Deconinck Eric (Orcid ID: 0000-0002-2374-1093)

Discrimination of legal and illegal *Cannabis spp.* according to European legislation using near infrared spectroscopy and chemometrics

Duchateau Céline[1,2], Kauffmann Jean-Michel[1], Canfyn Michaël[2], Stévigny Caroline[1], De Braekeleer Kris[1a] and Deconinck Eric[1,2a*]

[1]ULB- Faculty of Pharmacy – RD3 – Pharmacognosy, Bioanalysis and Drug Discovery Unit – Bld Triomphe, Campus Plaine, CP 205/6 – B – 1050 Brussels – BELGIUM

[2]Sciensano – Scientific Direction Physical and Chemical Health Risks - Medicines and health products
Rue Juliette Wytsmanstraat, 14 – 1050 Brussels – BELGIUM

[a]Equally contributing project leaders
[*]Corresponding author : Eric.Deconinck@sciensano.be

Keywords :

Cannabis

Near infrared spectroscopy

Chemometrics

Handheld

European legislation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/dta.2865

Abstract

Aerial parts containing cannabidiol can be purchased in a legal way but cannabis used as recreational drug is illegal in most European countries. Δ9-tetrahydrocannabinol is one of the main cannabinoids responsible for the psychotropic effect. European Union countries and Switzerland authorize a concentration of THC of 0.2% and 1.0% w/w, respectively, for smoking products and industrial hemp. Public health inspectors and law enforcement officers need to check the legality of samples. Therefore there is a need for innovative approaches, allowing quality control of these products in an easy way and preferably on site. In many countries, cultivation of industrial hemp is permitted if the THC content does not exceed 0.2% w/w. A portable equipment could be a useful measuring tool for farmers to check for the THC content at regular time. In this work, 189 samples were analysed with a benchtop and a handheld NIR device in order to create two classification methods according to European and Swiss laws. All samples were also analysed by GC-FID to determine their THC concentration. Supervised analysis was applied in order to establish the best model. For the first classification, the accuracy was 91 % for the test set with the benchtop data and 93 % for the test set with the handheld data. For the second classification, the accuracies were respectively 91 % and 95%. The obtained models, hyphenating spectroscopic techniques and chemometrics, enable to discriminate legal and illegal cannabis samples according to European and Swiss laws.

1. Introduction

The cannabis plant is known for diverses applications e.g. in the industry but more importantly for recreational and medical uses. Cannabis and its derived products are available on the Belgian market via both illegal and legal ways. Flowers containing cannabidiol (CBD) as well as medicinal cannabis can both be purchased in a legal way, but cannabis used as recreational drug is illegal in most European countries. Indeed, cannabis is the most widely consumed illicit drug in Europe and products contain generally a concentration of Δ9-tetrahydrocannabinol (THC) up to 15 % (w/w) [1,2]. Different popular classifications for cannabis exist considering three different species, different subspecies or only a single species : *Cannabis sativa* [3]. The plant is complex from a chemical point of view because it contains more than 500 different compounds as flavonoids, monoterpenes, sesquiterpenes, steroids and cannabinoids. The latter are considered as the main active constituents of the plant and are produced by the glandular trichomes localized in the aerial part of the plant female [4,5]. The concentration of these compounds depends on tissue type, age, variety, growth conditions, harvest time and storage conditions [6]. The THC content in cannabis plants may also vary according to genetic and environmental factors [7]. The most important cannabinoids are THC, cannabinol (CBN) and CBD, in their neutral form. CBN is the primary product of THC degradation. In the fresh vegetal material, the THC carboxylic acid precursor represents 90 % of the total THC (ie. acid and neutral forms of THC). The acid form is decarboxylated in the neutral form. Oxidative degradation of THC results in the formation of CBN when the vegetal material is exposed to heat [8–10]. THC is one of the main cannabinoids responsible for the psychotropic effect [11]. Most European Union countries authorize for the herbal products for smoking and for the agricultural hemp a concentration of THC of 0.2 % (w/w), while Switzerland allows 1.0 % (w/w) [1,12]. In Belgium, laws authorize a maximal concentration of 0.2 % (w/w) in THC [1,13].

CBD is not psychoactive but shows, although not scientifically proven, medicinal properties and modulates the THC metabolism by blocking its conversion to the more psychoactive 11-hydroxy-THC [11]. The medical properties attributed to this non-psychoactive molecule are described in the literature and are e.g. anti-anxiety actions, anti-psychotic effects, anti-inflammatory and immunomodulatory effects [11,14]. A lot of products called “CBD-based products” are selled as “pot pourri” [8] generally in order to circumvent the legislation. It has to be said that THC and CBD are neutral forms and do not exist in significant concentrations in the plant as such. Due to the upcoming availability of these CBD based products, seizures of cannabis samples have become more difficult due to the impossibility to discriminate cannabis flowers with a THC concentration higher than 0.2 % (w/w) by inspectors or law enforcement officers on site. For the moment the analysis of the samples should be performed in an accredited laboratory in order to check for the

THC concentration. The conventional methods used for this purpose are GC-MS, GC-FID and HPLC-DAD [9,15]. All these methods are slow and expensive because they require a complete sample preparation. Furthermore, they require trained personnel, are not ecological and are not suited for onsite analysis [15,16]. Moreover, the time between the seizure of the products and the laboratory results is long which can cause a blockage of legal products, with an economical loss for the distributor. Therefore, there is a need for an innovative methodological approach which allows quality control of these products on site. Near infrared spectroscopy (NIRS) combined with chemometrics has great potential in the analysis of plant natural products as reviewed by D. Cozzolino. Moisture, volatile substances and chemical compounds in herbal products can be analysed by NIRS because absorbd in the infrared region of the electromagnetic spectrum [17]. C. Mees *et al*. have described a relevant study using NIRS to discriminate leaves, in a powder format, of various coffee plants [18]. Cannabis aerial parts have already been analysed by NIRS and classifications have already been made. B. Borille *et al.* have used NIRS for the growth stage classification of cannabis cultivated in a greenhouse [19] and C. Sanchez-Carnerero *et al* have exploited the potential of NIRS to estimate the content of cannabinoids. They also compared the results obtained with Fourier-Transform NIRS and dispersive NIRS [4]. This technique is easy to use, rapid, relatively cheap, non-destructive and needs no or limited sample preparation (green technique).

Considerable attention has been given to the miniaturization of spectroscopic devices for infield measurements. Handheld devices offer interesting possibilities by allowing insitu analysis by reduction of size of the device. These properties allow a reduction of cost and easy transport. Moreover, they offer accuracy of measurement and high performance reliability [20]. Miniaturization of NIRS devices allows the operator to perform the analysis on-site [21]. As example, R. Risolutti *et al* have developed a screening test for the real time detection of cannabinoids in hemp flour by a miniaturized analytical platform based on the MicroNIR spectrometer monitoring of cannabinoids in hemp flours [22].Spectroscopic methods, however, produce highly informative spectra, containing a lot of data difficult to interpret. Chemometric tools allow the analysis of large and complex datasets and are better to be applied for the extraction of the information of interest [23].

Cultivation of industrial hemp is only permitted for allowed varieties and the cultivation is performed according to standard agricultural practices described in the EU execution rules n°1306/2013 [24]. In many countries, a list of approved cultivars exists [1,25]. The Federal Public Service of Internal Affairs collects and analyses the hemp for the farmers in Belgium. If the THC content exceeds 0.2 % (w/w), the products must be destroyed, resulting in an economic loss for the farmer. Due to the environmental influence on the THC content [7], it could be interesting to

develop an easy tool that can be used by the farmer to check for the THC content at regular time intervals. This will allow the farmer to adapt the harvest time and to prevent economic loss. In literature some applications of NIRS for the screening of cannabinoids have already be found. These applications focus on the various chemotypes and the growth stage of the cannabis plant[19,26].

