Use Of Handheld FT-NIR Sensors To Rapidly Quantify Cannabinoids of Hemp, in situ.

Presented in Partial Fulfilment of the
Requirements for Graduation with Research Distinction

Cameron Michael Jordan
Department of Food Science and Technology
The Ohio State University

2022
Research Advisor: Dr. Luis Rodriguez-Saona
Department of Food Science and Technology
The Ohio State University

Project Mentor: Siyu Yao
Department of Food Science and Technology
The Ohio State University

The College of Food, Agricultural, and Environmental Sciences
The Ohio State University
Abstract

Hemp is a crop that has agricultural, economic, and pharmaceutic potential yet is still being researched. Hemp is known to produce over 100 phytocannabinoids, including cannabidiol (CBD) and $\Delta^9$-tetrahydrocannabinol ($\Delta^9$-THC). Hemp is a plant that must contain less than 0.3% THC w/w, per the 2018 Farm Bill. Current analytical methods of High-Performance Liquid Chromatography- tandem mass spectrometry (HPLC-MS/MS), which is a selective and sensitive method, but is cost inefficient, time-consuming, and requires complex analysis. Fourier Transform Near Infrared (FT-NIR) is a non-destructive, non-invasive method with the potential to be added to inline production settings. The analysis of FT-NIR is almost instantaneous compared to HPLC-MS/MS with a fraction of the cost.

Hemp samples were scanned using a handheld FT-NIR scanner. Cannabinoids were extracted from hemp inflorescence and analyzed by HPLC-MS/MS. The two data matrices were correlated by Partial Least Squares Regression (PLSR) and a prediction model was generated. The prediction model allowed for the differentiation between drug-type (THC>0.3%) and fiber-type (THC<0.3%) hemp by their content of THCA (.27-.80%) and THC (.021-.056%), and the ability to quantify 4 different cannabinoids in a single measurement, including CBDA (7.7 – 20.7%). The prediction model yielded a small standard error of cross-validation and high correlation coefficient of cross-validation of Rcv>0.95. This experimentation shows the use of a small handheld scanner to provide a faster and cheaper analysis of a very heavily regulated crop with relatively no standardized methodology of analysis. This technology will benefit hemp growers and analysts of hemp material greatly.
1. Introduction

Hemp (*Cannabis sativa* L.) is a crop of increasing agricultural, economical, and pharmaceutical importance.¹ Researchers are trying to determine if drug type cannabis and hemp (fiber type) are of the same species or two different species in the same family Cannabaceae.² Hemp must have <0.3% w/w Δ⁹-Tetrahydrocannabinol (Δ⁹-THC, THC) to be cultivated per the Agricultural Improvement Act of 2018 (2018 Farm Bill).³ The recent discovery of the endocannabinoid system (ECS) as a regulator of cellular homeostasis has pushed cannabinoids such as cannabidiol and Δ⁹-tetrahydrocannabinol into a central position from the medical and drug research point of view. This role of the ECS potentially opens many doors to preventing and treating many diseases related to the ECS, such as Alzheimer’s disease, multiple sclerosis,⁴ obesity⁵, and epilepsy⁶, among many others. As a result of this increased clinical knowledge, Cannabis-derived products are increasingly being recognized as substances with a demonstrated medicinal value, and this is confirmed by the fact that its production and use are becoming legalized and regulated in many countries.⁷ Breeders of medicinal Cannabis varieties carry on a constant process of selective breeding in order to develop varieties aiming to improve the yield of certain cannabinoids and other secondary metabolites of interest for the pharmaceutical industry, or even to guarantee the absence or reduction of certain unwanted molecules, as a practical example can be mentioned the reduction of Δ9-THC in varieties that are considered fiber type and intended to be used for controlling convulsive seizures in epileptic children.⁸ The high intrinsic variability present in Cannabis plants (different brands, varieties, chemotypes, and gender) leads to great difficulty in obtaining a classification standard.⁹ Furthermore, conditions during the growth and storage of cannabis, such as environmental factors of cultivation (weather and altitude of cultivated area), the development stage of the plant at harvest time, as well as genetic
Accurate and reliable determination of THC and CBD is of great economic importance to stakeholders in the cannabis supply chain. The most common approaches consist of chromatographic methods, including High-Performance Liquid Chromatography followed by Mass Spectrometry (HPLC-MS) and Gas Chromatography coupled to Mass Spectrometry (GC-MS). Although mass spectrometry (MS) analysis provides selectivity and specificity for screening cannabinoids, it requires costly instrumentation, labor-intensive and complex sample pretreatment, well-trained technicians to operate the instrumentation, and is less amenable to be implemented for quality control at breeding and manufacturing facilities. This is limiting the growth of the hemp industry because the cost of the analytical instrumentation such as tandem HPLC/MS/MS is prohibiting in-plant QC analysis. Thus, most companies contract with accredited labs that have a 7-10 working day turnaround to provide results. Furthermore, it can be common for testing laboratories to overstate cannabinoid concentrations. Depending on geographic region, over 50% of edible products were labeled as having higher cannabinoid concentrations than the analysis stated, meaning that third-party labs often are pressured to increase THC and CBD content, this is also the case with whole hemp inflorescence. This could allow consumers to purchase and ingest less than paid for, reflecting negatively on producers and testing laboratories. The American Herbal Pharmacopeia highlights the necessity of standardizing the procedures behind Cannabis species analysis to provide an accurate analysis, considering the pharmaceutical nature of the phytocannabinoids.

