

# Cloud-Enabled Handheld NIR Spectroscopy: A Transformative Approach for Real-Time Forensic Analysis of Cannabis Specimens

Florentin Coppey,<sup>a</sup> Cédric Schelling,<sup>b, c</sup> Jean-Luc Veuthey,<sup>b, c</sup> and Pierre Esseiva<sup>\*a</sup>

<sup>a</sup> University of Lausanne, School of Criminal Justice, CH-1015, Lausanne, Switzerland,  
e-mail: pierre.esseiva@unil.ch

<sup>b</sup> School of Pharmaceutical Sciences, University of Geneva, CMU – Rue Michel Servet 1, CH-1211, Geneva 4, Switzerland

<sup>c</sup> Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, CMU – Rue Michel Servet 1, CH-1211, Geneva 4, Switzerland

Dedicated to Prof. *Robert Deschenaux* on the occasion of his retirement

© 2023 The Authors. *Helvetica Chimica Acta* published by Wiley-VHCA AG. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

In the past few years, there has been significant interest within the forensic community regarding the deployment of portable solutions that provide real-time results. This article introduces an innovative technology or technology architecture that enables the integration of a handheld device, specifically, *Viavi MicroNIR*, with a cloud-based system. This cloud system encompasses a server responsible for data processing and a mobile application acting as a user interface.

To demonstrate the transformative impact of this technology on field operators, the analysis of cannabis specimens has been utilized. System's capacity to distinguish between CBD-type and THC-type cannabis has been particularly highlighted, along with the remarkable congruence observed between the near-infrared (NIR) spectra and the reference analytical method involving ultra-high-performance liquid chromatography (UHPLC). The article will present the advantages of this application primarily focusing on its potential to alleviate the burden on laboratories by expediting routine illicit drug analysis. *Viavi MicroNIR* technology provides laboratory personnel with additional time to handle more complex cases, thereby enhancing overall efficiency.

**Keywords:** analytical methods, big-data analysis, cannabis analysis, forensic science, liquid chromatography, machine learning, near infrared (NIR), ultra-high-performance liquid chromatography (UHPLC).

## Introduction

*Cannabis sativa* L. is by far the most widely consumed illicit drug in the world.<sup>[1]</sup> It is also the most controversial product due to its potential for both recreational and medicinal purposes and legislation varying depending on the country.

This disparity extends to different forms that are available on the licit or illicit market. In Switzerland, for instance, marijuana and resin are the most encountered form. Marijuana refers to the dried vegetal forms such as buds (called marijuana) and resin is the concentrated product obtained by using a grinder to mechanically extract a fine powder that concentrates

the part of the plant (the trichomes) containing the highest concentration of  $\Delta^9$ -tetrahydrocannabinol (THC), the predominant cannabinoid responsible of the psychoactive effects. Additionally, other products can be found like cannabis oil extracts or edibles. The diversity of cannabis products is further reflected in the various methods of consumption, including smoking, inhalation, ingestion, drinking, and vaporization.

The cannabis plant contains a complex mixture of terpenes, flavonoids, phenolic derivatives, and more than 90 cannabinoids were present in the inflorescence. Among them, THC and cannabidiol (CBD) were the most abundant.<sup>[2]</sup> However, it is important to mention that CBD and THC are not biosynthesized in

the plant but are present in their cannabidiol acid (CBDA) and tetrahydrocannabinol acid (THCA) form.<sup>[3]</sup> To obtain the active forms of THC (tetrahydrocannabinol) and CBD (cannabidiol), a process called decarboxylation is required. Decarboxylation involves the removal of a carboxyl group from the cannabinoid molecules through the application of heat. This process converts the inactive forms of THC and CBD (THCA and CBDA, resp.) into their active and psychoactive forms.

In the context of cannabis consumption, one common method of achieving decarboxylation is by smoking the cannabis plant, often mixed with tobacco. When the cannabis plant material is exposed to heat through combustion, the decarboxylation process occurs, converting THCA and CBDA into THC and CBD, respectively. Smoking cannabis mixed with tobacco is a popular practice. This method effectively allows for the decarboxylation of cannabinoids through the heat generated during smoking.

The characterization of *cannabis sativa* L. is a significant topic in drug analysis as it has raised numerous questions within the community in recent years. Due to the varying legislation in different countries, where cannabis can be either a controlled or legal substance, several inquiries have emerged, primarily regarding the determination of potency and the ability to differentiate between high-THC, low-CBD content (THC-type) and low-THC, high-CBD content (CBD-type) products.<sup>[4]</sup>

To answer these questions, it is necessary to quantify the amount of THC. For instance, different limits can be defined to determine if one is dealing with CBD-type or drug-type cannabis. In Switzerland, the legal limit is set for THC at  $\leq 1\%$ ,<sup>[5]</sup> while in other European countries, the limit may vary from  $\leq 0.2\%$  to  $\leq 0.6\%$ .<sup>[4]</sup>

There is a significant demand for the availability of rapid analysis methods for forensic purposes. To tackle these challenges, various initiatives have been published, focusing on the deployment of rapid technologies for the analysis of illicit or licit drugs.<sup>[6–9]</sup> Spectroscopic techniques such as *Raman* or Near Infrared spectroscopy (NIR) have shown excellent results in identifying and quantifying illicit drugs.<sup>[10–13]</sup> This trend has also extended to cannabis analysis, with several publications demonstrating the potential of spectroscopy in analyzing cannabis samples.<sup>[14–16]</sup>

These rapid techniques based on spectroscopy have undeniable advantages, including their speed, portability, non-destructiveness, minimum sample preparations requirements. This opens a new possibil-

ity to deploy these techniques directly into the field in a various setting.<sup>[17]</sup> For example, in a forensic context, they can be used for determining the CBD-type or THC-type of substances.<sup>[17]</sup> In drug checking settings related to public health, they can inform consumers about the potency of drugs. Additionally, in the hemp industry, they can be utilized for monitoring the maturity of hemp production.

This article focuses on discussing the performance of a portable NIR spectrometer for the qualification and quantification of cannabis samples. The evaluation will be conducted by comparing the results with a well-established method commonly used for analyzing such samples: ultra-high-performance liquid chromatography (UHPLC) coupled with diode-array detection (DAD).<sup>[18,19]</sup> The UHPLC method measures the total THC (THC + THCA) and total CBD (CBD + CBDA) content.

As previously mentioned, the cannabis plant primarily contains the acidic forms of THC and CBD. However, when the product is consumed, such as through smoking, the acidic forms undergo decarboxylation to form THC. Therefore, measuring the total THC provides a better indication of the potency of the final product that will be consumed. The total CBD content of the specimens used in this study has also been systematically measured. Since the specimens are from police seizures, there is no prior information available regarding the type of cannabis. Two approaches have been evaluated for the treatment of NIR data. The first approach is more qualitative, aiming to differentiate between CBD-type and THC-type cannabis by employing exploratory data analysis techniques like principal component analysis (PCA). The second approach utilizes chemometric models based on artificial intelligence to predict the total THC content in marijuana and resin cannabis specimens.<sup>[7,16]</sup>

Finally, this article presents the advantages of combining spectroscopic approaches with the traditional chromatographic techniques to address the challenging workloads faced by laboratories. The use of these combined techniques provides instant results and enables the deployment of technologies directly into the field. This allows users to analyze seized samples independently, contributing to more efficient and expedited monitoring of the illicit drugs market.

## Results and Discussion

### Cannabis Samples Content Analyzed by UHPLC

To evaluate the potential of NIR approach in correctly identifying and classifying cannabis samples (CBD-type vs. THC-type) and predicting potency, 257 randomly selected cannabis specimens were obtained from police seizures and sent to the laboratory. These samples were initially analyzed using the reference separative UHPLC method for classification.

Based on the total THC and total CBD quantification obtained through UHPLC, the samples were categorized into four groups. The categories were determined based on their form (resin vs. marijuana) and their total THC and CBD content. The clustering is as follows: 109 marijuana specimens classified as THC-type, 27 marijuana specimens classified as CBD-type, 103 resin specimens classified as THC-type, and 18 resin specimens classified as CBD-type.

The summarized results of total THC and total CBD quantification for the 257 cannabis specimens are presented in *Figure 1*. For THC-type marijuana specimens, the mean THC content is 15.4% with most of the specimens ranging from 13.1 to 18.2%, and the mean CBD is 0.1% with most of the specimens ranging from 0.02 to 0.04%.

For resin THC-type specimens, the mean THC is 25.6% content with most of the specimens ranging from 23.9 to 29%, and the mean CBD is 1.4% with most of the specimens ranging from 1.0 to 1.89%.

For CBD-type marijuana specimens, the mean CBD content is 12.6% with most of the specimens ranging from 11.0 to 14.6%, and the mean THC is 0.9% with most of the specimens ranging from 0.5 to 0.9%.

For resin CBD-type specimens, the mean CBD is 18.1% content with most of the specimens ranging from 17.0 to 20.7%, and the mean THC is 0.8% with most of the specimens ranging from 0.5 to 1.9%.

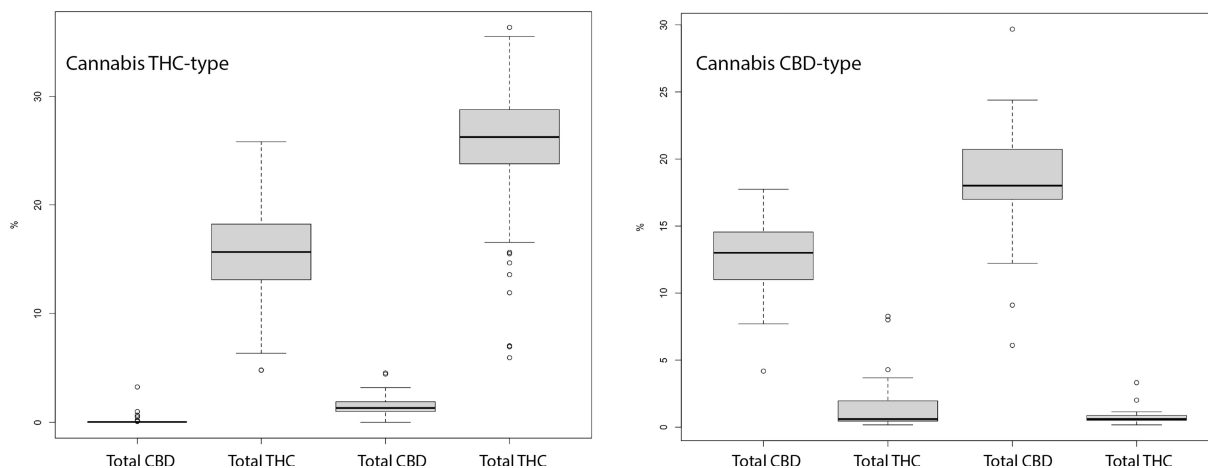
### Qualitative Approach: Differentiation between THC-Type and CBD-Type Cannabis by NIR

The potential for NIR spectra to distinguish between THC and CBD-type has been investigated using PCA visualization. *Figures 2* and *3* demonstrate excellent separation between marijuana and resin cannabis. This highlights the capability of MicroNIR to accurately assign the cannabis type based on prior classification using UHPLC. These results provide law enforcement organizations with an efficient and reliable means to quickly test cannabis seizures.

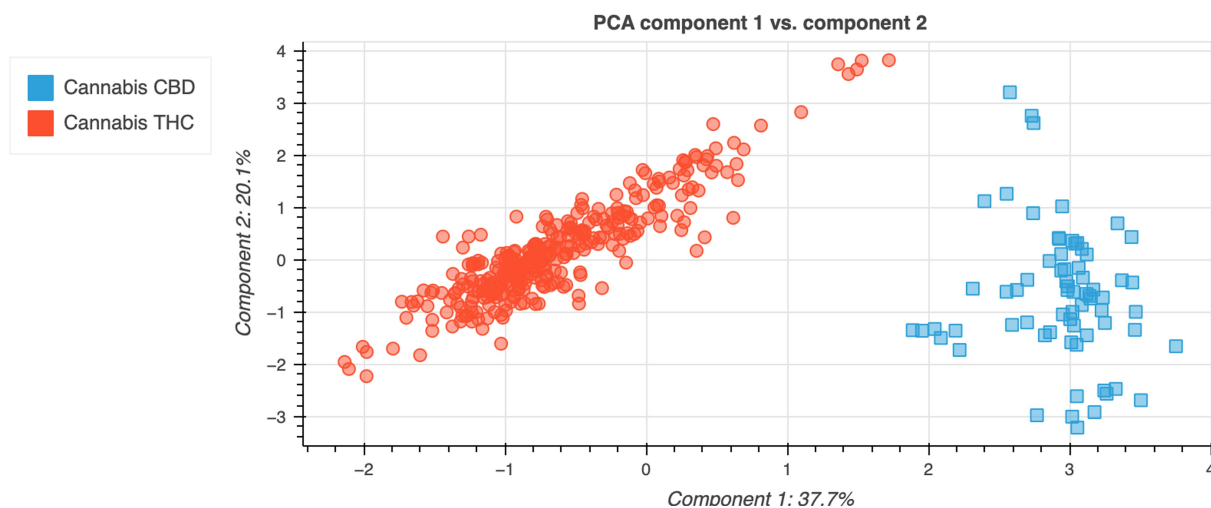
### Quantitative Models

The cannabis samples were categorized into marijuana and resin samples based on physical characteristics. For both groups, a training set consisting of 2/3 of the specimen and a validation set comprising the remaining specimens were used to develop the two models. The division into these datasets was performed using *Kennard–Stone*<sup>[20]</sup> strategy to maximize the variability of the specimens within each group.

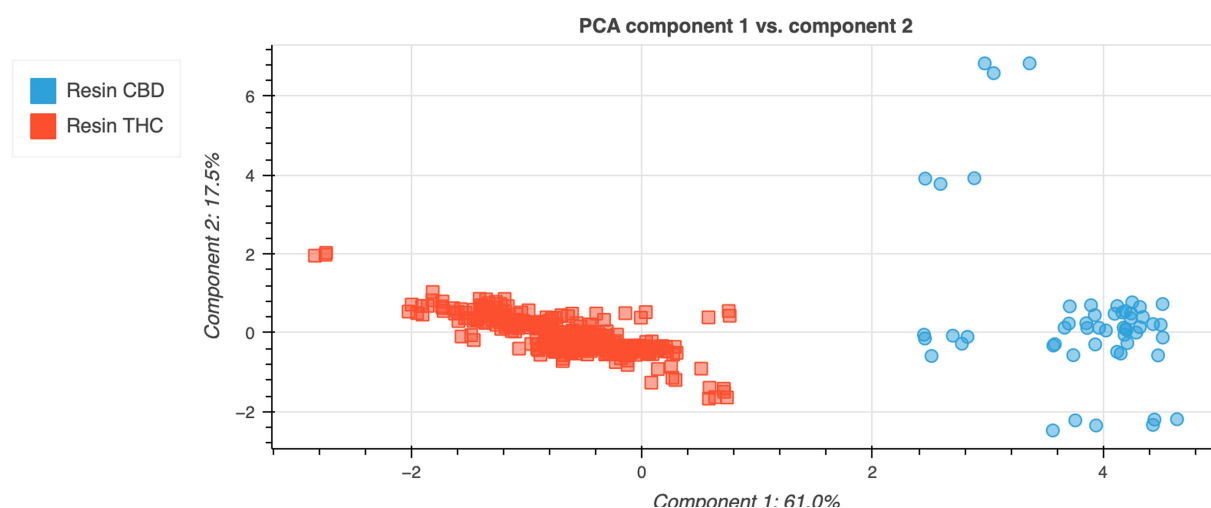
To evaluate the correlation between NIR THC prediction and UHPLC THC quantification, an acceptable limit of 2.5% in absolute difference was defined. The evaluation of regression models for marijuana specimens (see *Figure 4*) and resin specimens (see *Figure 5*) demonstrates that the results obtained with



**Figure 1.** Distribution of total THC and total CBD content analyzed by UHPLC-DAD for marijuana and resin specimens.



**Figure 2.** PCA separation of the two populations of marijuana samples (THC-type vs. CBD-type), using the scores of the two main principal components with the 2nd derivative and SNV pre-treatment of the raw spectra.



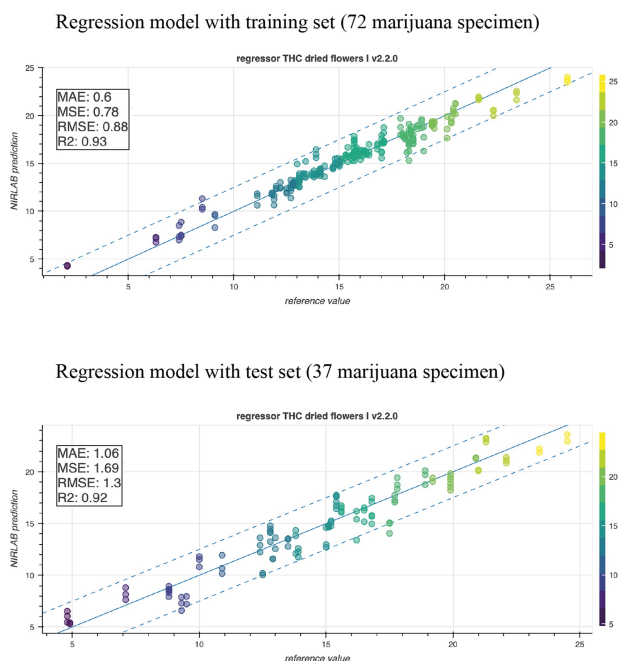
**Figure 3.** PCA separation of the two populations of resin samples (THC-type vs. CBD-type), using the scores of the two main principal components with the 2nd derivative and SNV pre-treatment of the raw spectra.

the NIR method are remarkably close to those obtained with the reference method, not only for the training groups, but also for the test groups. In fact, more than 95% of the specimens in the test set fall within the limit of 2.5%.