This study describes the use of NIRS, combined with chemometrics to develop an innovating classification model of cannabis and derived products according to European Union and Swiss laws. As secondary goal the applicability and the performance of a handheld device was also compared to the performance of a benchtop NIRS equipment.

For this purpose, spectra of 189 samples were measured with both a benchtop and a handheld device. All samples were analysed before, using GC-FID to determine their THC concentration. The data was subsequently explored for the discriminatory power of the NIR spectra, using unsupervised clustering techniques, i.e. Principal Component Analysis (PCA) and Hierarchical Clustering Analysis (HCA). Finally the spectral data were modelled as a function of the THC concentration using supervised classification techniques i.e. Partial Least Square-Discriminant Analysis (PLS-DA), *k*-Nearest Neighbours (k-NN) and Soft Independent Modelling of Class Analogy (SIMCA).

2. Materials and methods

2.1 Plant materials

For this study, 189 cannabis samples (aerial parts) were investigated. The samples were seized by the Belgian authorities on festivals and in night life settings. Herbal products for smoking were seized by Belgian authorities or voluntarily donated by distributors. The samples mentioned before were already dried by the manufacturers. Agricultural hemp samples were obtained from Belgian farmers. The samples were dried as quickly as possible and in all cases within 48 hours after delivery to prevent THC degradation to CBN. The plants were collected in an oven tray, and placed in an oven at 68 °C for a period of minimum 24 hrs according to the standard method prescribed by the European Directives [27]. Afterwards the dried samples were stripped of stems and seeds to retain only the aerial part.

2.2 Reagents and chemicals

Ethanol and dichloromethane were purchased from Biosolve (Valkenswaard, The Netherlands). Methyl-arachidate is a fatty acid methyl ester with 21 carbons, like the cannabinoids of interest, and was used as internal standard. It was purchased from Sigma-Aldrich (Darmstadt, Germany)

and its purity was ≥ 99.5 %. The reference standard used in this study, cannabinol, was from Lipomed (Arlesheim, Switzerland) and had a purity ≥ 98.5 %.

2.3 Standard

Five standards solutions containing 2-400 µg/mL cannabinol and 100 µg/mL of the internal standard were prepared in ethanol. The stock solution of the internal standard was prepared in dichloromethane.

2.4 Sample preparation

Dried samples were crushed by hand in a plastic bag to reduce the size of the flowers without using any sophisticated tools.

For NIRS analysis, each sample was divided into 3 glass vials. The height of the vegetal material in the vial had been approximatively 1.5-2 cm in order to render the analysis in diffuse reflectance mode efficient. The reflectance mode is designed to direct the diffusely reflected beam from the sample onto the detector [28].

For GC-FID analysis, according to the sample type, 1 g of agricultural hemp, 400 mg for CBD flowers or other CBD based smoking products and 80 mg for recreational cannabis were weighted in a brown glass volumetric flask of 50 mL. To this flask, 1 mL of 5 mg/mL internal standard solution was added and the sample was brought to volume with the extraction solvent (dichloromethane). The extraction was performed by sonication during 30 min. The supernatant was removed after 5 min.

2.5 GC-FID analysis

An officially validated method to quantify THC, currently applied by Sciensano (Brussels, Belgium), was used to analyze samples. THC content in the cannabis samples was determined by GC-FID using CBN as reference standard. The content of THC was determined using a response factor to convert the concentration of CBN to THC [29].

GC-FID analysis was carried out using an Agilent technologies 5973 system with a FID detector. The instrument was equipped with a capillary column (Agilent, DB-5ms, 30 m length, 0.25 mm internal diameter, module LTM) impregnated with an 0.25 µm phenyl-methyl-siloxane film. The injector temperature was 225°C with an injection volume of 2 µL in split mode (1/20). The carrier gas (He) flow rate was set at 1.5 mL/min. The temperature gradient started at 270°C which was held for one min. Then, the temperature was linearly increased at a rate of 10°C/min until 320°C. The FID detector temperature was set at 300°C. The device was controlled by MSD Chem Station software (E02.02.1431). All determinations were performed in duplicate.

2.6 Near infrared spectroscopic analysis

Each vial was analyzed 3 times, resulting in nine spectra per sample.

2.6.1 Fourrier-Transform NIRS : benchtop device

All samples were scanned with the Frontier MIR/NIR Spectrometer™ (PerkinElmer™, USA). Spectra were measured in reflectance mode using the Near Infrared Reflectance Accessory with a resolution of 8 cm^{-1} in the region of 10000 – 4000 cm^{-1} and consisted of an average of 16 scans. Background spectra were collected with a diffuse reflector provided by PerkinElmer™. The background was recorded between each sample (replicate). The spectra were recorded by subtracting the backgroung spectrumA correction of the background was applied using an arithmetic functions.

2.6.2 Dispersive NIRS : handheld device

The same samples were also scanned by the dispersive handheld device MicroPHAZIR™ RX Analyzer (Thermo Scientific™, USA). Measurements were performed in the reflectance diffuse mode with a resolution of 11 nm in the region of 6200-4000 cm^{-1}. An average of 5 scans was reported. The background was recorded automatically thanks to an inside sphere which enables to take a separate background.

2.7 Chemometric methods

2.7.1 Data preprocessing

Various factors such as the sample particle size can interfere with the information of interest in the NIR spectra. Different preprocessing methods can be used in order to remove these interferences. In this study, two pretreatments were retained after trying several pretreatments and different combinations. Standard Normal Variate (SNV) was used to remove the interferences due to pathlength effects. Raw spectra obtained with the benchtop were preprocessed by SNV followed by the second derivative. The deconvolution method of Savitzky-Golay used a second order polynomial and a window size of 17. Raw spectra obtained with the handheld device were processed by a first derivative followed by SNV. The deconvolution method of Savitzky-Golay used a third polynomial and a window size of 7 [30].

The benchtop NIR showed no information in the region from 10,000-6,000 cm^{-1}. Therefore the same spectral region as for the portable (i.e. 6,000 to 4,000 cm^{-1}) was taken into account.

2.7.2 Exploratory data analysis

2.7.2.1 PCA

PCA was used for exploratory (unsupervised) data analysis to reveal cluster tendency of the samples. It permitted to project the original data space onto a smaller latent variable space retaining the maximal amount of variance in the spectral data. The latent variables were defined as linear combinations of the manifest variables (in this study the intensity of the signal at a certain wavelength) describing the highest variance in the data (PC1) or the highest remaining variance (PC2, PC3,…). The aim of applying PCA was to visualize the spectra in two or three dimensional plots, defined by two or three PCs. As said before PCA was applied as exploratory analysis, i.e. it was checked if PCA on the spectral data could differentiate between the different types of samples defined based on the results obtained with GC-FID [31,32]. This allows to ensure that data can be modelled using supervised techniques and that good predictive results are not based on co-incidence or due to the modelling of noise in the data.

2.7.2.2 HCA

Hierarchical clustering allowed to cluster samples into groups, based on the similarity between the spectra of the different samples. Various distance metrics and clustering methods were tested to compare clustering performances. In this study, the Euclidean metric was chosen to measure the (di)similarity between samples using Wards algorithm (inner squared distance) [31].

2.7.3 Supervised analysis

2.7.3.1 Selection of a training and test set

The Duplex algorithm was applied to select training and test sets. The Euclidean distance was used as dissimilarity criterion. The two samples with highest Euclidean distance between them were selected and placed in a first set. The next two samples with the highest distance were placed in a second set (test set). This procedure continued until a predefined number of samples was selected in the test set. The first set and the remaining not selected samples composed the training set, the second set the test set. In this study, 30 % of the samples were used for the test set [31,33].