The industry is therefore in need of an alternative method that can provide data in a timely and economic manner, so they can comply with legal regulations and minimize costly recalls, loss
of reputation, and costs associated with litigation.\textsuperscript{14} In addition, a rapid method that could be used in-house would allow companies to make decisions that could improve the quality and safety of their products.\textsuperscript{17} Optical technology is rapidly developing, and instruments are already available commercially as portable, hand-held, and micro-devices that can be used when it is not practical or economical to use the more sophisticated and costly instruments used in research laboratories.\textsuperscript{18}

Miniaturation of near-infrared (NIR) spectrometers into commercially available systems has occurred within the last years, driven by developments in micro-electro-mechanical systems (MEMS) production.\textsuperscript{19} The developments of optical components, wavelength selectors and detectors that can be thermoelectrically air-cooled have enabled the miniaturization of spectrometers without sacrificing performance.\textsuperscript{19} Advantages of approaches based on NIR spectroscopy include low cost, small size, compactness, robustness, high throughput, and ease of operation for in-field routine analysis.

Fourier Transform- Near Infrared (FT-NIR) Spectroscopy collects the interference in the Near Infrared region of the electromagnetic spectrum\textsuperscript{20}, which can then be combined with chemometric methods to quantify the amount of a compound, in this experiment THC and CBD in hemp. FT-NIR provides a rapid and highly sensitive alternative testing method, that is more accessible to growers and producers of hemp products.\textsuperscript{21}

Although previous studies have reported the successful use of benchtop NIR spectroscopy and PLS regression to determine different cannabinoids in \textit{Cannabis sativa} L. plant material\textsuperscript{22}, to date miniaturized devices lack of accuracy and do not permit to achieve the suitable sensitivity to avoid true positive and false positive response. Consequently, miniaturized and one-touch devices providing the identification and quantification of cannabinoids in a completely automated and accurate platform, are a necessity.\textsuperscript{23} Thus, our objectives focus on leveraging a platform sensor
technology that combines near-infrared (NIR) spectroscopy and Partial Least Squares Regression (PLSR) algorithms to quantify CBD and THC levels, determining optimal hemp sample preparation methods and their impact on measurement accuracy, and characterizing and verifying the performance metrics of the sensor in terms of meeting the user’s needs.

2. Materials and Methods

2.1 Sample Information

Unhomogenized hemp inflorescence of variety ‘Tangerine’ samples (n=21) were obtained from the OARDC Weed Lab at The Ohio State University. Another set of samples (n=4) was obtained from a third-party industrial hemp grower in Ohio. Samples were then homogenized in a coffee grinder and sifted through a strainer to reduce the particle size of the hemp material.

2.2 FT-NIR Method Development

First, the unhomogenized hemp inflorescence was scanned using a NeoSpectra Scanner by SI-Ware (Si-Ware systems, Cairo, Egypt). The scanner used for analysis is shown in Figure 1, showing how the system can be used conveniently outside of a laboratory environment. The controls and settings for the Scanner were monitored using the NeoSpectra Collect smartphone application for iOS. Reference scans were taken periodically between samples using a reflective standard. The material scan was set to 10-second exposure times in duplicate measurements. During the reference scan, near-infrared light (7,000 to 4,000 cm\(^{-1}\)) was reflected from the material and the interference was measured, allowing for a spectrum to be obtained for each sample. After the whole inflorescence was scanned, the inflorescence was homogenized, and the particle size was reduced. Spectra were collected using the same scanner settings as the unhomogenized material. The instrument was cleaned with 70% v/v ethanol between samples to remove any unintended residual analyte.
2.3 uHPLC-MS/MS Reference Analysis

