). In figure 3, the probabilities for each sample obtained with SVM-DA are plotted. As shown in this figure, the samples with cocaine

are clearly separated (black square) from those without cocaine (grey rhomb), with the exception of the six misclassified samples. It is observed that samples with probabilities above 0.62 are correctly classified as 'cocaine detected' and that samples with probabilities below 0.10 as 'cocaine not detected'. Between 0.10 and 0.62 there is an overlap of non-cocaine samples (class 0) and cocaine samples (class 1). If the screening is done respecting these probabilities, 100% of the samples are correctly classified.

To further evaluate the boundaries, 14 white powders of proficiency tests were predicted by the SVM-DA model and they were all correctly assigned as 'cocaine not detected' (probabilities < 0.10) or 'cocaine detected' (probabilities > 0.62) (Table 3).

Considering the results obtained by cross-validation and validation (Table 1 and 2), it can be concluded that for this dataset SVM-DA performed better than PLS-DA.

Quantification of cocaine

In a second part of this study, we investigated the use of ATR-FTIR spectra with chemometric techniques to quantify cocaine in seized drug powders. The same machine learning SVM technique as used in the classification part was compared with the widely used PLS technique. To build these quantification models, only the samples with cocaine (n=378 spectra) were included in the dataset since it is aimed to only quantify the samples that are positively identified as 'cocaine detected' by the classification model. The cocaine concentrations of the samples in the dataset, determined by GC-FID analysis, are not normally distributed with a median concentration of 76 w%. This distribution is representative for the cocaine samples annually analysed in the NICC.

The performance of the SVMR and PLSR quantification models are shown in table 4. For each model using SVMR or PLSR, the prediction parameters RMSEC(V), RMSEP and R^2 values are presented. The SVMR models showed RMSECs and RMSECVs ranging between 4.0 and 6.7% (Table 4). The R^2 values remained almost constant (R^2c: 0.95-0.97 and R^2cv:

0.91-0.93). With PLSR similar results were obtained with RMSECs and RMSECVs ranging from 4.9% to 7.0%. Like SVMR, the R^2 values remained almost constant (R^2c: 0.94-0.95 and R^2cv: 0.90-0.92). These validation parameters indicated a good correlation with the calibration parameters. All SVMR models showed reasonable RMSEP values (ranging between 5.4% and 7.1%) and good coefficients of determination (R^2v) between the predicted and the GC-FID values (ranging between 0.90 and 0.95). With PLSR, comparable results were obtained with RMSEP values (ranging between 5.6% and 7.1%) and R^2v values (ranging between 0.89 and 0.95) (Table 4).

Figure 4 summarizes the results for the five datasets of SVMR and PLSR. SVMR and PLSR showed an accurate prediction (R^2v = 0.92) of the validation samples with a RMSEP of 6.3% and 6.5% respectively.

The difference of 0.2 between these RMSEPs is not significant[29] and it can be concluded that for the used dataset, the machine learning technique SVMR and the traditional technique PLSR quantification model performed equally.

The results for repeatability and intermediate precision, expressed as SD, for SVMR and PLSR are presented in Table 5. Repeatability varied between 1.0 and 2.6 for SVMR and between 0.5 and 2.7 for PLSR. Intermediate precision ranged between 1.6 and 4.2 for SVMR and between 1.6 and 3.8 for PLSR (Table 5). The z-scores for the proficiency samples ranged between -1.8 and 1.4 for SVMR and PLSR and can be considered successful ($|z| \leq 2$), with the exception of one proficiency sample predicted by PLSR (z = -3.6) (Table 6).

DISCUSSION

For routine screening analysis, it is important to have a fast and simple method to define the presence of a drug or a class of drugs. Therefore, colour tests are generally used by law enforcement agencies as well as by forensic laboratories. In the NICC laboratory, ATR-FTIR spectrometry is used as an initial screening, before performing GC-MS and GC-FID analyses.

The identification of cocaine in mixtures is, however, not always feasible. A correct identification is dependent on the library matching. Commercial libraries are expensive and are built up from spectra of pure substances. An in-house library can be of interest but making such a library is time-consuming and expensive. Moreover, the interpretation of identification results has to be performed by a trained operator. Commercially available instrument software for the analysis of mixtures exists but are also based on a library search. Therefore, the limitations are comparable (expensive libraries and need for trained staff for interpretation of results).

This paper shows that ATR-FTIR spectrometry combined with chemometric data processing can be used as a rapid method for the classification and quantification of cocaine in seized drug powders.

In the first part of our study, a comparison of PLS-DA and SVM-DA classification models was performed. During the validation, 19 misclassifications out of 515 were obtained with PLS-DA, resulting in a total misclassification error of 3.7%. In contrast, SVM-DA showed only 6 misclassifications which means that only 1.2% of all the samples was misclassified. By adding these 6 misclassified spectra into the calibration dataset, all samples were correctly classified with SVM-DA. This proofs that the misclassifications were inherent to the validation process, pointing out their uniqueness and necessary presence in the calibration dataset.

It should be emphasized that cocaine street samples have no fixed matrix. The type, the number and the concentration of adulterants and cutting agents vary. In future, new adulterants and cutting agents which are unknown to the developed models can appear on the illicit drug market. Based on a survey performed by the Drugs Working Group of the European Network of Forensic Science Institutes (DWG ENFSI)[3], it can be concluded that the presence of an adulterant or cutting agent in a sample is correlated to a geographical

provenance. For example, tetracaine was detected in cocaine samples seized in Spain, griseofulvin in Poland and creatine in Czech Republic and Serbia. The adulterants benzocaine, procaine, creatine, tetracaine, hydroxyzine and griseofulvin were reported as one of the three most frequent in cocaine street samples, whereas in Belgium these adulterants are rarely or never seized. SVM-DA was able to correctly classify these unknown adulterants/cutting agents (such as creatine and griseofulvin) as 'cocaine not detected', even when not yet included in the calibration dataset. Moreover, SVM-DA was able to correctly predict other controlled substances (such as heroin, MDMA, amphetamine and new psychoactive substances) and non-controlled substances (such as dextromethorphan) as 'cocaine not detected'. Consequently, the specificity of SVM-DA was almost 100%. It can be concluded that probabilities below 0.10 can be used as a criteria to classify a sample as 'cocaine not detected' and probabilities above 0.62 can be used to classify a sample as 'cocaine detected'. A probability between 0.10 and 0.62 is questionable and no decision can be made. These boundaries were obtained with double cross-validation and can obviously change in the future, depending on the appearance of newly seized samples. Concerning the probabilities obtained with double cross-validation with PLS-DA, it was observed that it is difficult to draw boundaries. Compared to SVM-DA, more overlap between the probabilities of samples with cocaine and without cocaine was observed, pointing out the superior results of SVM-DA over PLS-DA. It should be pointed out that these samples (that were randomly left out with cross-validation) were critical and are necessary in the calibration dataset. In future, however, the final model (including all calibration and validation samples) will be used to analyse newly seized samples.

To the best of our knowledge, the classification of cocaine with ATR-FTIR and chemometrics in a sufficiently high number of samples with different complexity has not yet been studied. Two studies described the use of ATR-FTIR spectra in combination with chemometrics but

only to classify cocaine samples according to their chemical form (hydrochloride or base).[30,31] In our study, however, the classification model detects cocaine as such without any information about the chemical form. In the future, if enough samples are available also a classification SVM model to classify cocaine samples according to their chemical form can be established.

Besides the classification of unknown powders as 'cocaine detected' or 'cocaine not detected', our study clearly demonstrated that ATR-FTIR combined with SVMR or PLSR can be used to quantify cocaine. SVMR and PLSR determined cocaine content with RMSEPs of 6.3% and 6.5% respectively, calculated over a wide working range from 4 to 99 w%. The calibration models can be seen as stable because these RMSEP values were close to the corresponding RMSECV values. There were 13 samples with SVMR and 12 samples with PLSR that had absolute quantification errors above 15 w%. Taking into account the small number of low concentrated cocaine samples in the dataset (only 25 samples below 30%) and the highly variable matrix of the samples, it can be concluded that the number of absolute quantification errors above 15 w% was low. These quantification results have an acceptable adequacy for the routine analyses since 95% of the cocaine seizures in our lab are above 30 w% cocaine concentration as previously mentioned. Consequently, the obtained quantification results show that the SVMR and PLSR models are fit for purpose, i.e. giving a quick estimation of the drug purity. If in future more low concentrated samples are available, the models will be updated.

A good repeatability and intermediate precision of the models was also demonstrated by SDs lower than 5. When measuring the lower concentrated cocaine sample of 34.2 w%, a somewhat higher SD of 4.2 and 3.8 for SVMR and PLSR on 8 different days was calculated, respectively. This can be due to the fact that low concentrated samples are more heterogeneous. However, the repeatability was in agreement with the higher concentrated

samples, with a SD lower than 3. The above described results are acceptable compared to the validation parameters of our GC-FID reference method. Replicate analyses of the same in-house QC sample of 69.0 w% and 34.2 w% showed SDs of 1.2 and 1.4 for repeatability and SDs of 1.4 and 1.2 for intermediate precision. Based on these results, it can be concluded that ATR-FTIR was repeatable and reproducible. However, it should be pointed out that the homogenization of the powders is critical for the quantification.

The quantification of cocaine with ATR-FTIR and multivariate techniques is already described in literature. [32–34] Penido et al. used Principal Component Regression (PCR) for the quantification of self-prepared binary mixtures of cocaine base and adulterants (benzocaine, caffeine, sodium carbonate and lidocaine), with quantification errors ranging from 11 to 19%.[32] Considering the fact that they used self-prepared, binary mixtures instead of the wide variety of street samples we used, we can conclude that the models of our study showed an excellent performance with lower prediction errors (RMSEPs ranging between 5.4% and 7.1% for SVMR and PLSR).

As far as the authors know, only two studies have been based on seized samples.[30,33,34] Rodrigues et al. used PLS-discriminant analysis (PLS-DA) for the classification of samples below and above 15 w% cocaine content. They mentioned in their article that a quantification model could not be developed due to high errors of prediction (>20%).[30] In contrast, Grobério et al. managed to develop a PLS model with a working range from 35 to 99 w% based on 184 samples. The validation results of 91 samples showed a RMSEP of 3%. According to the method validation, a precision of 2% and a LOD of 12% were reported. However, 14% of the training samples and 7% of the validation samples were excluded; this is in contrast to our study where all samples were included.[33] In a second study using an extended dataset, the same research group obtained comparable results (working range from 24.2 to 99.9%, RMSEP of 3.0%, precision of 2.0 and LOD of 11.8%).[34]

Considering the results of our study, it was decided to use the SVM algorithm in an authentic routine setting. Until now, our study is the first one applying machine learning techniques for both the classification and quantification of cocaine. SVM have been demonstrated to be very valuable for many applications due to their ability to find nonlinear, global solutions and to work with high dimensional input vectors.[11,12,21,35] Moreover, they have a unique learning ability, meaning that our models are dynamic and can be further improved.[9,12,21] For example, new adulterants and cutting agents and a representative number of street samples of this new composition appearing on the illicit drug market in the future can be easily added to the calibration dataset so the model can be updated.

Compared to screening by means of colour tests and ATR-FTIR as such, ATR-FTIR combined with SVM is selective, sensitive and easy to interpret. In comparison with the existing confirmation methods (such as the conventional chromatographic techniques), the ATR-FTIR technique combined with SVM can be used as a fast analysis without an extensive sample extraction procedure. Apart from the homogenization of the powders (which is critical for the quantification), there are no consumables or chemicals required (low cost). The entire process of classification (cocaine detected or cocaine not detected) and a reliable estimation of the cocaine content can be completed within less than four minutes per sample, emphasizing its potential for high throughput (large seizures) and on-site applications. Non-experts can then be trained and the feasibility of the method can be verified in the field. In future, models can also be developed for classifying and quantifying the adulterants and/or cutting agents present in street mixtures.

CONCLUSION

It is shown that the ATR-FTIR technique combined with SVM resulted in a significant improvement of the screening test to a reliable and straightforward classification and quantification tool. Two chemometric techniques (the commonly used PLS versus the more

advanced SVM) were compared and a validation was performed to classify and quantify a wide variety of cocaine mixtures. The proposed SVM classification model resulted in a "cocaine detected" or "cocaine not detected" answer. No further interpretation of the spectra was needed. The SVM quantification model gave a reliable estimation of the cocaine content which can help us to reduce the number of samples to be analysed with GC-FID in the cases of large seizures. In summary, the major advantages of these models were their high throughput, low-cost and quick classification and quantification capabilities in comparison to those of conventional chromatographic techniques. In future, the applicability of the ATR-FTIR technique combined with SVM for the classification and quantification of other drugs could be attempted.

REFERENCES

[1] E. Plettinckx, J. Antoine, P. Blanckaert, J. C. H. van Bussel. *Belgian National Report On Drugs. Chapter 10: Drug Markets.*, Operational Directorate Public Health And Surveillance, Scientific Institute Of Public Health, Brussels, **2013**.

[2] European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). *European Drug Report. Trends and Developments*, Publications Office Of The European Union, Luxembourg, **2015**.

[3] N. Gentile, J. Broséus, P. Esseiva, F. Besacier, F. Van Durme, K. Jalava. Results from the survey on the analysis of cutting agents sent to the ENFSI DWG laboratories. **2015**.

[4] United Nations Office on Drugs and Crime. *Rapid Testing Methods of Drugs of Abuse*, United Nations, New York, **1994**.

[5] National Institute of Standards and Technology (NIST). *Color Test Reagents/kits for Preliminary Identification of Drugs of Abuse*, National Institute Of Justice, United States of America, **2000**.

[6] Y. Tsumura, T. Mitome, S. Kimoto. False positives and false negatives with a cocaine-specific field test and modification of test protocol to reduce false decision. *Forensic Sci. Int.*, **2005**, *155*, 158.

[7] P. R. Griffiths, J. A. D. Haseth. *Fourier Transform Infrared Spectrometry*, John Wiley & Sons, **2007**.

[8] J. A. Fernández Pierna, B. Lecler, J. P. Conzen, A. Niemoeller, V. Baeten, P. Dardenne. Comparison of various chemometric approaches for large near infrared spectroscopic data of feed and feed products. *Anal. Chim. Acta*, **2011**, *705*, 30.

[9] M. O'Connell, T. Howley, A. G. Ryder, M. N. Leger, M. G. Madden. Classification of a target analyte in solid mixtures using principal component analysis, support vector machines and Raman spectroscopy, (Eds: H.J. Byrne, E. Lewis, B.D. MacCraith, E. McGlynn, J.A. McLaughlin, G.D. O'Sullivan, A.G. Ryder, J.E. Walsh), **2005**, p. 340.

[10] M. G. Madden, T. Howley. A machine learning application for classification of chemical spectra, in *Appl. Innov. Intell. Syst. XVI*, Springer, **2009**, pp. 77–90.

[11] J. A. Fernández Pierna, V. Baeten, A. M. Renier, R. P. Cogdill, P. Dardenne. Combination of support vector machines (SVM) and near-infrared (NIR) imaging spectroscopy for the detection of meat and bone meal (MBM) in compound feeds. *J. Chemom.*, **2004**, *18*, 341.
[12] T. Zou, Y. Dou, H. Mi, J. Zou, Y. Ren. Support vector regression for determination of component of compound oxytetracycline powder on near-infrared spectroscopy. *Anal. Biochem.*, **2006**, *355*, 1.
[13] V. N. Vapnik. *The Nature of Statistical Learning Theory*, Springer New York : Imprint: Springer, New York, NY, **2000**.
[14] N. Cristianini, J. Shawe-Taylor. *An Introduction to Support Vector Machines: And Other Kernel-Based Learning Methods*, Cambridge University Press, New York, NY, USA, **2000**.
[15] M. Barker, W. Rayens. Partial least squares for discrimination. *J. Chemom.*, **2003**, *17*, 166.
[16] P. Geladi, B. R. Kowalski. Partial least-squares regression: A tutorial. *Anal. Chim. Acta*, **1986**, *185*, 1.
[17] International standard. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories. **2005**.
[18] R. J. Barnes, M. S. Dhanoa, S. J. Lister. Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Appl. Spectrosc.*, **1989**, *43*, 772.
[19] R. Cogdill, P. Dardenne. Least-squares support vector machines for chemometrics: an introduction and evaluation. *J. Infrared Spectrosc.*, **2004**, *12*, 93.
[20] H. Chih-Wei, C. Chih-Chung, L. Chih-Jen. A Practical Guide for Support Vector Classification. Available on http:// www.csie.ntu.edu.tw/_cjlin/papers/guide/guide.pdf. **2016**.
[21] U. Thissen, M. Pepers, B. Üstün, W. J. Melssen, L. M. C. Buydens. Comparing support vector machines to PLS for spectral regression applications. *Chemom. Intell. Lab. Syst.*, **2004**, *73*, 169.
[22] D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke. *Handbook of Chemometrics and Qualimetrics-Part A*, Elsevier Science, Amsterdam, **1997**.
[23] R. Kramer. *Chemometric Techniques for Quantitative Analysis*, CRC Press, **1998**.
[24] R. G. Brereton. *Applied Chemometrics for Scientists*, John Wiley & Sons, Chichester:, **2007**.
[25] D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte, L. Kaufman. *Chemometrics: A Textbook*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, **1988**.
[26] H. A. Martens, P. Dardenne. Validation and verification of regression in small data sets. *Chemom. Intell. Lab. Syst.*, **1998**, *44*, 99.
[27] M. I. López, M. P. Callao, I. Ruisánchez. A tutorial on the validation of qualitative methods: From the univariate to the multivariate approach. *Anal. Chim. Acta*, **2015**, *891*, 62.
[28] *ASTM E1655-05: Standard Practices for Infrared Multivariate Quantitative Analysis*, ASTM International, West Conshohocken, PA, **2012**.
[29] T. Fearn. Comparing standard deviations. *NIRnews*, **1996**, *7*, 5.
[30] N. V. S. Rodrigues, E. M. Cardoso, M. V. O. Andrade, C. L. Donnici, M. M. Sena. Analysis of Seized Cocaine Samples by using Chemometric Methods and FTIR Spectroscopy. *J. Braz. Chem. Soc.*, **2013**, *24*, 507.

[31] M. C. A. Marcelo, K. C. Mariotti, M. F. Ferrão, R. S. Ortiz. Profiling cocaine by ATR–FTIR. *Forensic Sci. Int.*, **2015**, *246*, 65.
[32] C. A. F. de O. Penido, L. Silveira, M. T. T. Pacheco. Quantification of binary mixtures of cocaine and adulterants using dispersive Raman and FT-IR spectroscopy and principal component regression. *Instrum. Sci. Technol.*, **2012**, *40*, 441.
[33] T. S. Grobério, J. J. Zacca, M. Talhavini, J. W. B. Braga. Quantification of Cocaine Hydrochloride in Seized Drug Samples by Infrared Spectroscopy and PLSR. *J. Braz. Chem. Soc.*, **2014**, *25*, 1696.
[34] T. S. Grobério, J. J. Zacca, É. D. Botelho, M. Talhavini, J. W. B. Braga. Discrimination and quantification of cocaine and adulterants in seized drug samples by infrared spectroscopy and PLSR. *Forensic Sci. Int.*, **2015**, *257*, 297.
[35] J. A. Fernández Pierna, P. Volery, R. Besson, V. Baeten, P. Dardenne. Classification of modified starches by fourier transform infrared spectroscopy using support vector machines. *J. Agric. Food Chem.*, **2005**, *53*, 6581.

TABLES

Table 1: Confusion tables SVM-DA and PLS-DA.

SVM-DA			CROSS-VALIDATION		VALIDATION	
Model	#SVs	Predicted as class	Actual class 0	Actual class 1	Actual class 0	Actual class 1
1	14	0	104	1	32	1
		1	1	306	0	70
2	15	0	115	2	20	1
		1	2	293	0	82
3	19	0	104	2	29	3
		1	4	302	0	71
4	20	0	107	2	28	0
		1	1	302	1	74
5	16	0	109	2	27	0
		1	1	300	0	76

(#SVs = number of support vectors; class 0 = samples without cocaine; class 1 = samples with cocaine)

PLS-DA			CROSS-VALIDATION		VALIDATION	
Model	#LVs	Predicted as class	Actual class 0	Actual class 1	Actual class 0	Actual class 1
1	7	0	98	10	31	2
		1	7	297	1	69
2	7	0	106	7	18	1
		1	11	288	2	82
3	11	0	97	5	27	0
		1	11	299	2	74
4	11	0	102	3	27	4
		1	6	301	2	70
5	8	0	103	7	25	3
		1	7	295	2	73

(#LVs = number of latent variables; class 0 = samples without cocaine; class 1 = samples with cocaine)

Table 2: Summary of the validation results of the classification models.

Performance parameters	SVM-DA	PLS-DA
True positives	378	378
True negatives	137	137
False positives	1	9
False negatives	5	10
Sensitivity (%)	99	97
Specificity (%)	99	94
Efficiency (%)	99	96

Table 3: Results of proficiency samples for SVM-DA.

Proficiency test	**Target substance**	**Cocaine presence**	**SVM-DA probabilities**
2015/1	Cocaine	yes	1.00000
	Cocaine	yes	1.00000
	Cocaine	yes	1.00000
2015/2	Cocaine	yes	1.00000
	Nimetazepam	no	0.00247
	Ketamine	no	0.00198
	Methamphetamine	no	0.00073
2014/1	1-(3-Chlorophenyl)piperazine	no	0.00501
	Methamphetamine	no	0.00001
	Ketamine	no	0.00593
	Cocaine	yes	1.00000
2014/2	Amphetamine	no	0.00020
	Cocaine	yes	1.00000
	2-Pyrrolidinovalerophenone	no	0.03268

Table 4: Summary of the results of the quantification models.

SVMR		CALIBRATION			CROSS-VALIDATION			VALIDATION		
Model	#SVs	RMSEC	R^2c	bias	RMSECV	R^2cv	bias	RMSEP	R^2v	bias
1	200	4.0	0.97	0.204	6.5	0.92	0.273	6.4	0.92	0.423
2	193	4.7	0.96	0.209	6.3	0.92	0.410	6.7	0.90	1.109
3	294	5.0	0.95	0.437	6.7	0.91	0.504	5.5	0.94	0.224
4	204	4.2	0.96	0.308	6.2	0.92	0.163	5.4	0.95	0.422
5	290	4.1	0.97	0.212	6.3	0.93	-0.150	7.1	0.90	-0.949

(R^2c = determination coefficient of calibration, R^2cv = determination coefficient of cross-validation and R^2v = determination coefficient of validation; optimal parameters: cost = 0.1 for models 1, 4 and 5; cost = 0.032 for models 2 and 3; epsilon = 0.1 for models 1, 2 and 4; epsilon = 0.01 for models 3 and 5)

PLSR		CALIBRATION			CROSS-VALIDATION			VALIDATION		
Model	#LVs	RMSEC	R^2c	bias	RMSECV	R^2cv	bias	RMSEP	R^2v	bias
1	12	5.2	0.94	-0.002	6.5	0.92	0.035	6.9	0.91	0.578
2	13	4.9	0.95	-0.002	6.6	0.91	-0.016	7.1	0.89	1.447
3	12	5.4	0.94	0.009	7.0	0.90	0.003	5.7	0.93	-0.037
4	13	5.2	0.94	0.001	6.7	0.90	-0.090	5.6	0.95	-0.629
5	13	5.1	0.95	0.020	6.9	0.91	-0.101	6.9	0.89	-1.091

(R^2c = determination coefficient of calibration, R^2cv = determination coefficient of cross-validation and R^2v = determination coefficient of validation)

Table 5: Results of repeatability and intermediate precision for SVMR and PLSR.

SVMR					
Cocaine concentration (%)	**# measurements**	**Repeatability**		**Intermediate precision**	
		Mean (%)	SD	Mean (%)	SD
99	8	95.1	1.0	95.1	1.6
68	8	71.1	2.6	74.2	1.9
34	8	28.3	2.0	32.2	4.2

(optimal parameters for final model with 378 SNV-processed spectra: 237 SVs, cost and epsilon = 0.1)

PLSR					
Cocaine concentration (%)	**# measurements**	**Repeatability**		**Intermediate precision**	
		Mean (%)	SD	Mean (%)	SD
99	8	89.7	0.5	90.1	1.6
68	8	74.6	2.2	76.0	1.7
34	8	29.2	2.7	33.7	3.8

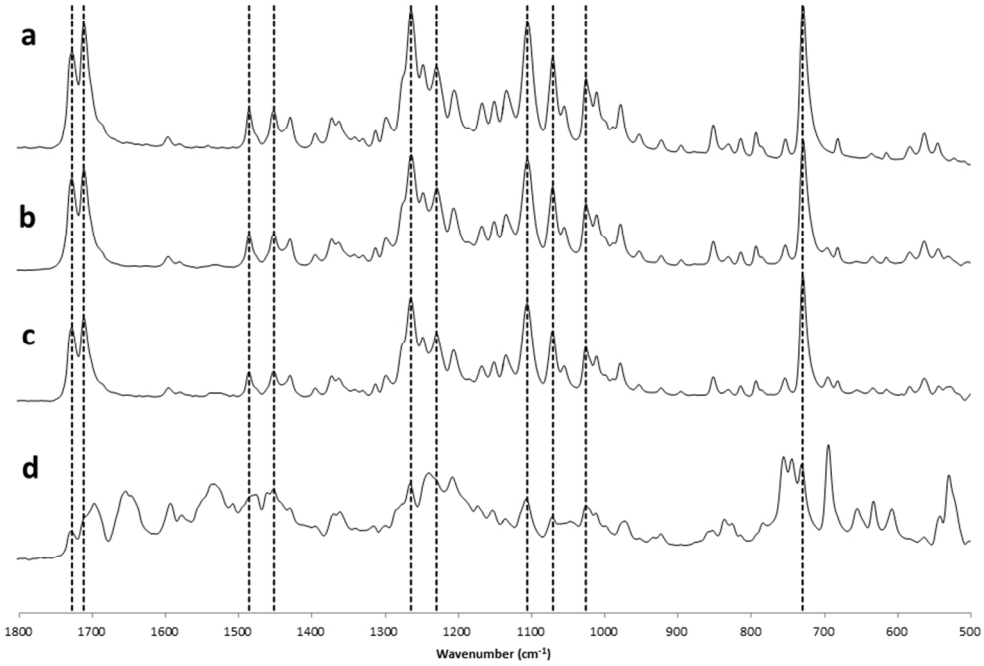
(optimal parameters for final model with 378 SNV-processed spectra: 12 LVs)

Table 6: Results of reproducibility for SVMR and PLSR.

Proficiency test	Mean laboratories (% cocaine concentration)	SD laboratories	SVMR (% cocaine concentration)	z-score*	PLSR (% cocaine concentration)	z-score*
2015/1	41.7	1.4	41.6	-0.1	36.7	-3.6
	64.5	1.7	63.1	-0.8	64.7	0.1
	41.5	1.4	43.5	1.4	39.0	-1.8
2015/2	87.1	6.5	83.4	-0.6	81.5	-0.9
2014/1	87.6	6.4	90.2	0.4	88.8	0.2
2014/2	72.4	4.6	75.6	0.7	75.2	0.6

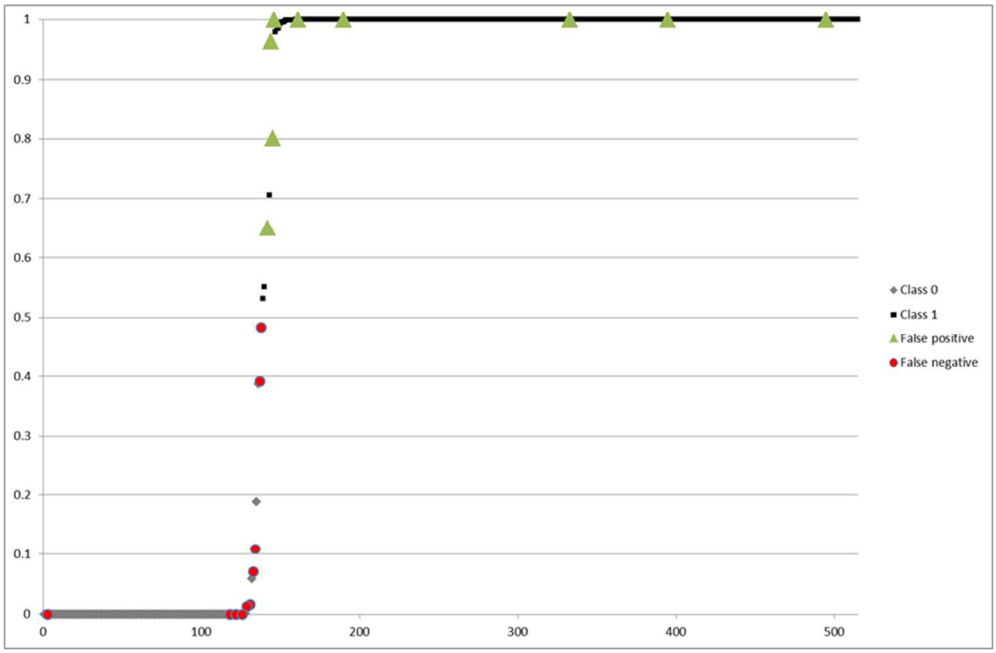
(*z-score = (predicted value – mean laboratories)/SD laboratories; excluding outliers (Grubbs test))

Figure 1. ATR-FTIR spectra of a pure reference cocaine (a) and three cocaine street samples (b-d) in the fingerprint region (1800-500 cm^{-1}).



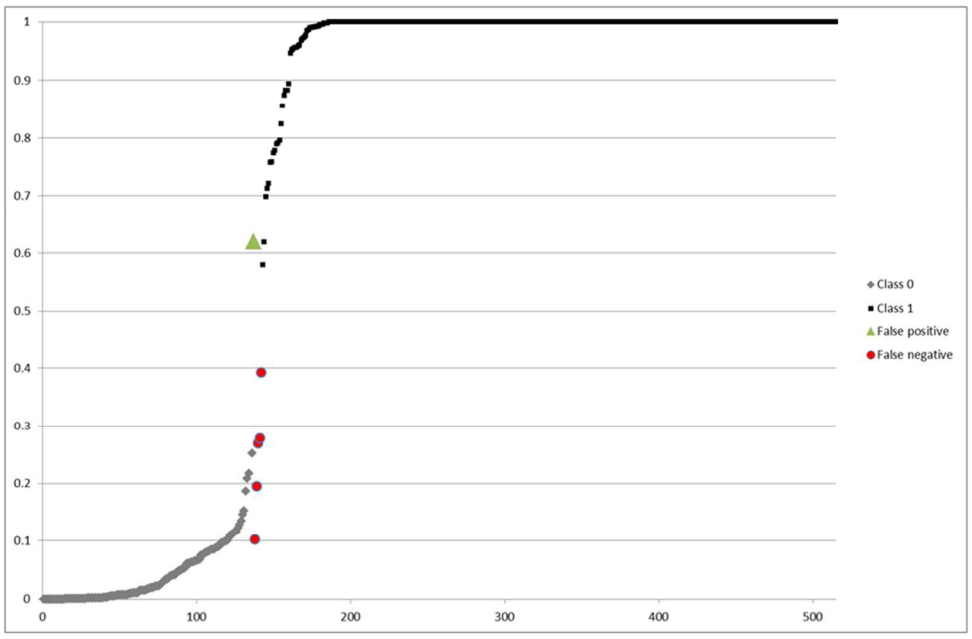
The main vibrational bands of cocaine hydrochloride are represented by black dashed lines.
Figure 1
254x190mm (96 x 96 DPI)

Figure 2. Probabilities plot PLS-DA with double cross-validation (5 random subsets).



False negatives represented by red circles and false positives by green triangles; class 0 = without cocaine represented by a grey rhomb; class 1 = with cocaine represented by a black square)
Figure 2
254x190mm (96 x 96 DPI)

Figure 3. Probabilities plot SVM-DA with double cross-validation (5 random subsets).

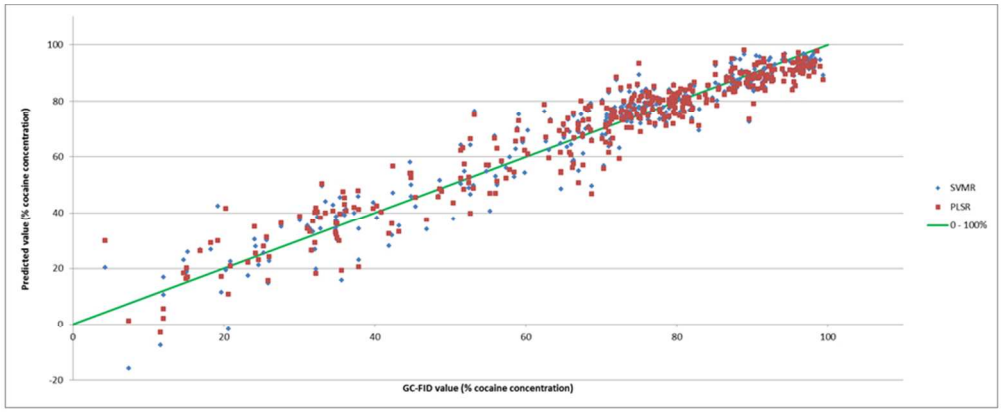


False negatives represented by red circles and a false positive by a green triangle; class 0 = without cocaine represented by a grey rhomb; class 1 = with cocaine represented by a black square

Figure 3

254x190mm (96 x 96 DPI)

Figure 4. Results of the validation with cocaine concentrations measured with GC-FID and predicted by SVMR (blue) or PLSR (red).



Predictions SVMR represented by blue rhombs; Predictions PLSR represented by red squares; range 0%-100 w% cocaine represented by green line
Figure 4
254x190mm (96 x 96 DPI)