2.7.3.2 SIMCA

SIMCA is a class modeling technique. A SIMCA model consists of a collection of PCA models, one for each class. The number of PCs of each class of the training set was selected by venetian blinds cross validation. The samples of the test set were projected in each subspace and were compared to it in order to calculate its distance from each class. Two distances were taken into account: the normalised Q and normalized Hotelling T^2 values. Confidence limits were set at 95 %. SIMCA attributed the samples of the test set using two different approaches: (1) each sample was assigned to one of the closest classes, (2) or assigned to none or one of the classes, if the threshold

was based on maximizing class specificity and sensitivity [31,34,35]. Statistical parameters used to evaluate the model were : precision, sensitivity, specificity, accuracy, error/non-error rate and ratio of not assigned samples.

2.7.3.3 *k*-NN

k-NN is a classification procedure assigning new samples to a class. Distances were calculated between the test samples and the samples of the training set. The Euclidian distance was used for this purpose. The *k* nearest samples of the training set were selected and the test samples were classified in the class to which the majority of the *k* nearest samples belonged. In this study, a 10–fold cross validation procedure was performed to select the optimal *k* value. The k value giving the highest correct classification rate (ccr) was selected as the optimal one [31,35].

2.7.3.4 PLS-DA

PLS is a supervised technique that shows analogy with PCA. PLS is also a projection technique in which latent variables are defined as linear combinations of manifest variables. These combinations reflect the highest (PLS factor 1) or highest remaining (PLS factor 2, 3 ,…) co-variance between the data and a response. PLS is a regression techniques used to model a continuous response like concentration. PLS-DA is a variant combining PLS and discriminant analysis, allowing PLS to be used as a classification technique. The optimal number of latent variables (PLS-factors) was selected by venetian blinds cross-validation. Two approaches were applied for classification of the samples : (1) each sample of the test set was assigned to one of the classes based on the probabilities calculated for each sample to belong to the different classes, (2) or tresholds were calculated by the Bayes theorem. In the latter case, samples of the test set were attributed to none or one of the classes. [35–37]. The same statistical parameters were calculated as for SIMCA.

2.8 Software

Data treatments were performed by using Matlab version R2018b (The Mathworks™, Natick, USA). The algorithm of PCA, HCA, Duplexx, *k*-NN, SNV and Savitzky-Golay were part of the ChemoAC toolbox (Freeware™, ChemoAC Consortium, Brussel, Belgium, version 4.1). The SIMCA and PLS-DA toolbox were downloaded from the Freeware Classification toolbox, version 5.3. [37].

3. Results and discussion

NIR spectra were collected using the benchtop and the handheld device. The spectral data of the fingerprint region (6,000-4,000 cm^{-1}) was subjected to chemometrics. Typical NIR absorption

bands appear in all the spectra and from both instruments. Raw spectra obtained with both devices are shown in Fig. 1 A and 1B. Different bands are shown in these raw spectra : the combination of stretching and bending vibrations of CH_2 and CH_3 functional groups (1), the first overtone and combinations bands of OH-group present in water (2), absorption of protein (3), aromatic hydrocarbons (4), fibers (5) and second overtone of stretching vibrations (bands assigned with lipids) (6). For bands (6), one band is not visible in the dispersive raw spectra because the region is less extended than for the benchtop. Preprocessed spectra obtained for the benchtop and for the handheld devices are shown in Fig. 2 and Fig. 3, respectively. Absorbance values for these both figures are different because pretreatment used for each one is not the same. Samples are indicated with different colours depending on the THC concentration: samples with a concentration less than 0.2 % (w/w) (green), samples with a concentration between 0.2 % (w/w) and 1 % (w/w) (blue) and higher than 1 % (w/w) (red). For the benchtop, green and red spectra were clearly separated, while only some small spectral differences were observed between the blue and green spectra. For the handheld device, red spectra were clearly separated from the green and blue spectra. The two latter were mixed between them and only few differences were observed between the blue and the green spectra.

3.1 Unsupervised analysis

3.1.1 Benchtop device

The score plots obtained with different data pretreatment procedures were compared and the best score plot was obtained with SNV followed by the second derivative (Fig. 4). The three first principal components (PCs) explain more than 97 % of the total variance (PC1= 87.0 %, PC2= 8.3 % and PC3=2.3 %). The samples with a THC content higher than 1 % (w/w) and lower than 0.2 % (w/w) were clearly separated. The samples with 0.2-1.0 % (w/w) THC were situated between these two groups and were partially mixed with those of less than 0.2 % (w/w) THC.

The HCA dendrogram (Fig. 5) showed three major clusters: the first cluster corresponded to samples with a concentration higher than 1.0 % (w/w) THC, the second cluster to samples with a concentration between 0.2 % (w/w) and 1.0 % (w/w) THC and the third cluster were samples with a concentration less than 0.2 % (w/w). Only few samples with a concentration between 0.2 % (w/w) THC and 1.0 % (w/w) THC were situated in the third cluster. Four samples were considered as outliers because the HCA dendrogram showed high dissimilarity for these samples towards the rest of the sample set. These samples were removed for subsequent analysis.

3.1.2 Handheld device

The best score plot for PCA was obtained with the first derivative followed by SNV (Fig. 6). The first three PCs explained more than 99 % of the variance (PC1= 97.6 %, PC2= 1.5 % and PC3=0.4 %). The clustering was comparable to the score plots obtained with the benchtop spectra, meaning that an acceptable clustering of the samples of the three concentration levels was obtained. The HCA dendrogram (Fig. 7) provided two major clusters: the first cluster were samples with a concentration higher than 1.0 % (w/w) THC and the second cluster were samples with concentration less than 1.0 % (w/w). The discrimination between samples with a concentration between 0.2-1.0 % (w/w) THC and samples with a concentration less than 0.2 % (w/w) THC could not be observed.

3.2. Supervised analysis

3.2.1. Binary classification based on European Union legislation (< 0.2 % (w/w) THC)

The classification of the cannabis samples was performed by a binary classification model. Classes were defined by the concentration of THC obtained with GC-FID. Class 1 represented the samples with a concentration higher than 0.2 % (w/w) THC and class 2 represented the samples with a concentration lower than 0.2 % (w/w) THC.

3.2.1.1. Benchtop device

The performances of *k*-NN, SIMCA and PLS-DA were compared and the best classification model was selected. The results are given in Table 1. The Duplex algorithm was used to select 56 samples for the test set (external validation), the 129 remaining samples formed the training set (cross-validation). The best classification was obtained with SIMCA. For each class, one PC was selected using venetian blinds cross-validation. Each sample of the test set was attributed to one of the two classes. The accuracy was 92 % (119/129) for the training set and 91 % (51/56) for the test set. For the test set, no samples of the second class were classified in the first class and 5 samples of the first class were classified in the second class. For the training set, one sample of the second class was classified in the first class and nine samples of the first class were assigned to the second class. The samples misclassified do not have a concentration at the limit of the two classes, so these errors have to be attributed to modelling errors, for which no logical explanation could be found. The studied classification parameters such as the error rate, non error rate, precision, class specificity and sensitivity showed a satisfying classification performance (Table 2). The confusion matrices are given in the supplementary data.

3.2.1.2. Handheld device

The same procedure was followed as for the benchtop data. Duplex was used to select 56 samples for the test set (external validation), the 133 remaining samples formed the training set (cross-

validation). The best classification was, as for benchtop, obtained with SIMCA. One and two PCs were selected for class 1 and class 2, respectively, using venetian blinds cross-validation. Each sample of the test set was attributed to one of the two classes. The accuracy was 97 % (129/133) for the training set and 93 % (52/56) for the test set. For the test set, no samples of the second class were classified in the first class and four samples of the first class were classified in the second class. For the training set, no samples of the second class were classified in the first class and four samples of the first class were assigned to the second class. Also here the misclassified samples do not have a THC concentration at the limit of the class border, so these misclassifications can only be attributed to errors of the obtained model. The classification parameters of the SIMCA model are collected in Table 2. The confusion matrices are given in the supplementary data.