After the FT-NIR scans were completed, the inflorescence was prepared for ultra-high performance liquid chromatography-tandem mass spectrometry (uHPLC-MS/MS). Cannabidiol (CBD), cannabidiolic acid (CBDA), Δ⁹-Tetrahydrocannabinol (Δ⁹-THC, THC), and Δ⁹-tetrahydrocannabinolic acid (Δ⁹-THCA, THCA) were extracted using methanol: water 4:1 v/v. 0.5 grams of homogenized hemp material was added to a 50 mL centrifuge tube along with 12 mL of LC-MS grade methanol (ThermoFisher Scientific Inc., Waltham, MA) and 3 mL of LC-MS grade water (ThermoFisher Scientific Inc., Waltham, MA). The centrifuge tube was vortexed for 2 minutes each and then ultrasonicated for 45 minutes using a ThermoFisher FS30H sonicator.
After sonication, the tube was centrifuged at 4500 rpm for 5 minutes using a Sorval ST8 Centrifuge (ThermoFisher Scientific Inc., Waltham, MA), and 1 mL aliquots were taken using a syringe. The aliquot was then filtered through a 0.2 µm syringe filter. 10 µL filtered sample was added to an HPLC vial along with 960 µL LC-MS grade methanol and 30 µL internal standard, (-)-Δ⁹-THC-D₃ (100 µg/mL) (Sigma-Aldrich, St. Louis, MO). The uHPLC used for quantification was a Nexera-i LC-2040C 3D (Shimadzu Corp., Kyoto, Japan) and the MS/MS used was an LCMS-8040 (Shimadzu Corp., Kyoto, Japan).

2.3.1 Chromatographic Conditions

Analytes were separated utilizing a reverse-phase Raptor ARC-18 column (2.7 µm, 150 × 3.0 mm dimensions) with an attached guard column (Restek, Bellefonte, PA) in a temperature-controlled 30°C environment. Solvents used for separation included (1) Solvent A, water with added 0.1% formic acid, and 4 mM ammonium formate, and (2) Solvent B, acetonitrile with added 0.1% formic acid. The solvent flow rate was set to 0.4 mL/min for the LC component and nebulizing gas flow rate of 3 L/min for the tandem mass spectrometers. The solvent gradient used for the separation was 0–15 min, 75% B; 15–18 min, 75–100% B; 18–20 min, 100–75% B; 20–23 min, 75% B. An electrospray ionization (ESI) technique was used for the MS parameters. 10 µL of analyte solution was injected into the column. Each sample run was completed in 20 minutes, with a 3-minute rinse period between each sample.

2.3.2 Quantification of Cannabinoids

Cannabinoids were quantified by taking the peak area under the total ion chromatogram (Figure 2). A calibration curve was completed each time with CBD, CBDA, Δ⁹-THC, Δ⁹-THCA, and IS (ThermoFisher Scientific Inc., Waltham, MA). The ratio of the analyte to IS peak area was used to calculate the concentration of the cannabinoids using the calibration curve. Cannabinoid
concentrations were obtained in triplicate measurements and averaged to obtain a value for each sample.

![Representative chromatogram of the analytes quantified using uHPLC-MS/MS](image)

**Figure 2.** Representative chromatogram of the analytes quantified using uHPLC-MS/MS

### 2.4 Data Analysis

Spectral data and uHPLC-MS/MS reference data were correlated using Partial Least Squares Regression (PLSR). Data analysis was done using Pirouette® (InfoMetrix Inc., Bothell, WA). The important regions for differentiation between samples were determined using a second derivative regression vector analysis. Regions of moisture content were filtered out of the analysis to reduce noise in the algorithm. PLSR projects the variables into a new plane, where the correlation is done from the projected plane. PLSR finds factors for both the reference data and the spectral data, which increases the signal to noise ratio by reducing the angles between the spectral and uHPLC-MS/MS eigenvectors. The reduction of the angles between the two eigenvectors by rotation of the vectors allows for data with more planarity, reducing the noise and increasing the correlating power of the matrices used in regression analysis.24 By reducing the
noise from the spectral data and reference data, this also decreases the potential for out-of-plane data points to skew the model. To validate the model, a 80:20 training: validation split was utilized. This means that for every 100 samples, 20 of the spectra are not included in the model and are instead cross validated. The cross validation shows that the prediction algorithm can accurately quantify the cannabinoids using the spectral data and confirming the prediction with the reference data. Ideally, the validation set should follow the same regression as the training model.