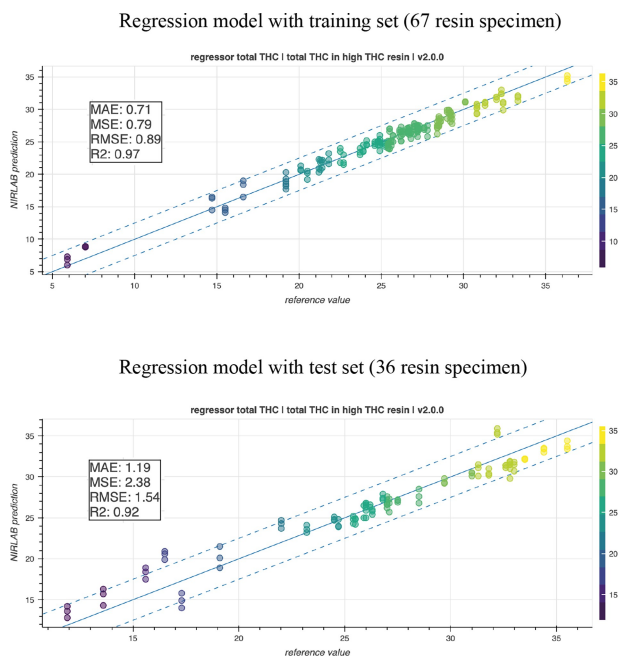
The results presented in this article highlight the potential of the MicroNIR to provide a valid information about the cannabis specimen. It can determine the type (CBD-type vs. THC-type) as well as the content of total THC. This approach offers new possibilities by bringing the laboratory to the field<sup>[21,22]</sup> and providing real-time information. This is particularly valuable for the law enforcement agencies in determining the legality of seized cannabis. The analysis is non-

destructive and chemical-free, making it an excellent alternative to other field tests, such as colorimetric reagents.

Additionally, MicroNIR is an excellent tool for selecting samples that require further analysis in the laboratory. As demonstrated in this article, NIR analysis is a powerful technology for identifying and quantifying most marijuana and resin specimens due to their high THC levels (between 13% and 29%) and clear NIR signals, enabling accurate qualification and quantification. If NIR prediction shows low THC concentration (<5%) or inconsistent results (e.g., identifying a CBD-type when a THC-type is expected), the specimen should be sent to the laboratory for further inves-



**Figure 4.** Summary of the performance of the quantitative regression model for the training set and the test set for the cannabis marijuana specimen. Each specimen has been analyzed in triplicate using three different NIR instruments.



**Figure 5.** Summary of the performance of the regression model for the training set and the test set for the cannabis resin specimen. Each specimen has been analyzed in triplicate using three different NIR instruments.

tigation. For low THC concentration, since the database may have limited samples in this range, analysis using a separative method such as UHPLC is necessary. In the case of CBD-type identification instead of THC-type, performing a complementary analysis is important. Synthetic cannabinoids like MDMB-4en-PINACA or HHC (Hexahydro cannabinol) can activate CBD cannabis, and analytical techniques like GC/MS are utilized to detect and identify these compounds. Portable NIR spectroscopy is again an effective approach for pre-screening unusual specimens, allowing the laboratory to focus on challenging cases and reducing its workload.

To illustrate this, potential cannabis specimens of the last two years provided by law enforcement agencies for forensic analysis were selected (the authors' university has a laboratory specializing in illicit drug analysis).

A total of 1,503 specimens were analyzed using portable NIR spectroscopy, resulting in 827 predictions of marijuana THC type, 328 predictions of marijuana CBD type, and 348 predictions of resin THC type. Following the quantitative triage approach, only 45 specimens (resin THC and marijuana THC) were further analyzed using chromatographic methods to confirm total THC content, which accounted for less than 4% of the total resin and marijuana THC specimens. This approach significantly saves time and reduces the workload on the laboratory.

This strategy allows to focus on CBD-type cannabis seized by the police. During this period, approximately 21% of all specimens (328) were identified as CBD types. As previously discussed, CBD-type cannabis can be adulterated with synthetic cannabinoids. Therefore, these specimens were analyzed using GC/MS screening method. Out of the 328 CBD-type specimens, only 16 (less than 5%) showed the presence of MDMB-4en-PINACA, indicating that this phenomenon is still limited in the Swiss French-speaking area of Switzerland.

NIR Technology enables direct analysis of cannabis specimens by the police, bypassing the need for an analytical laboratory. This allows for the collection of extensive data on the types of cannabis products present in the illegal market that would not otherwise be available. In Switzerland, cannabis consumption is generally not a priority for prosecuting authorities, resulting in police seizures often not being referred to forensic laboratories for analysis. As a result, there is little knowledge about the products available in the market and their potency. Since cannabis is the most used illicit drug, it is frequently seized. Therefore,

crucial information about the composition of this market becomes available, enabling its monitoring.

Real-time access to the data collected by the police authorities provides immediate for the field officers, enhancing the value of forensic data beyond mere data but as a source of intelligence.

This approach can also be applied in a drug checking system to inform the consumers about the potency of the product they intend to consume. Real-time and reliable results are crucial in such deployments.<sup>[23]</sup>

## Conclusion

In this article, several practical applications that exploit technical developments in the miniaturization of analytical instruments, and advancement in communication and data processing algorithms are presented. These applications highlight a general trend of shifting analytical capabilities outside of the traditional forensic laboratory setting, enabling instant results directly in the field. Portable technologies now go beyond conducting presumptive tests and provide highly valuable results comparable to separative reference methods.

Undoubtedly, there is still progress to be made, questions to be answered, and research efforts to be carried out regarding the deployment of portable technologies. However, the forensic community in general, and including forensic laboratories, should seize the opportunity presented by these innovative approaches. These technologies, rooted in the digital transformation of society, have the potential to bring disruptive changes to forensic laboratory operations and their interaction with law enforcement or prosecution authorities. As seen, they can contribute, at least partially, to reducing the workload of laboratories by decreasing the time required for routine analyses. They may also catalyze the reallocation of resources to more challenging cases that fully utilize the skills of employees, thereby revaluing their work.

The decentralization of forensic capabilities represents an important turning point that forensic and analytical laboratories must consider. However, it is important to mention that the rise of portable technologies does not mean the end of forensic laboratories. On the contrary, it redefines the contours of forensic laboratories, and places them at the center of the decision-making process, as they can generate tactical advantages and improve efficiency, such as real-time monitoring of illicit markets.

Reliable, real-time analysis also opens the use of this technology to other applications, particularly in the field of public health (*e.g.*, drug use and misuse). The ability to directly inform illicit drug users of the potency of the cannabis they plan to consume allows them to receive relevant and appropriate risk-reduction and safer-use messages.

## Experimental Section

### *Cannabis Specimens*

To assess the performance of the portable NIR spectrometer, 136 marijuana specimens and 121 resin specimens were randomly selected from police seizures conducted between 2020 and 2021 in the French-speaking part of Switzerland.

### *UHPLC Quantification Method*

*Chemicals and Reagents.* Methanol, ethanol, 2-propanol, water, and acetonitrile of UHPLC-MS grade were purchased from *Fischer Scientific* (Loughborough, UK). Formic acid was obtained from *Biosolve* (Valkenswaard, Netherlands). All phyto-cannabinoid standard solutions at 1 mg/mL in EtOH ( $\Delta^9$ -THC), in MeOH (CBD), in MeCN (CBDA) and in 2-PrOH (THCA–A) were obtained from *Lipomed AG* (Arlesheim, Switzerland).

*Standard and Calibration Solutions.* The standard stock solution containing  $\Delta^9$ -THC, CBD, THCA–A, CBD–A was prepared at a concentration of 250  $\mu\text{g/mL}$ . All stock solutions were stored at  $-20^\circ\text{C}$ . The standard stock solution was diluted in water/MeCN (3:7, v/v) to obtain the final concentrations of 1.95, 3.90, 7.81, 15.62, 31.25, 62.50 and 125.00  $\mu\text{g/mL}$  for each analyte.

*Preparation of Cannabis Samples.* Cannabis sample preparation consisted of a solid–liquid extraction using ethanol as extraction solvent. Then, 10 mL of ethanol was added to 500 mg of plant material in an *Ika* ultra tube drive system for agitation and grinding for 4 min at 3500 rpm with two glass beads 6 mm in diameter. The mixture was left at ambient temperature for 9 min, and a first centrifugation was carried out for 3 min at 3900 rpm. An aliquot was placed in a 5 mL *Eppendorf* which was centrifuged at 10'000 rpm for 3 min. The supernatant was diluted 50 times before injection into the chromatographic systems. If the measured sample has a concentration above the calibration range, the supernatant of the extraction is diluted so that it is within the assay range. This

extraction protocol has been previously optimized, and the robustness was tested by using a multivariate approach.<sup>[18]</sup>

### UHPLC Analysis

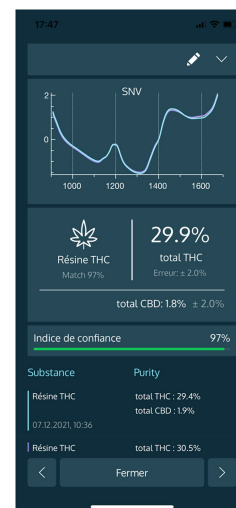
**Apparatus and Methodology.** All experiments were conducted using a *Waters Acquity UPLC H-class* system (*Waters*, Milford, MA, USA) equipped with a quaternary solvent manager, a sample manager with flow through needle (FTN) injector, and a column manager (CM). The wash solvent consisted of a mixture of MeCN, EtOH, and water in a ratio of 4:4:2 (v/v/v), while the purge solvent was a mixture of MeCN and water in a ratio of 7:3 (v/v), with a post-inject wash duration of six seconds. UV Detection at 214 nm was performed using a *Waters PDA* detector connected to the chromatographic system.

**UHPLC-UV Method Conditions.** The separation was carried out at 30 °C using an *InfinityLab Poroshell 120 EC-C18* column (150×2.1 mm, 2.7 μm) from *Agilent* (Santa Clara, USA), along with an *Infinity Lab Poroshell 120 EC-C18* guard column (5×2.1 mm, 2.7 μm) from *Agilent*. The mobile phase A consisted of water with 0.1 % formic acid, while the mobile phase B consisted of MeCN with 0.1 % formic acid.

The gradient profile was as follows: an isocratic mode at 68% B for 2.8 min, followed by an increase from 68% to 73% B in 0.5 min. This composition was held for 3.7 min, then increased to 95% B over 5.0 min, and held for 1.0 min. The percentage of B was then brought back to the initial condition in 0.5 min and maintained for 4.5 min to re-equilibrate the system. The flow rate was set at 0.5 mL/min, and the injection volume was 1 μL.

### Spectroscopy

The portable NIR Spectrometer used in this study is the *MicroNIR Onsite-W 1700* from *Viavi Solutions Inc.* (see *Figure 6*). It operates in the spectral region of 950–1650 nm and consists of a linear variable filter (LVF) directly connected to a 128-pixel linear indium-gallium-arsenic (InGaAs) array detector. The detector has a nominal spectral resolution of 6.25 nm. The signal-to-noise ratio of the instrument is 25,000, the integration time is set to 10 ms, and it collects 100 scans per analysis.



**Figure 6.** *MicroNIR Onsite-W 1700* and the mobile application.

The instrument is connected to the NIRLAB mobile application through Bluetooth as described by *Coppey et al.*<sup>[7]</sup> The mobile application, in turn, connects to a database where algorithms handle the previously acquired NIR signals to predict the type of illicit drugs (such as CBD-type or THC-type) as well as their content (in this case, the percentage of total THC or total CBD). These predictions are made using the predictive models developed in this article.

The technological architecture of NIRLAB is centered around an advanced web interface programmed in Python, which enables the development and maintenance of predictive models. This tool integrates the NIR spectra stored in the reference library with the reference data generated by the forensic laboratory to generate these models. The process of model development and updates involves combining the results obtained from NIR spectra with data from a chromatographic reference analytical method, such as UHPLC-DAD.

By utilizing machine learning algorithms, it becomes possible to create predictive models that are trained to estimate the desired value, whether it is qualitative or quantitative, based on the spectral data from the reference method. These algorithms use the spectral data as input to make predictions about the targeted value.

## NIR Measurements

In each specimen, four measurements are taken either around the flower or on various areas of the resin surface using direct contact, as depicted in *Figure 7*. These steps are repeated with three different *MicroNIR* devices to account for variability among instruments and to standardize measurements for UHPLC-DAD quantification.

Cannabis flowers, the primary products of the plant, exhibit substantial variability. To address this heterogeneity and enhance repeatability, a specific protocol is implemented. The detection capabilities of the instruments are designed to capture reflected light from a surface area of approximately 2 cm<sup>2</sup> on the flower. The spectra collected with the four measurements on the specimen are aggregated and averaged, thus producing a composite spectrum that more accurately represents the flower or the resin.

For model training and testing, each sample consists of a single flower of dimensions 2 cm by 1 cm or larger. It is important to emphasize that this study only examines flowers from recent harvests (< 1 year) that are preserved from excessive sunlight and heat. The potential influence of these external parameters on the models is not investigated in this study.

In addition, only 'consumer-ready' flowers are analyzed; these are post-harvest cured and dried, with a moisture level between 8 and 12%. Given water's significant impact on the NIR spectra in the range of 1300 nm to 1500 nm, the models trained here cannot be applied to fresh flowers.

The models were developed in Python using the scikit-learn framework version 0.24.2.<sup>[24]</sup> For the classification models that distinguish between THC-type and CBD-type flowers and resins, a tree-based pipeline optimization process was employed.<sup>[25]</sup> This process involves searching through the combination of classi-



**Figure 7.** Illustration of the analysis of marijuana and resin cannabis specimen using the *MicroNIR Onsite-W*.



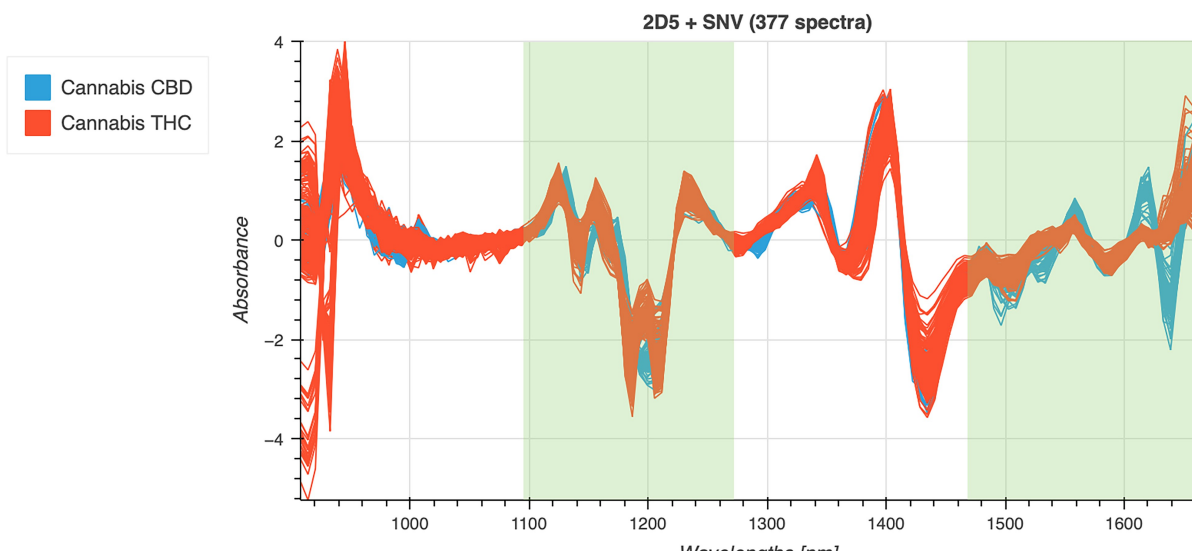
fication algorithms and parameters that yields the highest accuracy in cross-validation (Figure 8).

To enhance reproducibility across different devices, the spectra edges were truncated based on the manufacturer's recommendation. Consequently, the first and last five variables of the spectra were excluded from the training process. This step was taken to achieve better consistency and comparability between different devices.

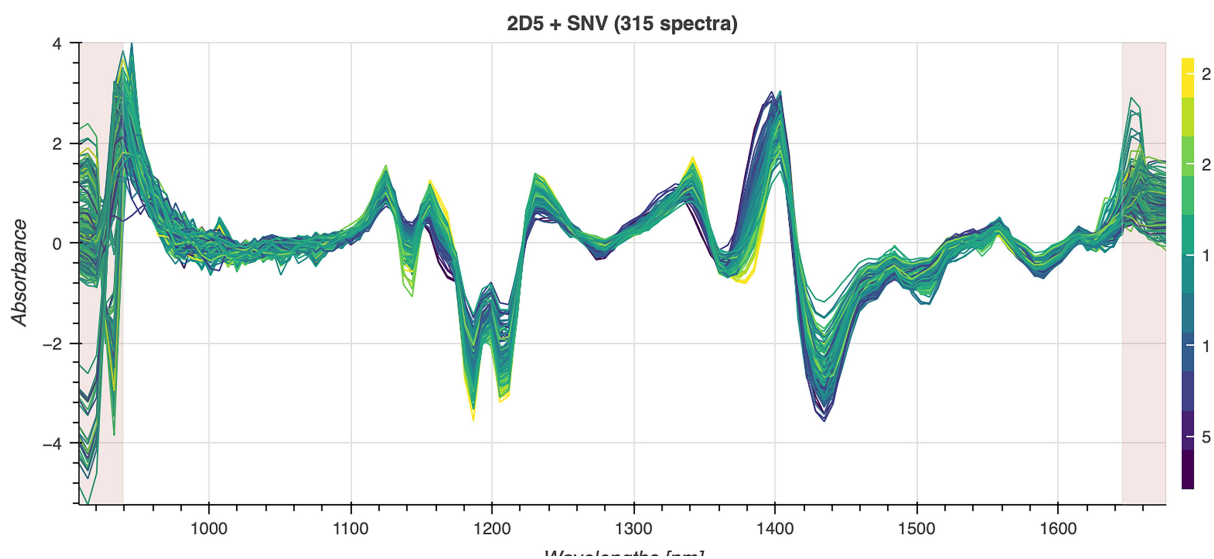
Following the initial model development, a second step of the analysis was performed.

The same method was also applied to train quantitative models for predicting the THC content of THC type flowers and resins (Figure 9).

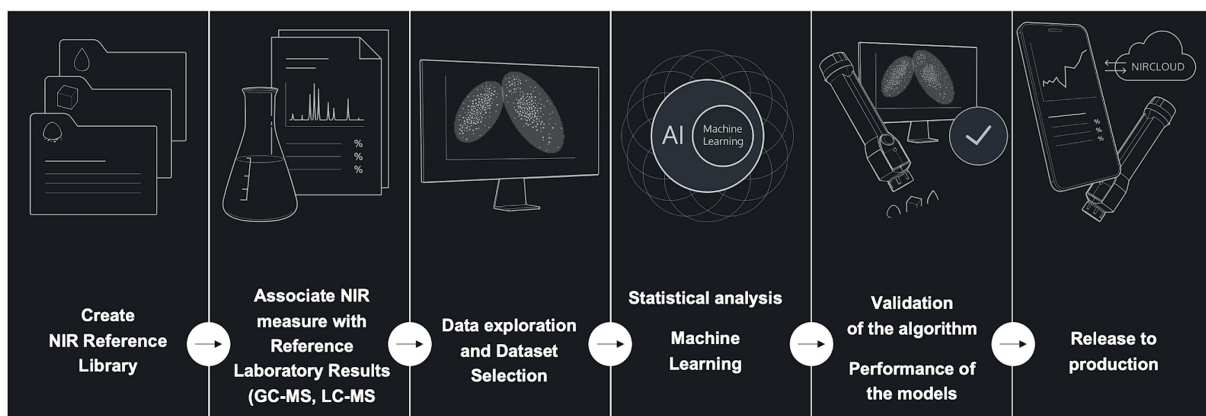
The resulting models in this study utilized an optimized pipeline comprising an ensemble of random forest and bagged trees classifiers for the classification of flowers and resins. For the quantitative analysis of



**Figure 8.** Spectra of CBD and THC flowers after pre-processing using *Savitzky–Golay* second derivative (smoothing of 5) followed by SNV. The green bands show the area where the differences are the most visible between the two cannabis types.



**Figure 9.** Spectra of THC type flowers used to train the model, pre-processing: *Savitzky–Golay* second derivative (smoothing of 5) + SNC. Color = THC potency in %.



**Figure 10.** Processes that govern the architecture of NIRLAB.

THC, an ensemble consisting of gradient boosting and k-nearest neighbors (KNN) regressors was employed.

These predictive models were trained using a library that can handle high-dimensional data well, prevents overfitting, and capture complex non-linear relationships without making assumptions of data distributions, unlike classical methods like PLS.

However, it is important to note that these models can be computationally intense, particularly with large datasets. Additionally, their 'black box' nature presents challenges in terms of interpretability. Automated machine learning systems also carry the risk of 'over-optimization', where the model becomes too closely fitted to the training data.

In comparison to classical methods like PLS and SVM, these automated techniques can provide higher accuracy by effectively modeling intricate patterns. However, they may lack simplicity in interpretation. On the other hand, PLS and SVM offer more straightforward interpretability but may struggle with complex data. The choice between these methods depends on factors such as data complexity, interpretability requirements, and available computational resources.

Once the prediction models are trained, a validation phase is conducted before the release of the new model (see *Figure 10*). This is to ensure the models are not overfitted and perform well on unknown data.

## Acknowledgements

Open Access funding provided by Universite de Lausanne.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Author Contribution Statement

*Florentin Coppey*: Writing, Reviewing, Editing Software, Conceptualization, Methodology, Investigation. *Cédric Schelling*: Reviewing, Methodology, Supervision. *Jean-Luc Veuthey*: Conceptualization, Methodology, Investigation. *Pierre Esseiva*: Writing, Reviewing, Editing, Supervision, Methodology, Conceptualization.

## References

- [1] 'World Drug Report 2021', can be found under <https://www.unodc.org/unodc/en/data-and-analysis/wdr2021.html>.
- [2] C. Citti, D. Braghiroli, M. A. Vandelli, G. Cannazza, 'Pharmaceutical and biomedical analysis of cannabinoids: A critical review', *J. Pharm. Biomed. Anal.* **2018**, 147, 565–579.
- [3] S. H. Burstein, 'The Cannabinoid Acids: Nonpsychoactive Derivatives with Therapeutic Potential', *Pharmacol. Ther.* **1999**, 82, 87–96.
- [4] A. Slosse, F. Van Durme, J. Eliaerts, N. Samyn, D. Mangels, Y. Vander Heyden, 'Analytical strategies for herbal Cannabis samples in forensic applications: A comprehensive review', *WIREs Forensic Sci.* **2023**, 5, e1479.

- [5] M. Hädener, S. König, W. Weinmann, 'Quantitative determination of CBD and THC and their acid precursors in confiscated cannabis samples by HPLC-DAD', *Forensic Sci. Int.* **2019**, *299*, 142–150.
- [6] R. Deidda, P.-Y. Sacre, M. Clavaud, L. Coïc, H. Avohou, P. Hubert, E. Ziemons, 'Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers', *TrAC Trends Anal. Chem.* **2019**, *114*, 251–259.
- [7] F. Coppey, A. Bécue, P.-Y. Sacré, E. M. Ziemons, P. Hubert, P. Esseiva, 'Providing illicit drugs results in five seconds using ultra-portable NIR technology: An opportunity for forensic laboratories to cope with the trend toward the decentralization of forensic capabilities', *Forensic Sci. Int.* **2020**, *317*, 110498.
- [8] M. Parrilla, A. Slosse, R. Van Echelpoel, N. F. Montiel, A. R. Langley, F. Van Durme, K. De Wael, 'Rapid On-Site Detection of Illicit Drugs in Smuggled Samples with a Portable Electrochemical Device', *Chemosensors* **2022**, *10*, 108.
- [9] R. F. Kranenburg, H.-J. Ramaker, S. Sap, A. C. van Asten, 'A calibration friendly approach to identify drugs of abuse mixtures with a portable near-infrared analyzer', *Drug Test. Anal.* **2022**, *14*, 1089–1101.
- [10] O. Y. Rodionova, A. L. Pomerantsev, 'NIR-based approach to counterfeit-drug detection', *TrAC Trends Anal. Chem.* **2010**, *29*, 795–803.
- [11] L. de O. Magalhães, L. C. Arantes, J. W. B. Braga, 'Identification of NBOMe and NBOH in blotter papers using a handheld NIR spectrometer and chemometric methods', *Microchem. J.* **2019**, *144*, 151–158.
- [12] E. Deconinck, C. Ait-Kaci, A. Raes, M. Canfyn, J.-L. Bothy, C. Duchateau, C. Mees, K. De Braekeleer, L. Gremaux, P. Blanckaert, 'An infrared spectroscopic approach to characterise white powders, easily applicable in the context of drug checking, drug prevention and on-site analysis', *Drug Test. Anal.* **2021**, *13*, 679–693.
- [13] C. Pasquini, 'Near infrared spectroscopy: A mature analytical technique with new perspectives – A review', *Anal. Chim. Acta* **2018**, *1026*, 8–36.
- [14] L. Sanchez, C. Filter, D. Baltensperger, D. Kurouski, 'Confirmatory non-invasive and non-destructive differentiation between hemp and cannabis using a hand-held Raman spectrometer', *RSC Adv.* **2020**, *10*, 3212–3216.
- [15] C. Duchateau, J.-M. Kauffmann, M. Canfyn, C. Stévigny, K. De Braekeleer, E. Deconinck, 'Discrimination of legal and illegal *Cannabis spp.* according to European legislation using near infrared spectroscopy and chemometrics', *Drug Test. Anal.* **2020**, *12*, 1309–1319.
- [16] R. Deidda, F. Coppey, D. Damergi, C. Schelling, L. Coïc, J.-L. Veuthey, P.-Y. Sacré, C. De Bleye, P. Hubert, P. Esseiva, É. Ziemons, 'New perspective for the in-field analysis of cannabis samples using handheld near-infrared spectroscopy: A case study focusing on the determination of  $\Delta^9$ -tetrahydrocannabinol', *J. Pharm. Biomed. Anal.* **2021**, *202*, 114150.
- [17] 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* **2017**, *173*, 318–323.
- [18] J.-M. Roussel, C. Schelling, M. Righezza, J.-L. Veuthey, 'Application of prediction intervals to the interpretation of the robustness study of a UHPLC method for the separation of cannabinoids', *J. Pharm. Biomed. Anal.* **2022**, *220*, 114977.
- [19] R. Deidda, C. Schelling, J.-M. Roussel, A. Dispas, C. De Bleye, É. Ziemons, P. Hubert, J.-L. Veuthey, 'The analysis of cannabinoids in cannabis samples by supercritical fluid chromatography and ultra-high-performance liquid chromatography: A comparison study', *Anal. Sci. Adv.* **2021**, *2*, 2–14.
- [20] R. W. Kennard, L. A. Stone, 'Computer Aided Design of Experiments', *Technometrics* **1969**, *11*, 137–148.
- [21] C. Roux, B. Talbot-Wright, J. Robertson, F. Crispino, O. Ribaux, 'The end of the (forensic science) world as we know it? The example of trace evidence', *Philos. Trans. R. Soc. London Ser. A* **2015**, *370*, 20140260.
- [22] S. A. Cole, 'Response: Forensic Science Reform: Out of the Laboratory and into the Crime Scene', *Tex. Law Rev. See Also* **2012**, *91*, 123–136.
- [23] Trans European Drug Information, 'TEDI GUIDELINES DRUG CHECKING METHODOLOGY', 2022, [https://www.tedinetwork.org/wp-content/uploads/2022/03/TEDI\\_Guidelines\\_final.pdf](https://www.tedinetwork.org/wp-content/uploads/2022/03/TEDI_Guidelines_final.pdf).
- [24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, 'Scikit-learn: Machine Learning in Python', *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- [25] T. T. Le, W. Fu, J. H. Moore, 'Scaling tree-based automated machine learning to biomedical big data with a feature set selector', *Bioinformatics* **2020**, *36*, 250–256.

Received March 30, 2023

Accepted July 7, 2023