3.2.2. Tertiary classification based on the European Union and Swiss legislation

The classification of the cannabis samples was repeated using a tertiary classification model. Classes were defined by the concentration of THC obtained with GC-FID. Class 1 represented the samples with a concentration > 1.0 % (w/w) THC , class 2 represented the samples with a concentration between 0.2 – 1.0 % (w/w) THC and class 3 represented the samples with a concentration < 0.2 % (w/w) THC.

3.2.2.1. Benchtop device

As for the binary classification, the performances of *k*-NN, SIMCA and PLS-DA were compared and the best classification model was selected. The results are given in Table 1. The Duplex algorithm was used to select 56 samples for the test set (external validation), the 129 remaining samples formed the training set (cross-validation). The best classification models were obtained with PLS-DA. As mentioned before two approaches were followed i.e. the classical one assigning each sample to a class and the Bayes theorem. For the classical approach five PLS factors were chosen for the optimal model. Although a ccr (correct classification rate) or cross validation of 92% (119/129) was obtained, only 47 of the 56 samples (84%) in the test set were correctly classified. Therefore it was chosen to select the PLS-DA model obtained with the Bayes theorem as optimal model. The reasoning was that, although a certain percentage of the samples can be unclassified, it is, in the context of this study, better to have a higher accuracy, to avoid misclassifications. The unclassified samples based on this model should additionally be sent to an accredited laboratory in order to confirm their legality. For the optimal model four PLS factors were selected using venetian blinds cross-validation. A total of 9 % (12/129) of the samples for the training set and 20 % (11/56) for the test set were unasigned, respectively. The accuracy was 96 % (112/117) for the training set and 91 % (41/45) for the test set. For the test set, three samples of the third

class were classified in the second class and one sample of the first class was classified in the third class. For the training set, one sample of the third class was classified in the first class, one sample of the first class was assigned to the second class, one sample of the third class was assigned to the second class and one sample of the first class was assigned to the third class. No explanation for these misclassifications could be found. The classification parameters of the PLS-DA model are collected in Table 3. The confusion matrices are given in the supplementary data.

3.2.2.2. Dispersive NIR handheld device

The same performances were compared and the best classification model was selected. The results are given in Table 1. The Duplex algorithm was used to select 56 samples for the test set (external validation), the 133 remaining samples formed the training set (cross-validation). The best classification models were obtained with PLS-DA. Also here the PLS-DA model following the classical approach showed a promising ccr for cross validation of 99% (128/129), although for the test set only 49 of the 56 samples were correctly classified. For the same reasons as for the benchtop the model obtained applying Bayes theorem was preferred. Three PLS factors were selected using venetian blinds cross-validation and 12 % (16/133) of the samples of the training set and 21 % (12/56) of the test set were not assigned to a class, respectively. The accuracy of the model was 96 % (113/117) for the training set and 95 % (42/44) for the test set. For the test set, two samples of the second class were classified in the third class. For the training set, one sample of the first class was classified in the second class and one sample of the first class was assigned to the third class. Also here, no explanation for these misclassifications could be found. The classification parameters for the model are collected in Table 3. Confusion matrices are given in the supplementary data.

4. Conclusion

A methodology for classification of different cannabis samples has been developed using two near-infrared spectroscopic devices: a benchtop and a handheld one. The samples were classified according to European Union and Swiss legislations, based on the THC concentration obtained with GC-FID. The advantage of the NIRS is that the samples are measured with almost no pretreatment and that the analysis is fast, ecological and has possibilities for on-site analysis. The presented approach would be helpful for inspecting authorities and farmers. In this study, 189 aerial parts of cannabis samples were subjected to analysis. Based on the results obtained from exploratory analysis, supervised chemometric tools were applied to reveal differences between the samples. The classification was performed according to the concentration of THC determined by GC-FID. Unsupervised analysis with HCA and PCA revealed some clusters. For the benchtop, three clusters were observed according to the tertiary classification. For the handheld device, two

clusters were observed according to the binary classification. The supervised analysis permitted to create models allowing one to classify new cannabis samples according to the European and Swiss legislation using a benchtop and a handheld devices. The SIMCA models obtained for the binary classification of the test set spectra had an accuracy of 91 % and 93 % for the benchtop and the handheld devices, respectively. For the tertiary classification the best models were obtained with PLS-DA, applying Bayes theorem and this for both devices. These models showed an accuracy of 91 % and 95 % for the test set obtained with the benchtop and the handheld devices, respectively. The rate of unclassified samples for the two devices was 20 % and 21 %. It was chosen to favor a higher accuracy of the models and thus allowing a certain percentage of the samples to be unclassified, since a higher risk of misclassification could lead to unnecessary legal procedure. Unclassified samples should then follow the classical route and be send to an accredited laboratory in order to check their legality. One of the main reasons of the lesser performance of the tertiary models is probably the fact that the intermediary class was only represented by 11 samples. The models in this study could therefore be improved by increasing the number of samples, especially for the intermediary class.

When comparing the results of the models obtained with the spectra recorded with the handheld device and the ones obtained with the benchtop it can be noted that no significant differences between the performance of the models could be observed. This mean that the handheld device, evaluated in this study, gave similar results as the benchtop device and therefore can be used for the screening of cannabis samples both by inspectors and law enforcement officers as well as farmers.

5. Acknowledgements

The authors want to acknowledge Dr. Marie Vandeput for critical reading of the manuscript ; the authors were also grateful to Corenthin Mees for his helpful suggestions. The authors would like to thank the department of Pharmacy and Analytical Chemistry of ULiège for the providing of the dispersive NIR handheld device.

6. Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or
not-for-profit sectors.

7. References

[1] European Monitoring Centre for Drugs and Drug Adiction. *Cannabis Legislation in Europe*, **2017**.

[2] S. P. F. Belge. "http://justice-belgium.be," Available at: https://justice.belgium.be/fr/themes_et_dossiers/securite_et_criminalite/drogues/cannabis, **n.d.**

[3] S. Farag, O. Kayser. *Chapter 1 - The Cannabis Plant: Botanical Aspects*, Elsevier Inc., **2017**.

[4] C. Sánchez-Carnerero Callado, N. Núñez-Sánchez, S. Casano, C. Ferreiro-Vera. The potential of near infrared spectroscopy to estimate the content of cannabinoids in Cannabis sativa L.: A comparative study. *Talanta*, **2018**, DOI 10.1016/j.talanta.2018.07.085.

[5] L. A. Ciolino, T. L. Ranieri, A. M. Taylor. Commercial cannabis consumer products part 1: GC–MS qualitative analysis of cannabis cannabinoids. *Forensic Sci. Int.*, **2018**, DOI 10.1016/j.forsciint.2018.05.032.

[6] C. M. Andre, J. F. Hausman, G. Guerriero. Cannabis sativa: The plant of the thousand and one molecules. *Front. Plant Sci.*, **2016**, DOI 10.3389/fpls.2016.00019.

[7] U. Avico, R. Pacifici, P. Zuccaro. Variations of tetrahydrocannabinol content in cannabis plants to distinguish the fibre-type from drug-type plants. *Bull. Narc.*, **1985**.

[8] C. scientifique de l'Agence fédérale pour la S. de la C. alimentaire Alimentaire. *THC(Tétrahydrocannabinol) Dans Les Denrées Alimentaires d'origine Animale : Score Des Dangers et Seuils d'actions*, **2017**.

[9] B. T. Borille, M. González, L. Steffens, R. S. Ortiz, R. P. Limberger. Cannabis Sativa: a Systematic Review of Plant Analysis. *Drug Anal. Res.*, **2017**, *1*, 1–23.

[10] I. G. Trofin, G. Dabija, D. I. Vaireanu, L. Filipescu. Long term storage and cannabis oil stability. *Rev. Chim.*, **2012**, *63*, 293–297.

[11] E. Russo, G. W. Guy. A tale of two cannabinoids: The therapeutic rationale for combining tetrahydrocannabinol and cannabidiol. *Med. Hypotheses*, **2006**, DOI 10.1016/j.mehy.2005.08.026.

[12] T. swiss Authorities. "www.ch.ch," Available at: https://www.ch.ch/fr/cannabis/, **n.d.**

[13] S. P. Health. "www.health.belgium.be," Available at: https://www.health.belgium.be/en/herbal-products-smoking-0, **n.d.**

[14] F. Grotenhermen. Les cannabinoides et le systeme des endocannabinoides. *Cannabinoids*, **2006**.

[15] C. Citti, D. Braghiroli, M. A. Vandelli, G. Cannazza. Pharmaceutical and biomedical analysis of cannabinoids: A critical review. *J. Pharm. Biomed. Anal.*, **2018**, DOI 10.1016/j.jpba.2017.06.003.

[16] A. K. Hewavitharana, G. Golding, G. Tempany, G. King, N. Holling. Quantitative GC-MS analysis of Δ9-tetrahydrocannabinol in fiber hemp varieties. *J. Anal. Toxicol.*, **2005**, DOI 10.1093/jat/29.4.258.

[17] D. Cozzolino. Near infrared spectroscopy in natural products analysis. *Planta Med.*, **2009**, DOI 10.1055/s-0028-1112220.

[18] C. Mees, F. Souard, C. Delporte, E. Deconinck, P. Stoffelen, C. Stévigny, J. M. Kauffmann, K.

De Braekeleer. Identification of coffee leaves using FT-NIR spectroscopy and SIMCA. *Talanta*, **2018**, *177*, 4–11.

[19] B. T. Borille, M. C. A. Marcelo, R. S. Ortiz, K. de C. Mariotti, M. F. Ferrão, R. P. Limberger. Near infrared spectroscopy combined with chemometrics for growth stage classification of cannabis cultivated in a greenhouse from seized seeds. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, **2017**, *173*, 318–323.

[20] C. A. Teixeira Dos Santos, M. Lopo, R. N. M. J. Páscoa, J. A. Lopes. A review on the applications of portable near-infrared spectrometers in the agro-food industry. *Appl. Spectrosc.*, **2013**, *67*, 1215–1233.

[21] M. Jamrógiewicz. Application of the near-infrared spectroscopy in the pharmaceutical technology. *J. Pharm. Biomed. Anal.*, **2012**, DOI 10.1016/j.jpba.2012.03.009.

[22] R. Risoluti, G. Gullifa, A. Battistini, S. Materazzi. Monitoring of cannabinoids in hemp flours by MicroNIR/Chemometrics. *Talanta*, **2020**, *211*, 120672.

[23] R. Kumar, V. Sharma. Chemometrics in forensic science. *TrAC - Trends Anal. Chem.*, **2018**, DOI 10.1016/j.trac.2018.05.010.

[24] L. A. C. Europ. *Règlement d'exécution (UE) N° 1306/2014 de La Comission*, **2014**.

[25] D. de l'agriculture et de la Pêche. "https://lv.vlaanderen.be," Available at: https://lv.vlaanderen.be/nl/plant/akkerbouw/melden-van-teeltintentie-hennep, **n.d.**

[26] N. Wilson, M. Heinrich. The Use of Near Infrared Spectroscopy to discriminate between THC-rich and hemp forms of Cannabis. *Planta Med.*, **2006**, DOI 10.1055/s-2006-950060.

[27] ONUDC. *Méthodes Recommandées Pour l'identification et l ' Analyse Du Cannabis et Des Produits Du Cannabis*, **n.d.**

[28] B. Jonathan. Chapter 5 : Diffuse Reflectance Spectroscopy, in *Mod. Tech. Appl. Mol. Spectrosc.*, (Ed: M.F. Wiley-Interscience), New-York, **1998**, p. 145.

[29] A. J. Poortman-Van Der Meer, H. Huizer. A contribution to the improvement of accuracy in the quantitation of THC. *Forensic Sci. Int.*, **1999**, *101*, 1–8.

[30] P. A. Gorry. General Least-Squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method. *Anal. Chem.*, **1990**, DOI 10.1021/ac00205a007.

[31] E. R. Ziegel, D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. de Jong, P. J. Lewi, J. S. Verbeke. Handbook of Chemometrics and Qualimetrics, Part B. *Technometrics*, **2000**, DOI 10.2307/1271476.

[32] B. M. Wise, N. B. Gallagher, R. Bro, J. M. Shaver, W. Windig, R. S. Koch. *Chemometrics Tutorial for PLS _ Toolbox and Solo*, **2006**.

[33] R. D. Snee. Validation of Regression Models: Methods and Examples. *Technometrics*, **1977**, DOI 10.2307/1267881.

[34] S. Wold. Pattern recognition by means of disjoint principal components models. *Pattern Recognit.*, **1976**, *8*, 127–139.

[35] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *J. Pharm. Biomed. Anal.*, **2007**, *44*, 683–700.

[36] N. F. Pérez, J. Ferré, R. Boqué. Calculation of the reliability of classification in discriminant partial least-squares binary classification. *Chemom. Intell. Lab. Syst.*, **2009**, DOI

10.1016/j.chemolab.2008.09.005.

[37] D. Ballabio, V. Consonni. Classification tools in chemistry. Part 1: Linear models. PLS-DA. *Anal. Methods*, **2013**, DOI 10.1039/c3ay40582f.

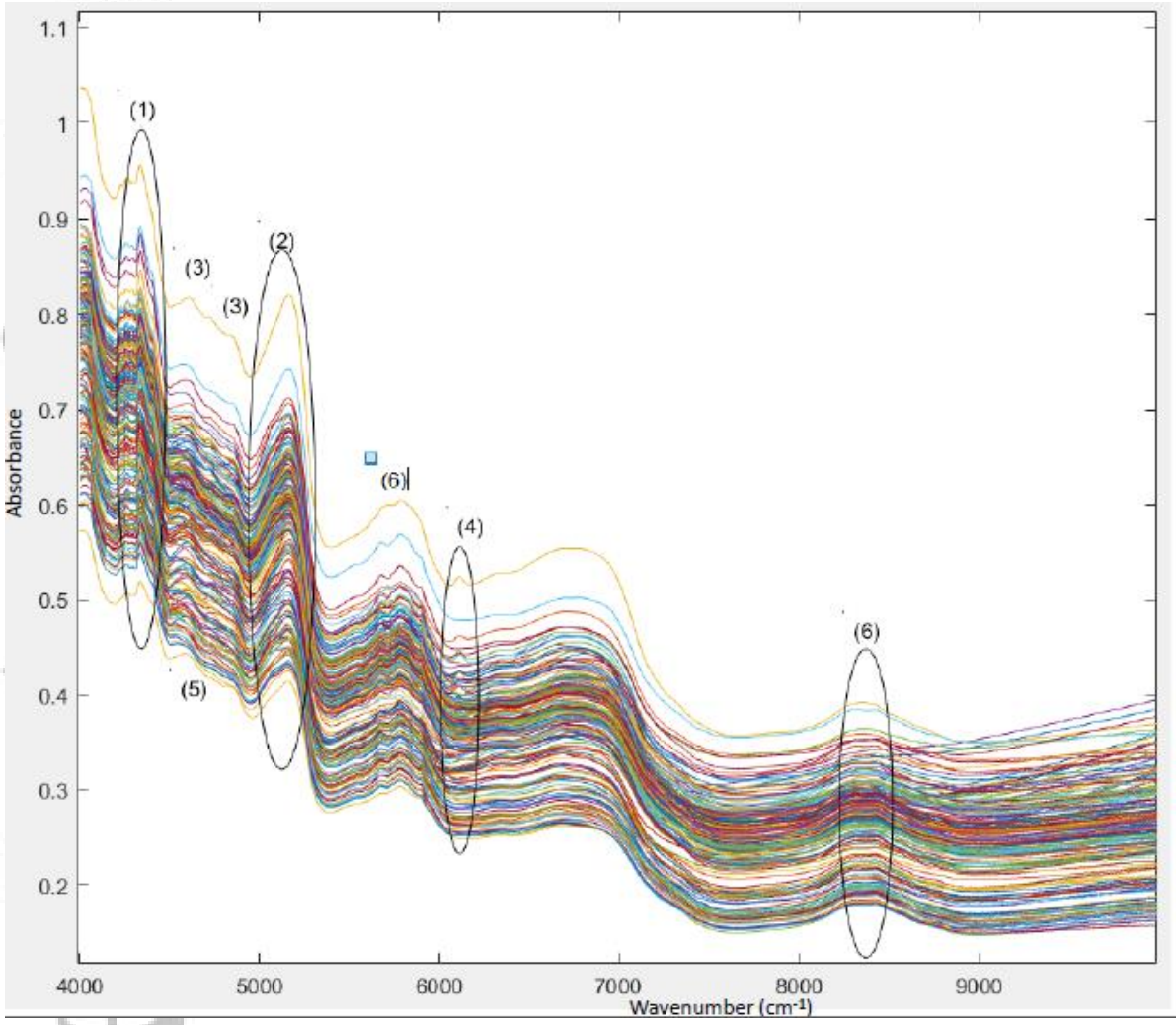

Fig.1 : NIR absorbance raw spectra obtained with the benchtop in the wavenumber region 10,000 – 4,000 cm^{-1} (A) and the dispersive absorbance raw spectra obtained with the handheld in the wavenumber region 6,000 – 4,000 cm^{-1} (B).

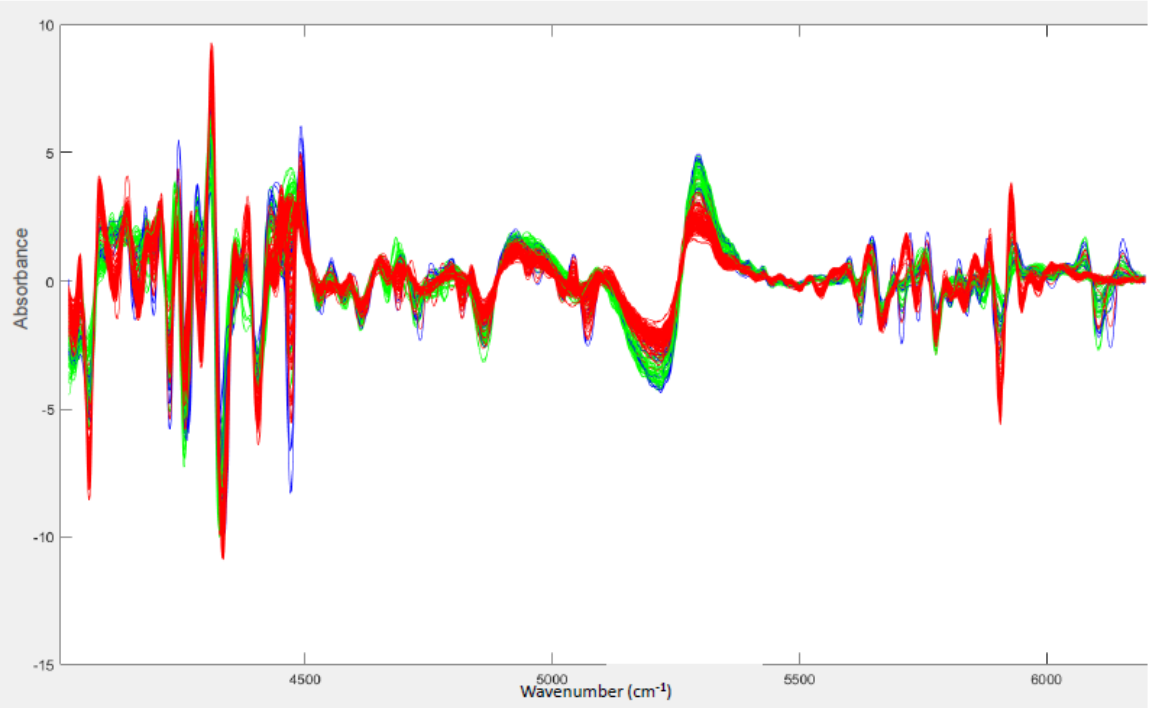

Fig. 2 : NIR spectra obtained with the benchtop for all samples after SNV followed by the second derivative.

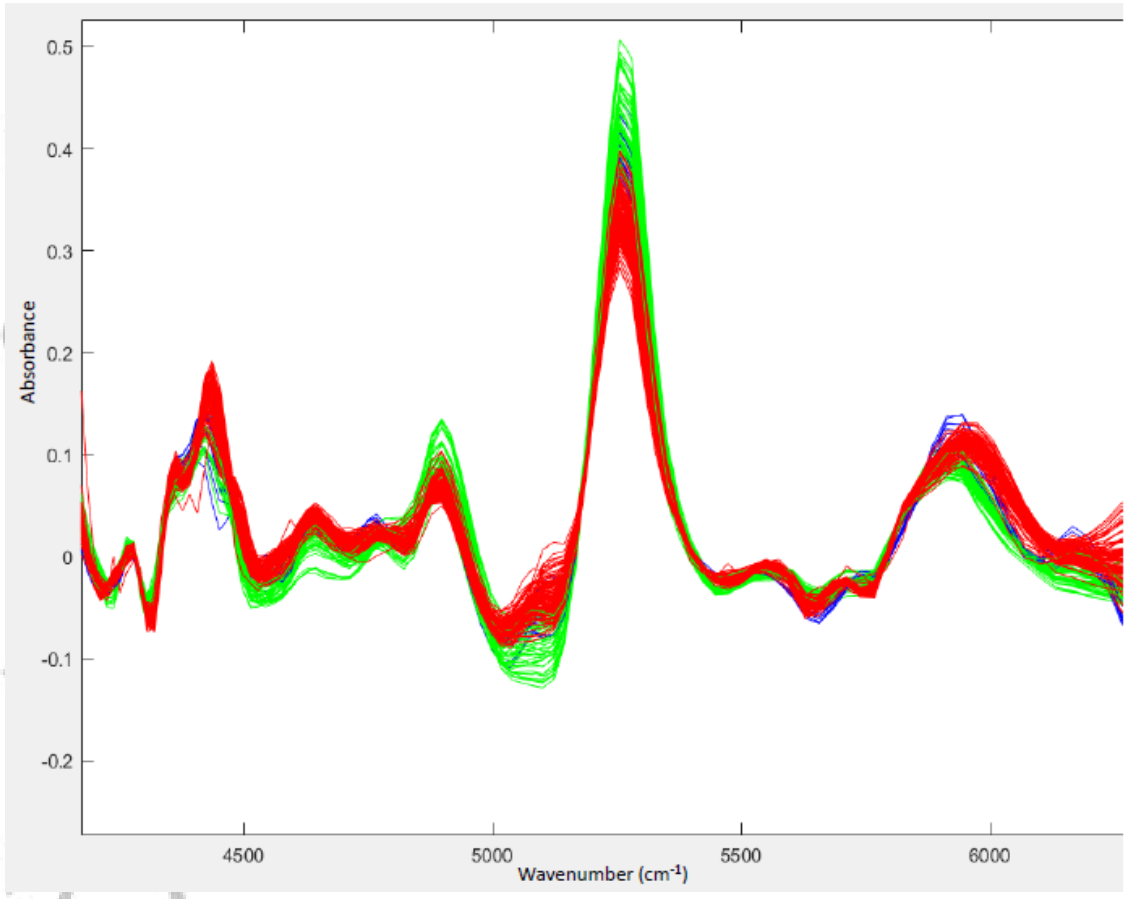

Fig. 3 : NIR spectra obtained with the handheld device for samples after the first derivative followed by SNV.

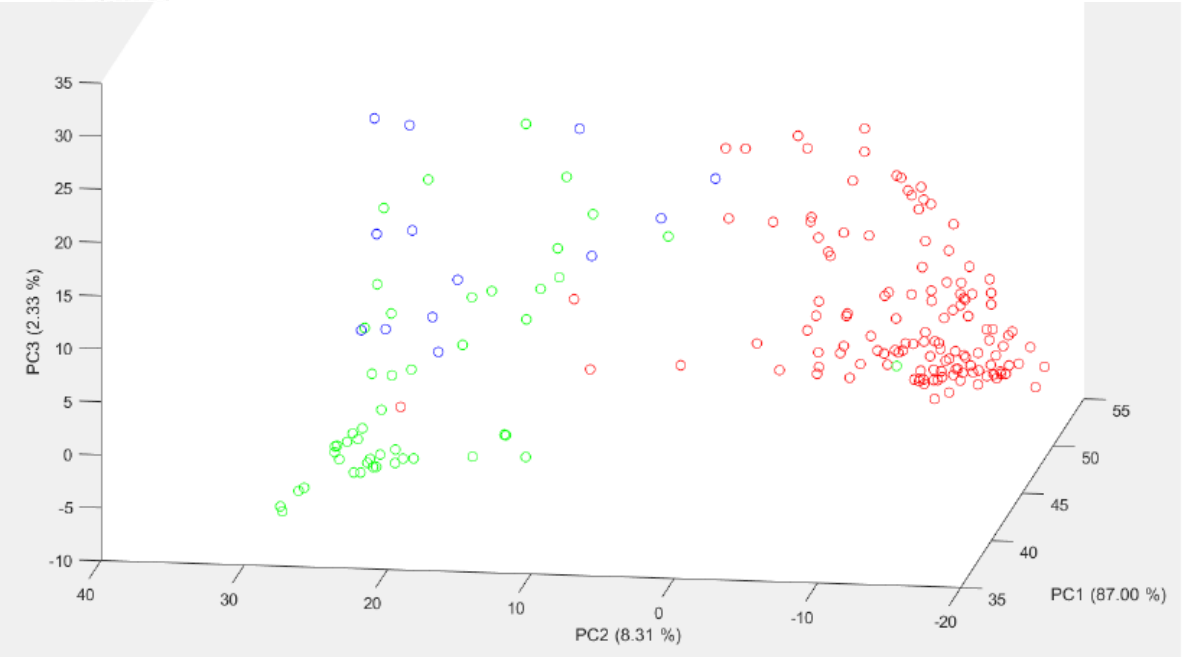

Fig. 4 : PCA plot obtained with the NIR benchtop spectra. Samples are indicated with different colors depending of the THC concentration : lower than 0.2 % (w/w) (green), between 0.2 – 1 % (w/w) (blue) and higher than 1 % (w/w) (red).

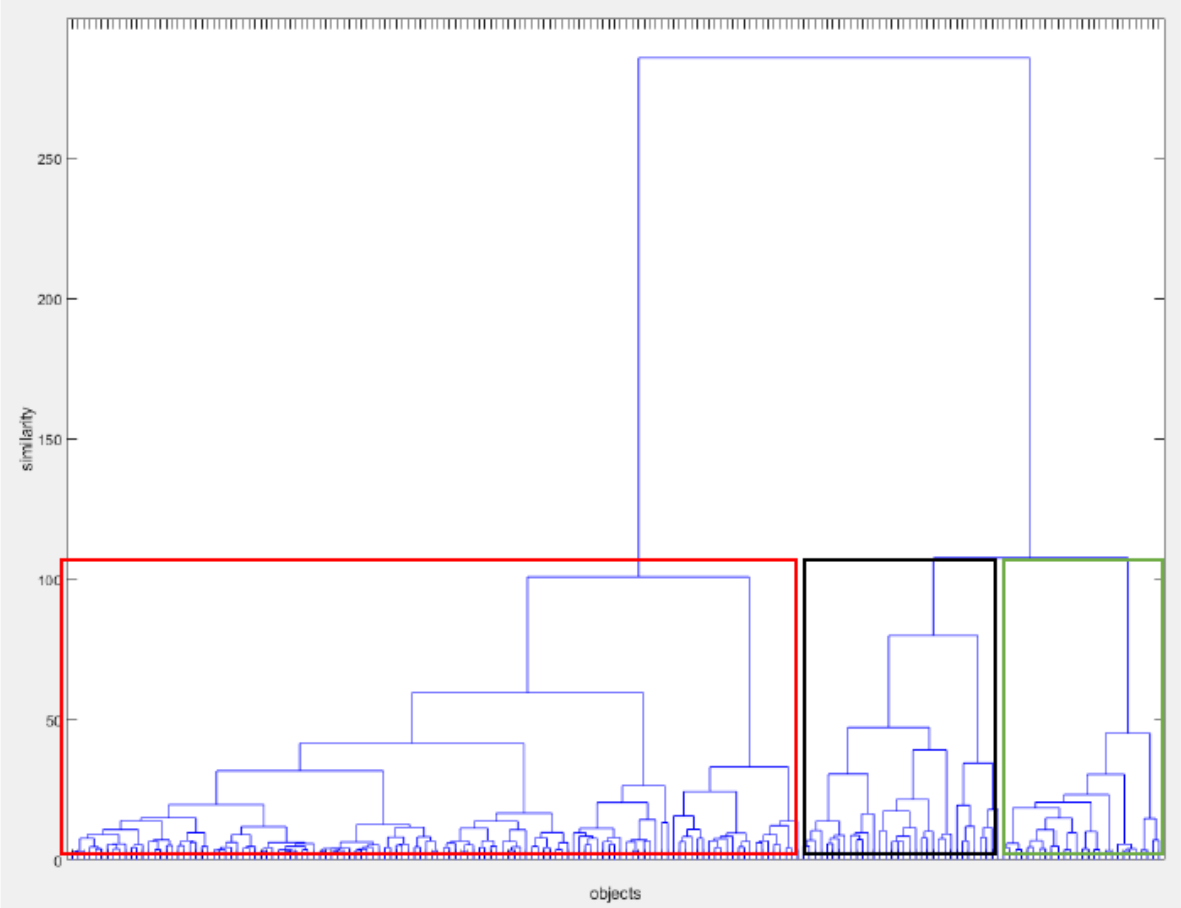

Fig. 5 : Dendrogram constructed via hierarchical clustering on NIR spectra obtained with the benchtop. Samples are grouped depending the THC concentration : lower than 0.2 % (w/w) (green), between 0.2 – 1 % (w/w) (black) and higher than 1 % (w/w) (red).

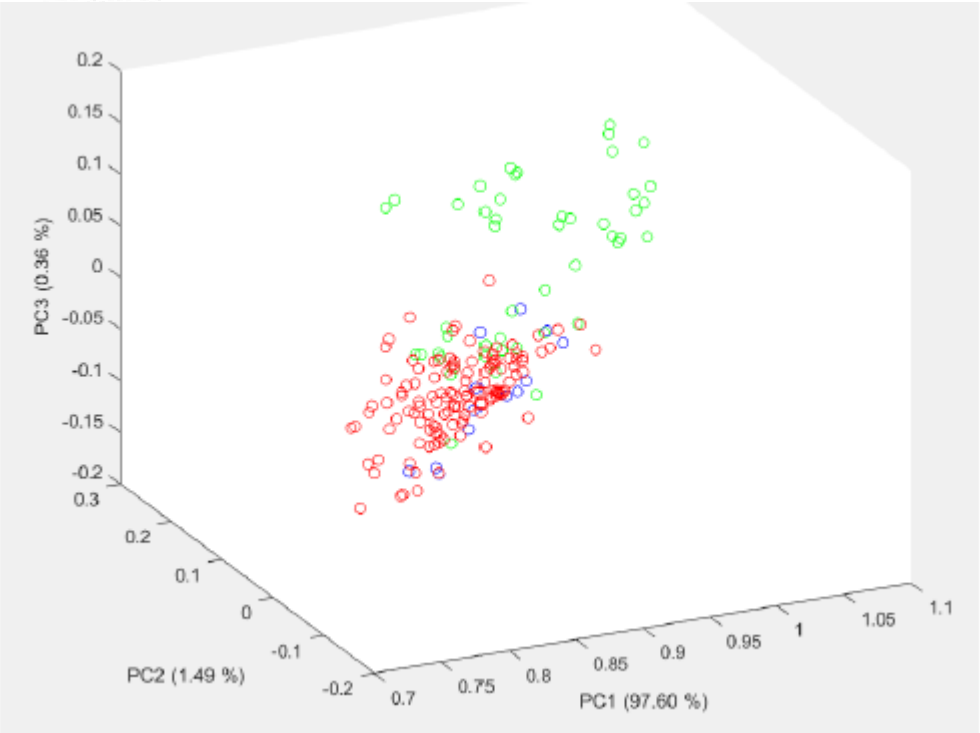

Fig. 6 : PCA plot obtained with the NIR handheld spectra. Samples are indicated with different colors depending of the THC concentration : lower than 0.2 % (w/w) (green), between 0.2 – 1 % (w/w) (blue) and higher than 1 % (w/w) (red).

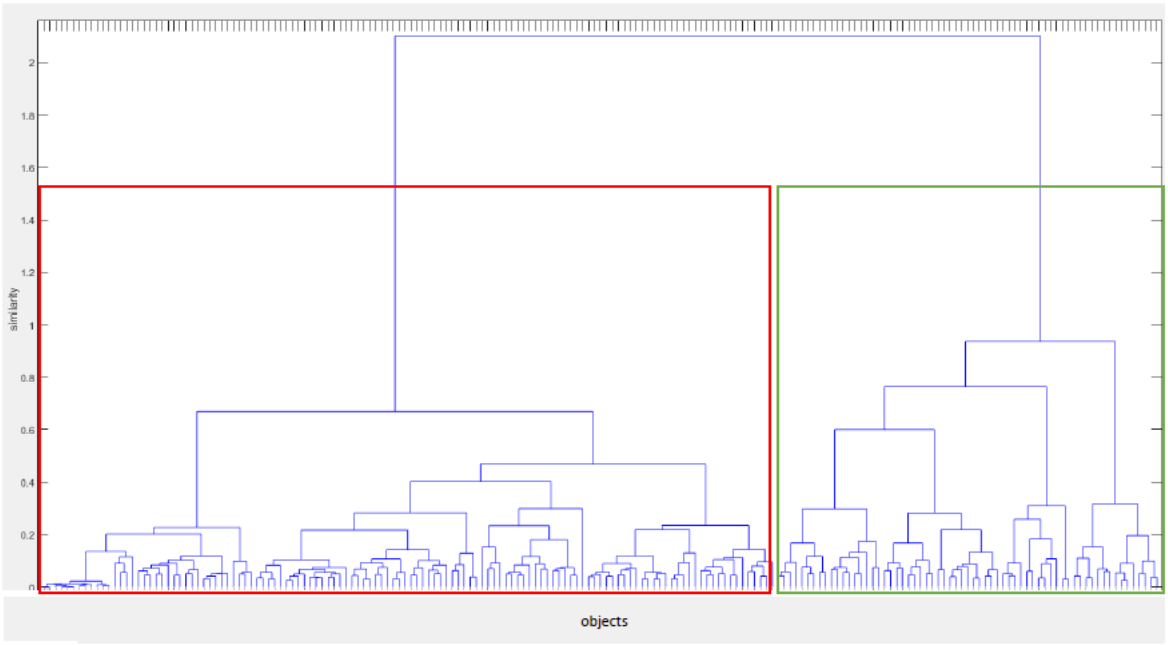

Fig. 7 : Dendrogram constructed via hierarchical clustering on NIR spectra obtained with the handheld device. Samples are grouped depending the THC concentration : lower than 1 % (w/w) (green) and higher than 1 % (w/w) (red).

Table 1 : Correct classification rate (ccr %) obtained with PLS-DA, k-NN and SIMCA for benchtop and handheld NIR spectra

Devices	**Benchtop**	**Handheld**
Preprocessing	2nd derivative -SNV	SNV-1st derivative
	2 classes	
	PLS-DA (1)	
Number of PLS-factors	**6**	**7**
Cross-validation (ccr %)	**96**	**98**
External validation (ccr %)	**85**	**91**
	kNN	
Number of K	**5**	**3**
Cross validation (ccr %)	**92**	**98**
External validation (ccr %)	**87**	**91**
	SIMCA	
Number of PCs	**1-1**	**1-2**
Cross validation (ccr %)	**92**	**97**
External validation (ccr %)	**91**	**93**
	3 classes	
	PLS-DA (2) Bayes theorem	
Number of PLS-factors	**4**	**3**
Cross-validation (ccr %)	**97** (113/117)	**98** (114/116)
External validation	**91** (41/45)	**95** (42/44)
	PLS-DA (1)	
Number of PLS-factors	**5**	**13**
Cross-validation (ccr %)	**92**	**99**
External validation (ccr%)	**84**	**88**
	k-NN	
Number of K	**7**	**3**
Cross validation (ccr %)	**92**	**95**
External validation (ccr %)	**87**	**91**
	SIMCA	
Number of PCs	**1-1-1**	**3-2-2**
Cross validation (ccr %)	**89**	**95**
External validation (ccr %)	**80**	**84**

Table 2 : Binary classification based on Belgian legislation - Classification statistics for cross-validation and test set

BENCHTOP											
SIMCA				Class 1 : > 0.2 % (w/w) THC				Class 2 : < 0.2 % (w/w) THC			
	Accuracy	NER	ER	Precision	Specificity	Sensitivity	PCs	Precision	Specificity	Sensitivity	PCs
Cross-validation	0.92	0.94	0.063	0.99	0.96	0.91	1	0.74	0.91	0.96	1
Test set	0.91	0.93	0.071	1.00	1.00	0.86		0.81	0.86	1.00	

HANDHELD							
SIMCA				Class 1 : > 0.2 % (w/w) THC			
	Accuracy	NER	ER	Precision	Specificity	Sensitivity	PCs
Cross-validation	0.97	0.98	0.019	1.00	1.00	0.96	1
Test set	0.93	0.95	0.05	1.00	1.00	0.89	

Table 3 : Tertiary classification based on Belgian and Switzerland legislation - Classification statistics for cross-validation and test set

BENCHTOP													
PLS-DA				Bayes	Class 1 : > 1 % (w/w) THC			Class 2 : 0.2 -1 % (w/w) THC			Class 3 : < 0.2 % (w/w) THC		
PLS factors : 4	Accuracy	NER	ER	Not assigned	Precision	Specificity	Sensitivity	Precision	Specificity	Sensitivity	Precision	Specificity	Sensitivity
Cross-validation	0.96	0.95	0.05	0.093	0.99	0.96	0.98	0.25	0.97	1.00	0.95	0.99	0.86
Test set	0.91	0.92	0.083	0.20	1.00	1.00	0.97	0.25	0.93	1.00	0.92	0.97	0.79
HANDHELD													
PLS-DA				Bayes	Class 1 : > 1 % (w/w) THC			Class 2 : 0.2 – 1 % w/w THC			Class 3 : < 0.2 % (w/w) THC		
PLS factors: 3	Accuracy	NER	ER	Not assigned	Precision	Specificity	Sensitivity	Precision	Specificity	Sensitivity	Precision	Specificity	Sensitivity
Cross-validation	0.96	0.91	0.09	0.12	1.00	1.00	0.97	0.6	0.98	0.75	0.91	0.98	1.00
Test-set	0.95	0.83	0.17	0.21	1.00	1.00	1.00	1.00	1.00	0.50	0.87	0.94	1.00