3. Results and Discussion

3.1 Spectral Data

Spectral data obtained from the handheld scanner for both powdered and whole inflorescence, represented in Figure 3, show that the scanner generates reproducible data, and samples provide similarly shaped spectra. The replicable spectral curves make the generated prediction model more accurate, decreasing the error in cross-validation with the reference data. The regions shown are important NIR regions of important biomolecules such as lipids, proteins, carbohydrates, and water. Although these molecules are not quantified in this study, they will still have matrix effects on the quantification of target analytes. Spectral data is consistent with that of literature benchtop systems, as well as the spectral data on powdered samples from a similar portable prototype equipped with the Micro spectrometer (Neospectra) using a 3-LED source. The focus of this experimentation was to use a 5-LED source in a commercial NIR sensor equipped with an Indium-Gallium-Arsenide (InGaAs) detector and determine if there were any correlative differences between the whole and powdered inflorescence. The difference between the two portable systems is that the 5-LED source should increase the signal being transmitted to the detector allowing for a higher signal-to-noise (SNR) ratio. The increase in SNR is due to higher amounts of light being reflected off the sample and being received in the interferometer.
3.2 Cannabinoid Concentration Prediction Model

The numerical results in Table 1 show the correlation and low error associated with making a prediction from the spectral data. The powder and whole inflorescence data have Rcv $\geq 0.95$, showing a strong linear correlation between the spectral and reference data sets. The prediction model shows that 4 cannabinoids can be simultaneously quantified in a single measurement from a complex matrix. The range of cannabinoids allows for an accurate prediction. CBDA has the largest range of % w/w, making it the most predominant cannabinoid in the analyzed hemp samples.

Table 1. Table of cannabinoid ranges and PLSR prediction data

<table>
<thead>
<tr>
<th>Sample Presentation</th>
<th>Analyte</th>
<th>Range (%)</th>
<th>Factors</th>
<th>SECV</th>
<th>Rcv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flower</td>
<td>CBDA</td>
<td>7.7 - 20.7</td>
<td>9</td>
<td>0.89</td>
<td>0.95</td>
</tr>
<tr>
<td>Powder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flower</td>
<td>THCA</td>
<td>0.27 - 0.80</td>
<td>10</td>
<td>0.032</td>
<td>0.98</td>
</tr>
<tr>
<td>Powder</td>
<td></td>
<td></td>
<td></td>
<td>0.03</td>
<td>0.98</td>
</tr>
<tr>
<td>Flower</td>
<td>CBD</td>
<td>0.27 - 0.87</td>
<td>6</td>
<td>0.23</td>
<td>0.97</td>
</tr>
<tr>
<td>Powder</td>
<td></td>
<td></td>
<td></td>
<td>0.21</td>
<td>0.98</td>
</tr>
<tr>
<td>Flower</td>
<td>THC</td>
<td>0.021 - 0.056</td>
<td>8</td>
<td>0.035</td>
<td>0.95</td>
</tr>
<tr>
<td>Powder</td>
<td></td>
<td></td>
<td></td>
<td>0.016</td>
<td>0.96</td>
</tr>
</tbody>
</table>
The data displayed in Table 1 shows the linearity of the cannabinoid analytes. CBDA was the analyte with the largest range of quantification, which is consistent with literature values in cultivated hemp (CBD content 6.4-25.4%).\textsuperscript{26} The lowest range of concentrations was Δ\textsuperscript{9}-THC, the decarboxylated form of Δ\textsuperscript{9}-THCA. The low concentration of the neutral THC molecule shows that the cannabinoid that needs more monitoring is the acidic THCA molecule, with concentrations being closer to the 0.3% limit. The acidic cannabinoids are biosynthesized directly from cannabigerolic acid (CBGA) and then later decarboxylated through ambient heat and light, meaning that the acidic forms, CBDA and Δ\textsuperscript{9}-THCA, will be more predominant than CBD and Δ\textsuperscript{9}-THC.\textsuperscript{27} Regarding the factors, the lowest number of factors that explain most of the variance is desired.\textsuperscript{28} If too many factors are used, the prediction model could include noise that is not beneficial for the regression. Flower samples used more factors than the powder samples, this could be due to the particle size difference allowing for less light scattering when the particle size is smaller in a powdered form. A higher amount of light scattering allows for less signal to be obtained by the instrument, increasing the prediction error.\textsuperscript{29} The bands also exhibit a lower resolution by loss of light not reflected into the detector. The powder samples had less standard error than the flower samples, which is expected. Although the powder samples have less error and higher Rcv values than their respective whole flower samples, the whole flower still provides an adequate signal for a prediction model. The use of less factors for the prediction model allows for less noise to be included for the cross validation. The powder samples use less factors than the whole inflorescence, providing less noise allows for a more accurate prediction algorithm. 

Our results exhibit excellent signal-to-noise ratios and good linearity in predicted vs. reported CBD levels (Figure 4) supporting development of field-deployable sensor devices. The cannabinoid ranges determined by a benchtop NIR system displays data also show good linearity
along the ranges of CBD (0.01-13.3%) present in hemp, but having samples with higher THC content (average THC= 4.0%) than applicable for the analyzed data set. Although, another study using a benchtop system of industrial hemp displays similar ranges of cannabinoids (total THC: 0.06-.16%; total CBD: 2.2-5.4%), the portable NIR system used in current experimentation displays a better correlation with less latent variables or factors (LV, factors=7).

Figure 4. PLSR plots for CBDA and Δ⁹-THCA analytes from powder and whole hemp inflorescence.

Growers are also looking for a non-destructive analysis, meaning that powdering the sample does not allow for the analyzed sample to be returned to the product flow. Regarding the mislabeling of hemp products, the scanner could allow for a standardized analytical method that
removes the human error of performing an extraction or analysis of cannabinoids. The scanner does not require a trained HPLC-MS/MS technician, which reduces the human handling of the analysis. The analysis can be completed in 10 seconds by an untrained professional such as a grower of hemp, or a quality assurance technician.

The samples used in this study were grown in a greenhouse located on Ohio State’s campus. The samples were also direct cuttings or clones of the same mother plant. This means that there could be little genetic variation in the samples. This could skew the model for this specific chemovar of hemp. Ideally, for a more applicable and accurate model, samples from multiple geographic regions should be analyzed. Multiple chemovars should be analyzed, as well as drug type cannabis to differentiate drug type and fiber type cannabis more accurately. A larger range of analyte concentrations would also be beneficial for the model.

5. Conclusion

The field of hemp and cannabinoid research is still in its younger years. By applying a faster and alternative testing method with modern technology, quality control becomes easier for production facilities. FT-NIR allowed for prediction models for neutral $\Delta^9$-THC (range: 0.021-0.056%) and CBD (range: 0.27-0.87%) as well as the two acidic forms THCA (0.27-0.80%) and CBDA (7.7-20.7%), providing an accurate (Rcv≥0.95) prediction model with low standard error of cross validation. The FT-NIR method development and prediction model provides a rapid and cost-effective analysis to an expensive and highly regulated industry. Hemp and hemp products have strict legal regulations, yet analytical labs often mislabel due to the lack of standardized methodologies and pressure from growers to provide higher cannabinoid content profiles. FT-NIR
combined with chemometrics provides a selective and sensitive method that rivals the accuracy of the reference uHPLC-MS/MS.

With the potential to be added to a product flow, the rapid FT-NIR sensors could greatly impact the analysis of hemp. Other cannabinoids as well as terpenes have the potential to be quantified using FT-NIR combined with chemometric methods. The scanner also has the potential to quantify important biomolecules like lipids, protein, and carbohydrate content- which are important for food processors or those using hemp in the food industry. The handheld scanner could be used to determine the cannabinoid concentrations of the hemp inflorescence over the growing period, meaning an almost instantaneous decision of the plant’s future could be determined faster than the traditional HPLC-MS/MS.

Acknowledgments

Thank you to the OARDC Weed Lab and Cedar Valley Farms for providing samples for my research. I would also like to thank all my Food Science and Technology peers and professors for helping me prepare for my presentations. FoodSURE was integral to the beginning of my research journey and graduate school preparations. A special thank you to the coordinators of the program, Abigail Sommer and Fenfen Tang. To all my lab mates: the advice you gave me allowed me to get up to par in being a researcher. Most importantly, I would like to thank my family for supporting me through my educational endeavors and being a rock-solid support system.

Reference:


