KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI

COLLEGE OF SCIENCE

FACULTY OF PHYSICAL AND COMPUTAIONAL SCIENCES

DEPARTMENT OF PHYSICS

CHARACTERIZATION OF SOME LOCAL GHANAIAN HERBAL ALCOHOLIC BEVERAGES (BITTERS) USING RAMAN SPECTROSCOPY

BY

ABAYIE DEBORAH | AMETEPEH EMMANUEL | RODGERS EPHRAIM

(SUPERVISORS: DR. AKYANA BRITWUM, DR. MICHAEL K.E. DONKOR, & DR. GEORGE O. DWAPANYIN)

AUGUST, 2024

CHARACTERIZATION OF SOME LOCAL GHANAIAN HERBAL ALCOHOLIC BEVERAGES (BITTERS) USING RAMAN SPECTROSCOPY

A REPORT SUBMITTED TO THE FACULTY OF PHYSICAL AND COMPUTATIONAL SCIENCES, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI, IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF BSC PHYSICS

BY

ABAYIE DEBORAH | AMETEPEH EMMANUEL | RODGERS EPHRAIM

(SUPERVISORS: DR. AKYANA BRITWUM, DR. MICHAEL K.E. DONKOR, & DR. GEORGE O. DWAPANYIN)

AUGUST, 2024

DECLARATION

We hereby declare that this thesis is the result of our research undertaken by ABAYIE DEBORAH, AMETEPEH EMMANUEL and RODGERS EPHRAIM towards the award of BSc. Physics Degree in the Department of Physics, KNUST, and except for where references have been made, no part of this research work has been submitted for the award of any other degree.

Student Name	**Signature**	**Date**
Abayie Deborah	.....	20/09/2024
Ametepeh Emmanuel	.....	20/09/2024
Rodgers Ephraim	.....	20/09/2024
Dr. Akyana Britwum **(Supervisor)**	.....	20/09/2024
Dr. Michael Kweku Edem Donkor **(Supervisor)**	.....	20/09/2024
Dr. George Okyere Dwapanyin **(Supervisor)**	.....	20/09/2024
Prof. Francis Kofi Ampong **(Head of Department)**	.....	20/09/2024

DEDICATION

To the Almighty God, our lovely families, all lecturers and staff at the Department of Physics, our colleagues at the Department of Physics, and our esteemed advisor, Dr. George Dwapanyin, for your unwavering guidance and support throughout this journey.

ACKNOWLEGEMENT

Throughout the course of this study, we have been supported and guided by the expertise of many individuals. To start with, we would like to acknowledge our dedicated supervisors and express our profound appreciation for all they did to help bring this work to fruition: Dr. Akyana Britwum, Dr. Michael K.E. Donkor, and Dr. George O. Dwapanyin.

We extend our gratitude to the Department of Physics for providing access to facilities and resources that were essential for completing this work. A special thanks is due to the University of St Andrews, UK, for loaning us the Raman Spectrometer, which was crucial for our research.

Finally, we would like to thank God, our families, and our supportive friends and colleagues for their unwavering support throughout the course of this work.

ABSTRACT

In Ghana, local alcoholic bitters are popular due to their perceived medicinal properties, yet limited scientific research has explored their chemical composition, particularly concerning ethanol, the main alcohol ingredient. This study employs Raman spectroscopy, a non-destructive analytical technique recognized for identifying molecular compositions through vibrational modes, to analyse the ethanol content in 13 well-known Ghanaian bitters brands sourced from different regions of Kumasi. Using a portable Raman spectrometer, detailed spectra of the bitters were obtained, with a standard reference Raman ethanol spectrum serving as a calibration baseline. Each sample, prepared by extracting 5 mm of the bitter into borosilicate glass vials to prevent cross-contamination, was analysed, and the spectral data processed using OceanView software. Further analysis involved Principal Component Analysis (PCA) and Partial Least Squares Regression (PLSR). PCA revealed that fluorescence and ethanol were the primary factors influencing spectral differences among the bitters, indicating the diversity in their composition. PLSR demonstrated a strong positive correlation, confirming the consistency of ethanol content across different brands. Moreover, Leave-One-Out Cross-Validation (LOOCV) verified the unique spectral signatures of each brand, achieving 100% classification accuracy. The results affirm that Raman spectroscopy, combined with PCA, is a powerful tool for characterizing the chemical composition and quality of Ghanaian bitters, supporting their legitimacy and consistency. The study also validates current labelling practices concerning alcoholic content (ABV), ensuring regulatory compliance and enhancing consumer safety in the herbal alcoholic beverage industry.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

The Ghanaian food and beverage industry has historically played a crucial role in the country's economy, having a major impact on both the general nutrition of the population and the country's overall economic development. The beverage industry is one of the most active in this area as it is divided into many different companies and is mostly structured along production and distribution lines. The manufacture of alcoholic beverages and carbonated soft drinks (CSDs) is particularly notable in this field since each adds distinctively to the market environment (Sector Industry Study Ghana Food and Drink Industry by Strategy and Research Department, 2019).

Ghana has a sizable market for alcoholic beverages, with locally produced alcoholic beverages serving as the main industry drivers. A greater percentage of alcoholic drinks consumed in Ghana are manufactured locally, showing the significance of locally made alcoholic drinks in satisfying Ghanaian customers' tastes and preferences. The lesser percentage is made up of different foreign alcohol brands, which diversifies the products available on the market. Ghanaian herbal alcoholic bitters which fall in the Spirits category are the most popular alcoholic beverages consumed locally. These bitters are preferred for their flavor, cost-effectiveness in comparison to other liquors, and potential health advantages. Typically, they are made from a blend of herbs, roots, barks, and other botanical materials (Ghana: Ghana Alcoholic Beverages Report | USDA Foreign Agricultural Service, 2024) . The market for bitters has grown globally as a result of increased consumption; it was estimated to be worth USD 10460 million in 2019 and is projected to rise at a compound annual growth rate (CAGR) of 3.2% from 2021 to 2026, when it will reach USD 13070 million (Global Bitters Market – Market Reports World, 2021).

Production of bitters from a variety of sectors has increased in response to the growing demand for both local and worldwide Ghanaian bitters. But the sudden increase in demand has raised concerns about the authenticity and safety of the bitters on the market (MyJoyOnline, 2016). These bitters are prized for their distinctive taste of ethanol, which is the only deemed safe alcohol for human consumption and specifically herb blends which each add different flavors and possibly health benefits.

On the other hand, there may be serious health hazards if substandard bitters are manufactured using industrial alcohols or by using poor distillation methods and are misrepresented as commercial products (Lapierre et al., 2024) . Different combinations of ethanol and herbal ingredients used in production determine the uniqueness of each brand of herbal bitters. More exact and advanced techniques are required because the existing methods for evaluating and guaranteeing the quality of these bitters frequently fall short of fully capturing the complexity of their chemical makeup. In a market where there is fierce competition this kind of analysis is essential for upholding quality control as well as each brands authenticity and identity.

Ghanaian herbal bitters do not have a clear classification system like other alcoholic beverages such as beer and whisky which are categorized according to their Alcohol by Volume (ABV) and other unique characteristics (Balekundri & Mannur, 2020). To ensure the safety quality and authenticity of alcoholic beverages sophisticated laboratory techniques like nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry, liquid chromatography (LC) and gas chromatography (GC) are commonly used. But these procedures can be costly and time-consuming especially when dealing with the intricate chemical compositions of herbal bitters (Tabago et al., 2021).

The procedure of preparing samples typically requires several processes, including the process of extraction, which is not appropriate for continuous online monitoring. Additionally, exposing end goods to be analyzed has a direct effect on concentration and integrity of the final product. This limits the practicality of these methods for routine quality control, especially in large-scale production environments where online, real-time monitoring is preferable (Andre de Villiers et al., 2012) .

To improve upon this limitation, Raman spectroscopy is employed as an analytical tool.

The whole field of Raman spectroscopy is founded on the Raman effect, which is named after one of its discoverers, an Indian scientist C. V. Raman, who along with K. S. Krishnan identified the effect in organic liquids in 1928 (D.J. Gardiner, 1989) . Spectroscopy on the other hand is the scientific study of how electromagnetic radiation interacts with matter. Various spectroscopic techniques rely on phenomena such as emission, absorption, fluorescence or scattering (Barth, 2007). Raman spectroscopy is used in research to provide both qualitative and quantitative studies for characterizing a wide range of substances (Andrea & Filippo, 2021). While quantitative analysis establishes the analyte concentration inside the sample, qualitative analysis seeks to identify the constituents of the sample (Dr. Shinde, 2023) . Raman spectroscopy stands as a unique technique enhanced by augmented machine learning and a label-free method to gain molecular information from samples. It serves as a powerful tool for analyzing alcoholic samples without harming or contaminating the sample itself (Luo et al., 2022).

Raman spectroscopy as an analytical technique is a non-destructive chemical analysis, where laser light beam is shone on a sample, and by measuring the way light interacts with the sample's molecules, it generates a unique "fingerprint." This fingerprint is a spectrum composed of peaks corresponding to the specific types of molecules present in the chemical composition of the sample. This fingerprint allows scientists to not only identify the various

molecules by qualitative analysis but also determine their relative abundances by quantitative analysis (Zhang et al., 2023).

Raman spectroscopy has been deployed in versatile fields such as material science, where it helped researchers to utilize it to characterize and identify nanomaterials like graphene and carbon nanotubes, analyzing their structure to understand their properties. In polymer science, Raman spectroscopy helps determine the composition, crystallinity and orientation of polymers, aiding in both research and quality control processes (Wu et al., 2018).

Similarly, the field of forensics and security also leverages Raman spectroscopy's power, crime scene investigators can employ it to analyze trace evidence like paint chips, fibers, and even explosives residues, differentiating materials with exceptional precision (Marisia et al., 2019) . Gemologists can utilize Raman analysis to distinguish genuine gemstones from synthetic ones based on their unique Raman fingerprint (Pandiscia, 2022) .

Beyond its application in vast fields, Raman spectroscopy is also employed as a guardian of both safety and quality control in alcoholic beverages (McHugh, 2021). Raman spectroscopy is used in ensuring consumer safety, checking the lacing or adulteration of ethanol, determining the concentration of the ethanol and also gauge the overall constituents of alcoholic beverages known as the congeners (Cleveland et al., 2007).

Unlike other analytic techniques which are traditional laboratory-based, Raman spectroscopy, provides much more portable versions for sample analysis. This eliminates the need for transporting samples to a dedicated lab, streamlining the process and enabling real-time assessments to the authenticity of the products been produced. Raman spectroscopy is employed by means of using a Raman spectrometer which generates data from samples as light interacts with these samples (*BRAVO Handheld Raman Spectrometer | Bruker*, 2022).

The vast amount of data generated by Raman spectrometer helps to provide a detailed chemical fingerprint of the sample being used. This data is handled by Machine learning (ML)

models which have emerged as powerful tools for this purpose. By applying the machine learning algorithms, analysis is made on the large amount of spectral data identifying subtle patterns and relationships that might otherwise go unnoticed. This allows for gathering of valuable information about the sample's composition and structure for more accurate predictions (Qi et al., 2023).

These machine learning techniques, may be divided into many sub-methods such as discriminant function analysis (DFA), partial least-squares followed by a discriminant analysis (PLS-DA), and linear discriminant analysis (LDA). Also, models based on artificial neural networks (ANNs) like support vector machines (SVM), multilayer perceptron's (MLP), and convolutional neural networks (CNNs) are used. In addition, models based on regression trees, such as random forests (RF) and classification and regression trees (CART) are employed (Lussier et al., 2020).

However, the machine learning techniques used in this research are Leave-One-Out Cross-Validation (LOOCV), Partial Least Squares Regression (PLSR) and Principal Component Analysis (PCA). Principal Component Analysis (PCA) is a multivariate technique that operates in an unsupervised manner and is used to analyze the inherent structure of the data. PCA reduces the dimensionality of the data set by finding an alternative set of coordinates, the principal components (Mobili et al., 2010).

The Principal Component Analysis (PCA) results usually reveal the distinct cluster patterns corresponding to different types of alcohol, demonstrating the technique's effectiveness in differentiating between various alcoholic beverages. The Principal Component Analysis (PCA) highlights the specific Raman shifts associated with key molecular features, such as ethanol and other chemical compounds. These components are crucial for identifying lacing substances, as adulterated samples will exhibit unique spectral signatures that deviate from the normal (Chi et al., 2023).

Partial least squares (PLSR) regression on the other hand is a technique that reduces the predictors to a smaller set of uncorrelated components and performs least squares regression on these components, instead of on the original data. Partial least squares (PLSR) regression models are developed to correlate the Raman spectra to quantify alcohol content and different substances based on the spectral data (Wold et al., 2001).

Also, Leave-one-out Cross Validation or LOOCV is a type of cross-validation method that involves leaving out one sample from the training set and using the remaining samples to train the model (Brownlee, 2020a) . To guarantee the Partial least squares (PLSR) regression models' robustness and generalizability, Leave-one-out Cross Validation (LOOCV) is used for validation. Every iteration uses one observation as the test set and the rest of the observations as the training set. The dependability of the Partial least squares (PLSR) regression models is then verified by averaging the performance measures throughout all iterations. The validity of the models is then demonstrated by a consistency in the outcomes across various subsets of the data (Lestyo et al., 2016).

The integration of Principal Component Analysis (PCA), Partial least squares (PLSR) regression, and Leave-one-out Cross Validation (LOOCV) with Raman spectroscopy offers an effective framework for the accurate identification and characterization of the unique chemical signatures of various herbal bitters brands, with a focus on both the ethanol content and the specific herbal compounds that define each product. This methodology does not only improve and enhance the quality control measures in the alcohol industry but also ensures consumer safety and trust in the products they purchase.

1.2 RESEARCH PROBLEM

Ghanaian herbal bitters are highly favored by consumers for their perceived health benefits and deep cultural significance, with many believing that the herbs and botanicals used in these drinks offer various therapeutic effects. Despite their widespread popularity, there is limited scientific research on the chemical composition of these bitters, especially regarding the ethanol content and the specific herbal compounds that give each brand its unique characteristics.

The distinctiveness of each herbal bitters brand is shaped by the different combinations of ethanol and herbal ingredients used in production. However, the current methods for analyzing and ensuring the quality of these bitters fall short in capturing the full complexity of their chemical makeup. There is a clear need for a more precise and sophisticated approach to analyzing and differentiating these products, which is essential not only for quality control but also for preserving the authenticity and identity of each brand in a competitive market.

The central research problem, therefore, is to establish a comprehensive analytical approach that can accurately identify and characterize the unique chemical signatures of various herbal bitters brands, with a focus on both the ethanol content and the specific herbal compounds that define each product. Raman spectroscopy, combined with advanced statistical techniques like Principal Component Analysis (PCA) and Partial Least Squares Regression (PLSR), presents a promising solution to this challenge. This research aims to investigate the effectiveness of these methods in providing a detailed and reliable analysis of Ghanaian herbal bitters.

1.3 AIM OF RESEARCH

This research thesis aims to characterize local Ghanaian herbal alcoholic beverages known as bitters using Raman spectroscopy.

The specific objectives of the research are to;

- Obtain samples of commonly consumed bitters.
- Perform Raman spectroscopy on the samples to obtain spectra.
- Analyze Raman spectra to characterize the bitters.
- Determine if there are any differences in the products sold by different manufacturers.
- Determine if the alcoholic content stated on the products are consistent with the actual ABV.

1.4 JUSTIFICATION

The consumption of herbal bitters is widespread in Ghana, with a lot of consumers valuing these beverages for their medicinal properties and unique flavors derived from various herbs and botanicals. However, to maintain consumer trust and ensure product integrity, it is crucial to thoroughly understand and consistently monitor the chemical composition of these products. Specifically, accurately measuring the ethanol content and identifying the distinct herbal compounds in each bitters brand are essential for quality control and preserving the uniqueness of each product.

Given the subtle differences in herbal ingredients that contribute to each brand's identity, Raman spectroscopy emerges as an ideal tool due to its ability to provide detailed molecular insights with minimal sample preparation. By integrating Raman spectroscopy with PCA and PLSR, this research aims to develop a method that can precisely characterize the ethanol content and the specific herbal components of various bitters brands.

This approach not only ensures accurate labeling and authenticity of herbal bitters but also builds consumer confidence by guaranteeing that the products they purchase adhere to high standards of quality and consistency. Additionally, the ability to differentiate between brands based on their unique chemical signatures will help protect the intellectual property and market position of each brand, safeguarding against imitation or adulteration.

CHAPTER 2

THEORETICAL BACKGROUND

2.1 OVERVIEW

This chapter aims to explain the fundamental principles and applications of Raman spectroscopy, providing the necessary context for its use in analysing complex mixtures like herbal alcoholic beverages. The chapter will end with an introduction and discussion about the data analysis techniques used in this study, namely Principal Component Analysis (PCA), Leave-One-Out Cross-Validation (LOOCV) and Partial Least Squares Regression (PLSR), which are essential for the accurate interpretation of spectroscopic data.

Raman spectroscopy is a highly effective analytical method that investigates molecular vibrations by measuring the inelastic scattering of monochromatic light. It is widely utilized to analyse chemical structures, physical properties, and molecular composition. It identifies substances by their unique spectral patterns (often referred to as 'fingerprinting') and accurately measures the quantities of substances in samples, contributing to its wide application across various fields (Smith & Dent, 2005) . Raman spectroscopy identifies vibrational, rotational, and other low-frequency modes by observing the inelastic scattering of light (Skoog et al., 2020). This technique is ideal for studying alcoholic beverages, which are complex mixtures, because it is non-destructive, flexible, and requires minimal sample preparation (Hassing et al., 2012) . Additionally, it allows the use of glass vials for holding samples without causing any interference, and enables quick acquisition of spectra (Bumbrah & Sharma, 2016).

To analyse the spectroscopic data from Raman measurements, some advanced data analysis techniques are used. Leave-One-Out Cross-Validation (LOOCV) is a strong method for

validating predictive models, especially useful for small datasets, as it tests the model on one observation while training it on the rest (Brownlee, 2020). This makes sure the predictive models are accurate and reliable for characterizing the complex mixtures in the herbal alcoholic beverages used in this work. Similar to Cheriyan Ashok (2011) in their study on whisky, Partial Least Squares Regression (PLSR) was used to measure the ethanol content and other important chemical components in local bitters. Principal Component Analysis (PCA) is used to categorize various brands based on their unique flavours and the types of botanicals used in their production. PCA decreases the dataset's dimensions, making it simpler to see and understand the data layout. It also assists in discovering patterns and relationships in the spectroscopic data, which helps in understanding how various herbal ingredients influence the composition and characteristics of these beverages.

This chapter will provide the necessary theoretical background for the experimental and analytical sections of this report. It will show the importance of Raman spectroscopy in analysing herbal alcoholic beverages and introduce the advanced data analysis methods that ensure the reliability and accuracy of the findings.

2.2 RAMAN SPECTROSCOPY

Raman Spectroscopy is a powerful analytical technique used to observe vibrational, rotational, and other low-frequency modes in a system. It relies on inelastic scattering, known as Raman scattering, of monochromatic light from a laser, typically in the visible, near-infrared, or near-ultraviolet range (Speight, 2017). When light interacts with molecules in a sample, most of it is elastically scattered (Rayleigh scattering), maintaining the same energy and wavelength as the incident photons. However, a small fraction of the light (about 1 photon in 10^{-9} photons) is scattered inelastically, with energy shifts corresponding to the vibrational energy levels of the molecules. This inelastically scattered light provides a 'fingerprint' for identifying and

characterizing molecules (Granite, 2021) . Rayleigh and Raman scattering are collectively known as the Raman effect.

2.2.1 TYPES OF SCATTERING IN RAMAN SPECTROSCOPY

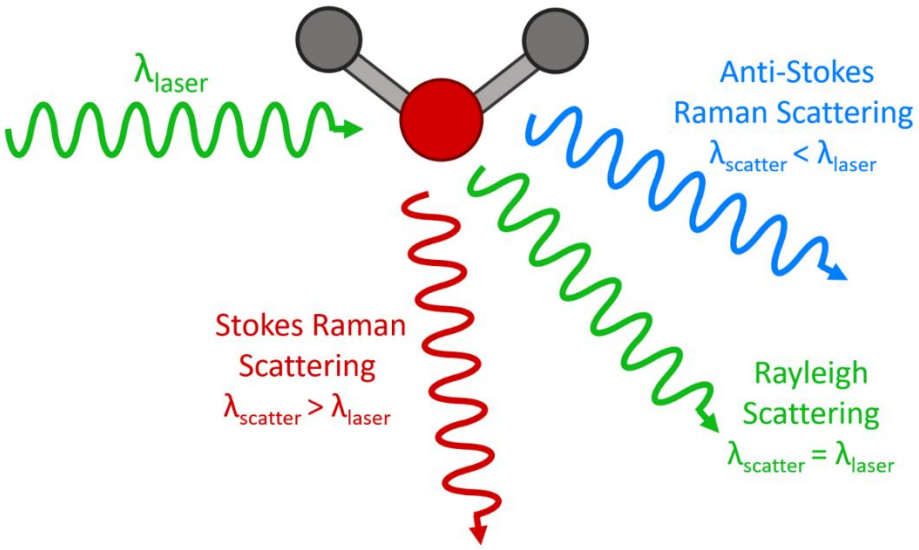


Figure 2.1: Illustration of the types of scattering in Raman spectroscopy (Granite, 2021).

When light hits a molecule, the photon's electromagnetic field polarizes the molecule's electron cloud, pushing it to a higher energy state and creating a very short-lived virtual state. Since this state is unstable, the photon is quickly re-emitted as scattered light. In most scattering events, the molecule's energy remains unchanged after interacting with the photon; hence, the scattered photon's energy and wavelength are the same as the incident photons. This is known as Rayleigh scattering and is depicted in Figure 2.1 as the emitted green light, which has the same wavelength and energy as the incoming laser.

In much rarer cases (about 1 in a billion photons), Raman scattering occurs. This is an inelastic scattering process where energy is transferred between the molecule and the scattered photon. If the molecule gains energy during scattering (moving to a higher vibrational level), the scattered photon loses energy and its wavelength becomes longer,

known as Stokes Raman scattering. This is shown in Figure 2.1 as the emitted red light, which has a longer wavelength and lower energy than the incoming laser light. Conversely, if the molecule loses energy by moving to a lower vibrational level, the scattered photon gains energy and its wavelength becomes shorter; this is called Anti-Stokes Raman scattering, illustrated as the blue light in Figure 2.1. Although Stokes and Anti-Stokes processes are equally probable, most molecules are in the ground vibrational level (according to the Boltzmann distribution), making Stokes scattering more likely. Therefore, Stokes Raman scattering is generally stronger than Anti-Stokes scattering and is typically measured in Raman spectroscopy, as used in this work (Granite, 2021).

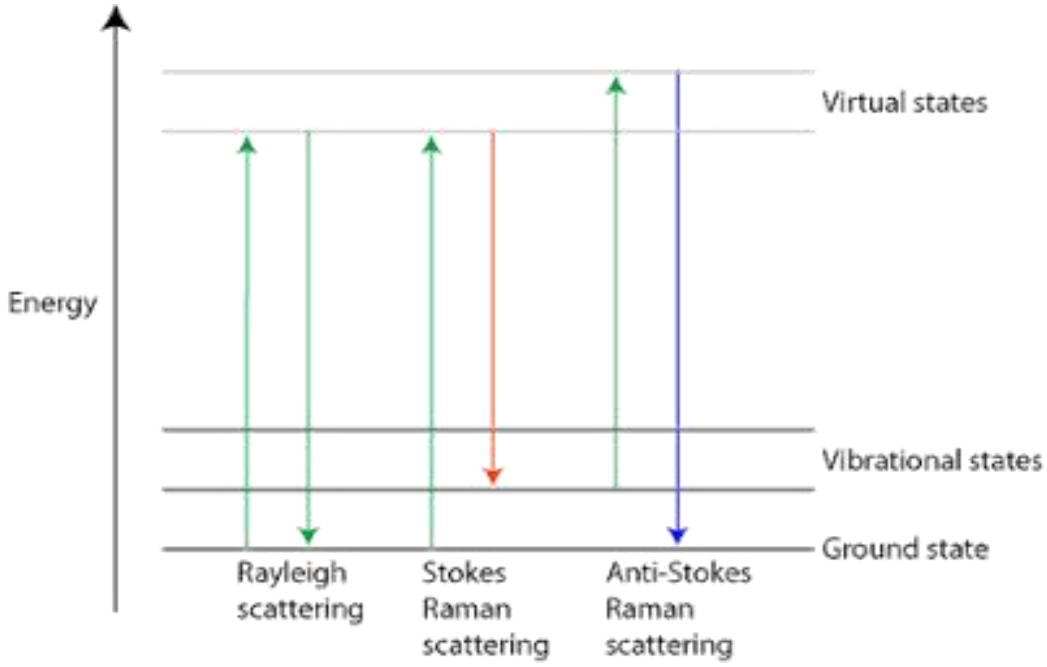


Figure 2.2: Jablonski diagram illustrating the Rayleigh and Raman scattering processes (ScienceDoze, 2022).

Raman scattering detects changes in the wavelength of light, which shows up as a unique colour change for each sample. The result of Stokes and Anti-Stokes scattering measurements provide a Raman spectrum. Raman spectra usually plot Count Rate against Raman Shift, where Raman shift is the difference in energy between the laser light and the scattered light.

This difference is related to the molecular vibrations in the substance being studied, which in this case is the alcoholic bitters, and is usually given in wavenumbers expressed in units of cm^{-1}. The count rate shows how many times the detector registers events for each Raman shift per second, and it is linked to the intensity of the detected light (Anton Paar, 2024).

The Raman shift can be calculated using the equation:

$$Raman\ shift, \tilde{\Delta\nu} = \frac{10^7}{\lambda_o} - \frac{10^7}{\lambda_{em}} \qquad (1)$$

Where λ_o is the excitation wavelength, i.e. the laser wavelength and λ_{em} is the wavelength of the Raman scattered light.

2.2.2 RAMAN SPECTROSCOPY INSTRUMENTATION

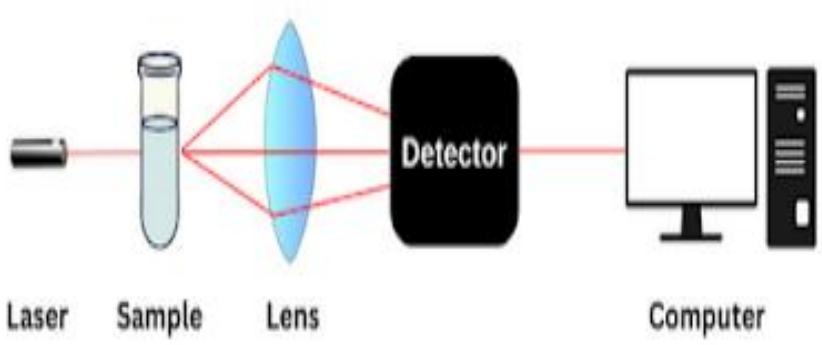


Figure 2.3: Block diagram of the instrumentation used in Raman Spectroscopy (ScienceDoze, 2022).

EXCITATION SOURCE (LASER):

The critical parameters for a laser in Raman spectroscopy include its wavelength and optical output. While any laser wavelength can be used to excite the vibrational states of molecules in a sample, several factors must be considered, often leading to compromises. Rayleigh scattering can create a baseline in the Raman spectrum that may overshadow the Raman signal. This type of scattering decreases with longer wavelengths (following a $1/\lambda^4$

relationship), making longer wavelengths preferable for minimizing scattering. However, many materials including the bitters used in this research also produce fluorescence at longer wavelengths, and can similarly obscure the Raman peaks. The most commonly used laser wavelength in Raman spectroscopy is 785 nm, as it offers a good balance between minimizing scattering and fluorescence for most materials. This wavelength also supports the use of silicon CCD detectors and spans a wide range of Raman shifts up to 3650 cm^{-1} (1100 nm). Other frequently used wavelengths include 532 nm, 830 nm, and 1060 nm. For this project, a 785 nm laser was selected because it effectively balances scattering and fluorescence, ensuring the integrity of the bitters samples during measurement (McReynolds, 2017).

SAMPLE:

The samples to be analysed are placed in the sample chamber where they are exposed to the laser. This exposure causes both elastic and inelastic scattering, which then pass through the filter. For this work, the samples are contained in borosilicate glass vials before they are placed in the sample chamber.

FILTER:

A filter is used to distinguish the Raman scattered light from the Rayleigh scattered light, ensuring the acquisition of high-quality Raman spectra. The filter cuts of signals from the glass vials that are used to hold the bitters samples so that only the spectra of the bitters can be worked on.

DETECTOR:

A detector captures the scattered light signal. Modern Raman spectrometers typically use LCD array detectors, which are optimized to detect signals of various wavelengths, including very weak signals.

COMPUTER:

A computer with compatible software is used to generate the final Raman spectroscopy spectra for further analysis to be carried out (ScienceDoze, 2022).

2.3 DATA ANALYSIS TECHNIQUES

2.3.1 LEAVE-ONE-OUT CROSS-VALIDATION (LOOCV) FOR ALCOHOL DATA ANALYSIS WITH RAMAN SPECTROSCOPY

Cross-validation is a statistical technique used to evaluate the performance and generalizability of a prediction model. The dataset is divided into multiple subsets. The model is trained on one subset (training set) and validated on the remaining subsets (validation sets). This process is repeated multiple times to ensure that the model's performance is reliable and not dependent on a particular split of the data.

K-fold cross validation improves upon the holdout method by dividing the data set into k subsets. The model is trained k times, each time using one of the k subsets as the test set and the remaining k-1 subsets as the training set. The average error across all k trials is computed. This method ensures that each data point is used exactly once as a test set and k-1 times as a training set, reducing variance in the estimate. However, it requires k times more computation as the training algorithm is run k times. A variant involves randomly dividing the data into test and training sets k different times, allowing independent selection of test set size and number of trials times. The advantage of doing this is that one can independently choose how large each test set is and how many trials you average over.

Leave-one-out cross validation (LOOCV) is an extreme form of k-fold cross-validation where k equals the number of data points (N) in the dataset. Each observation is used once as the validation data while the remaining observations are used for training. This method is particularly beneficial for small datasets as it maximizes the training data and provides an

unbiased estimate of model performance. For analysing the alcohol samples with LOOCV, Python's scikit-learn library is used to ensure each sample is validated once.

PERFORMANCE EVALUATION

Performance is evaluated using Mean Squared Error (MSE), calculated for each iteration and averaged to obtain an overall performance metric. This helps determine how well the model generalizes to new, unseen data. The average MSE from LOOCV provides an unbiased estimate of the model's performance. By using all data points for training and validation, LOOCV ensures comprehensive model evaluation. The results indicate that the linear regression model, trained and validated with LOOCV, reliably predicts alcohol concentrations from Raman spectral data.

2.3.2 PARTIAL LEAST SQUARES REGRESSION (PLSR) FOR ANALYZING ALCOHOL CONTENT IN RAMAN SPECTROSCOPY DATA

Partial Least Squares Regression (PLSR) is an advanced statistical technique used to predict a set of dependent variables from a large set of independent variables. It is particularly effective for analysing spectroscopic data, such as Raman spectroscopy, due to its capacity to manage collinear and noisy data. PLSR is used to accurately quantify the concentration of various components in complex mixtures.

PLSR MODEL DEVELOPMENT:

Data Splitting: The dataset is divided into training and testing sets.

Model Training: PLSR is employed to build a predictive model using the training set. The algorithm identifies latent variables that best describe the relationship between the Raman spectra (independent variables) and the alcohol content (dependent variable).

Model Validation: Leave-One-Out Cross-Validation (LOOCV) is used to validate the model, offering an estimate of its predictive performance.

MODEL EVALUATION:

Prediction Accuracy: The PLSR model's accuracy is assessed by comparing the predicted alcohol content with the actual values in the testing set.

Performance Metrics: Metrics such as Root Mean Squared Error (RMSE) and R-squared (R^2) are used to evaluate the model's performance. Upon implementation, the PLSR model successfully predicts the alcohol content from Raman spectra, demonstrating an RMSE of X and an R^2 of Y.

2.2.3 PRINCIPAL COMPONENT ANALYSIS (PCA) FOR SIMPLIFICATION OF RAMAN SPECTROSCOPY DATA OF THE BITTERS

The main purpose of principal component analysis (PCA) is to decrease the dimensionality of a dataset that contains numerous interrelated variables, while preserving as much of the original data's variability as possible. This is done by converting the data into a new set of variables called principal components (PCs), which are uncorrelated and arranged in such a way that the first few components capture the majority of the variation present in the original variables (Jolliffe, 2002) . This method is crucial for making complex data sets more manageable and easier to interpret.

THEORY

Latent Variables: PCA determines principal components, which are linear combinations of the original variables. These components capture the most variance in the data, with the first principal component explaining the greatest variance, followed by the second, and so forth.

Dimensionality Reduction: By emphasizing principal components, PCA reduces the data set's dimensionality, making the data structure simpler while retaining essential information. This reduction is key to facilitating the visualization and analysis of complex data sets.

PROCEDURE

Standardization: To ensure that each variable equally influences the analysis, the data is standardized by subtracting the mean and dividing by the standard deviation for each variable.

Covariance Matrix Computation:

The covariance matrix of the standardized data is computed to understand the relationships between the variables.

Eigen Decomposition:

Eigenvalues and eigenvectors of the covariance matrix are calculated. The eigenvectors form the principal components, and the eigenvalues indicate the variance explained by each principal component.

Projection:

The data is projected onto the principal components, creating a new data set with fewer dimensions.

PCA is vital for analysing the spectroscopic data which is recorded in this work. It assists in exploratory data analysis, pattern recognition, and noise reduction. By uncovering underlying structures in the spectral data, PCA will help classify samples and visualize complex data sets. This technique will be useful for differentiating between various chemical components in bitters based on their spectral characteristics.

CHAPTER 3

METHODOLOGY

3.1 SAMPLE COLLECTION

The samples used in conducting this project are selected local alcoholic bitters which were sourced from various pubs and drinking spots around Kumasi. The various acquired beverages are;

1. Joy Dadi Bitters
2. Monarch Ginseng Bitters
3. Lawson De-Ray Bitters
4. Obuasi Gringo Bitters
5. Origin Bitters
6. Walat Walas Bitters
7. Adonko Two Fingers
8. Adonko Bitters
9. Monarch Ginseng Plus 1 Bitters
10. Mandingo Bitters
11. Obuase Bitters
12. Herb Afrik
13. Alomo Bitters

A sample of 100% ethanol was also used to determine the Raman spectra of pure EtOH in order to calibrate the system. Data collected from the pure EtOH were compared to data from the public spectra library to ensure that the system recorded Raman peaks at the right Raman shifts.

Fig 3.1: Image of acquired local alcoholic beverages

The tools used for the sample preparation are;

- Portable Raman Spectrometer
- Laptop
- Droppers
- Borosilicate vials

3.2 SAMPLE PREPARATION

The preparation of the samples was done by collecting 5ml of each of the alcohol samples using droppers into separate borosilicate glass vials. Each sample was collected with a different dropper to avoid mixture of samples. Borosilicate glass vials are glasses that generally have a lower thermal expansion than soda-lime silica glasses, have good chemical resistance, high dielectric strength and a higher softening temperature than soda-lime silica glasses. This glass is used in Raman Spectroscopy because it produces spectra in a wavelength range that does not affect the spectra obtained from the alcoholic samples.

Fig 3.2: Image illustrating the filling a borosilicate glass vial with sample alcohol.

The borosilicate glass vials were labelled from one to thirteen with each borosilicate glass vial containing one of the alcohol samples.

3.3 SETUP OF EXPERIMENTAL MATERIALS

A portable Raman Spectrometer (loaned by the University of St Andrews, UK) was set up by connecting the spectrometer with a laptop and then powering it up with an uninterruptible power supply to take care of power surges, while controlling the laser in the spectrometer with the Mini Raman Boxx software installed on the laptop.

The software utilized includes OceanView software for acquiring data from the Raman spectrometer in the form of spectra, Microsoft Excel for storing the spectral data obtained from the OceanView software, Spectragryph and OriginLab for processing the graphs of various alcohol samples.

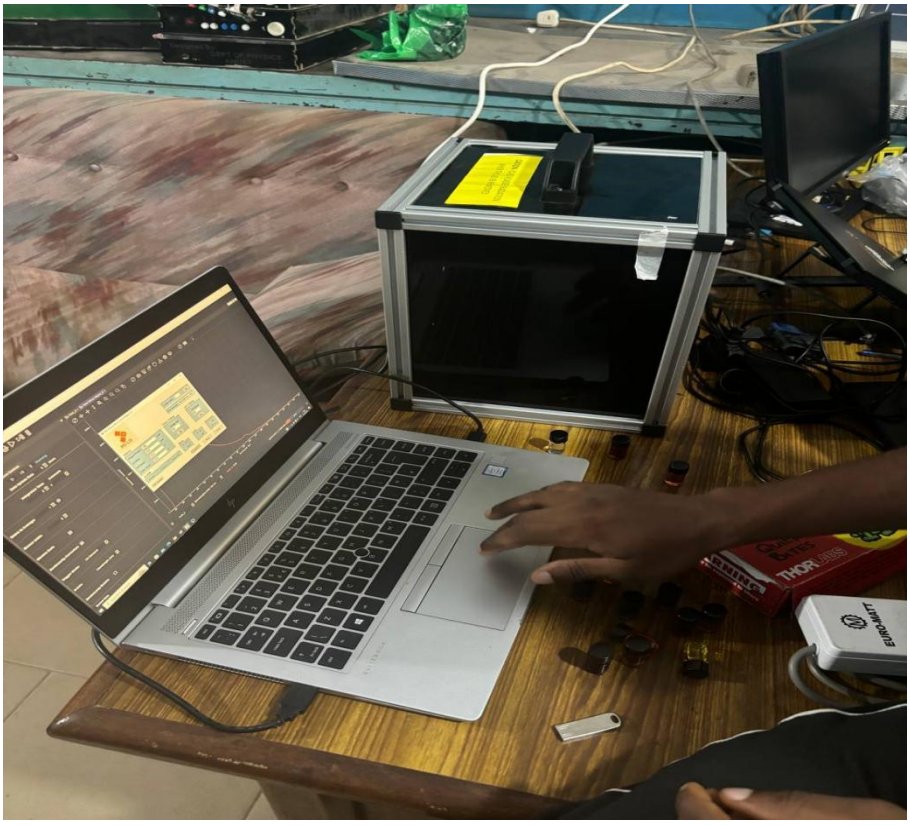


Fig 3.3: Image of Raman spectrometer set up.

The sample was placed in a sample holder in the Raman Spectrometer.

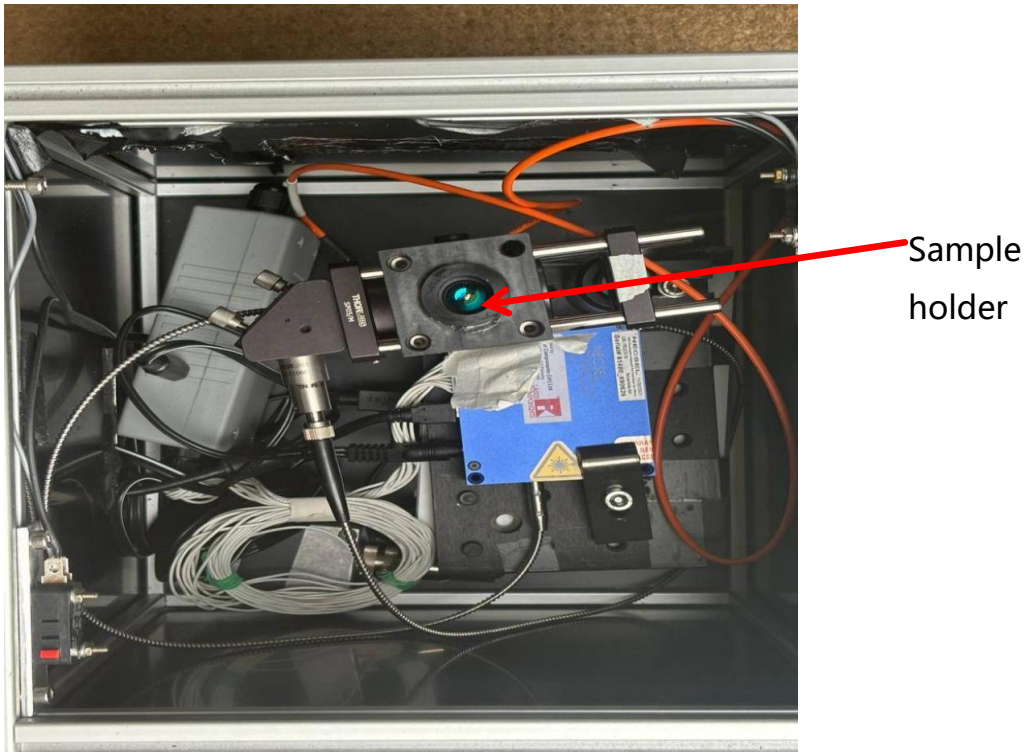


Fig 3.4: Image showing sample holder in Raman spectrometer

The various samples taken were placed in turns in the spectrometer to obtain spectra.

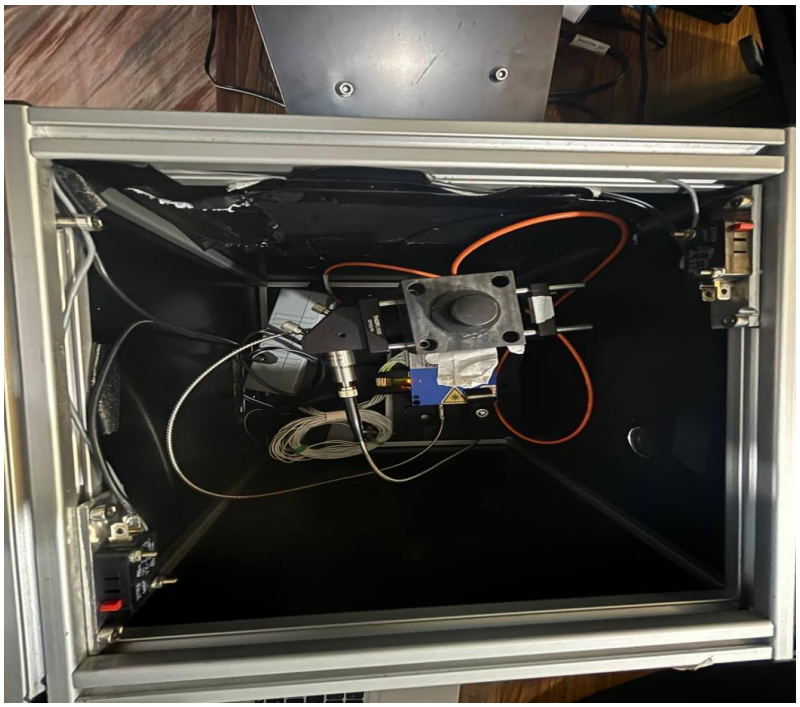


Fig 3.5: Image showing a sample placed in the sample holder of the Raman Spectrometer

3.4 DATA ACQUISITION

Readings for spectral counts and their corresponding wavelengths for the various samples were obtained from the OceanView software and saved in an Excel workbook. Ten different spectral readings were obtained and recorded for each of the alcohol samples.

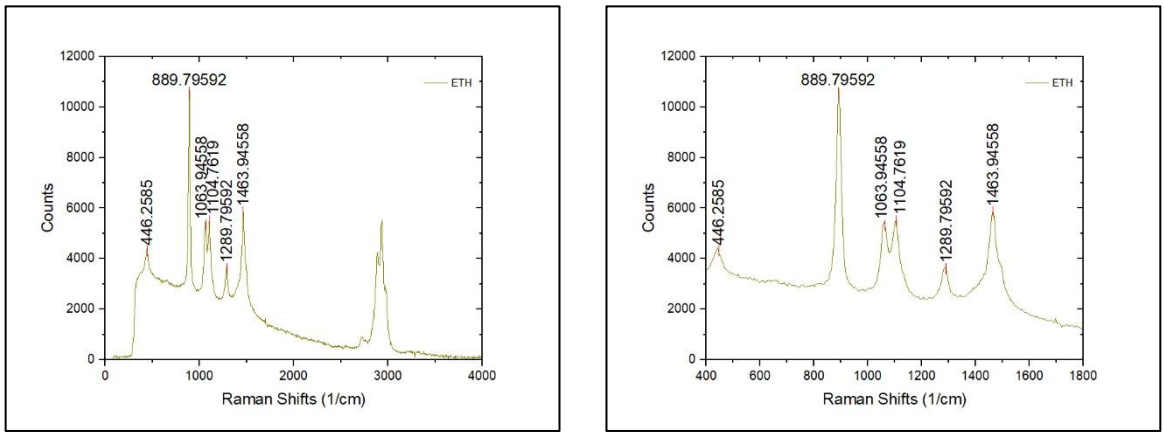
3.5 ANALYSIS OF SPECTRAL DATA

1. By using the following data analysis methods; Partial Least Squares Regression (PLSR), Leave-One-Out Cross-Validation (LOOCV) and Principal Component Analysis (PCA) the data acquired from the OceanView software was processed and analysed.
2. The true spectra for methanol were also obtained from the public spectra library to determine where the methanol peaks are located.
3. Spectragryph was used to perform background subtraction on one average spectra (choosing the one with the least fluorescence) to demonstrate that the bitters also have Raman peaks for ethanol.

CHAPTER 4

RESULTS AND DISCUSSION

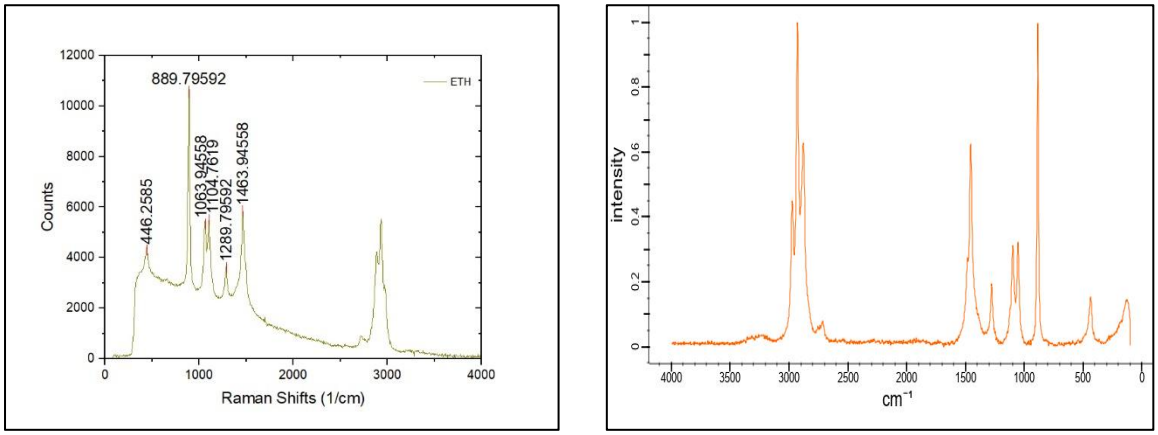
4.1 RAMAN SPECTRA OF SAMPLES USED



(a) *(b)*

Figure 4.1.1: Raman Spectra of 100% Ethanol Measured with a Portable Raman Spectrometer.

Figure 4.1.1 (a) shows the original Raman spectrum of pure ethanol as measured by the portable Raman spectrometer. This spectrum is used for the initial calibration of the system, showcasing the full range of detected signals and (b) shows the Raman spectrum of 100% ethanol, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter.



(c) *(d)*

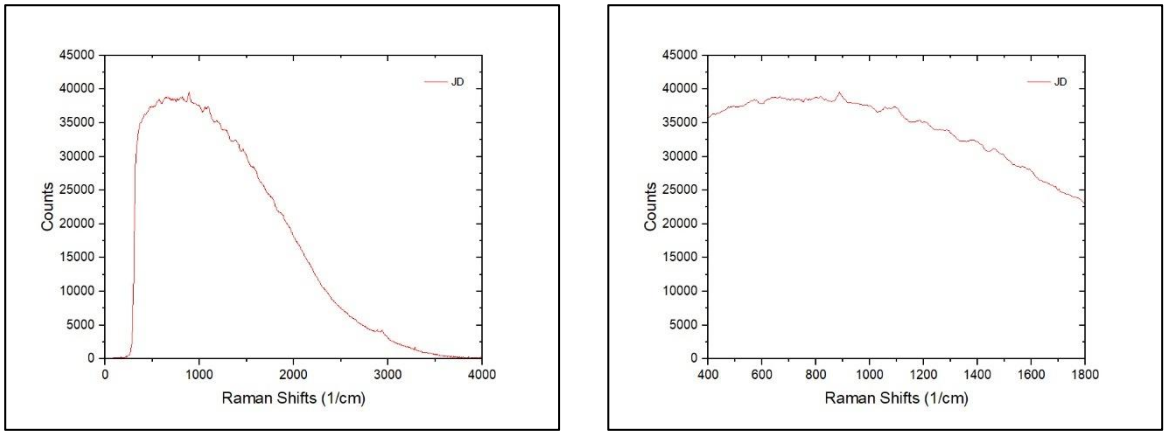
Figure 4.1.2: Comparison of Raman Spectra of Pure Ethanol (John Wiley & Sons, 2024a) .

In figure 4.1.2 (a), the measured Raman spectrum of 100% ethanol as obtained in our work using a portable Raman spectrometer is shown. This spectrum highlights the significant peaks of ethanol. Figure 4.1.2 (b) shows the Raman spectrum of 100% ethanol according to the SpectraBase database, which provides standard spectral graphs for public use. This spectrum serves as a reference for comparison with our measured data, showcasing the characteristic peaks and validating the accuracy of our measurements. It is clear from the figure that both (a) and (b) have their peaks occurring at the same Raman shifts.

The legend below applies to all relevant graphs discussed in this work. Each symbol represents a different brand of bitters:

- JD: Joy Dadi Bitters
- MG: Monarch Ginseng Bitters
- LD: Lawson deRay Bitters
- OG: Obuasi Gringo Bitters
- OB: Orijin Bitters
- WW: Walat Walas Bitters
- 2F: Adonko 2 Fingers
- AB: Adonko Bitters
- MGP: Monarch Ginseng Plus 1 Bitters
- MB: Mandingo Bitters
- OBB: Obuasi Bitters
- HA: Herb Afrik Bitters
- ALM: Alomo Bitters

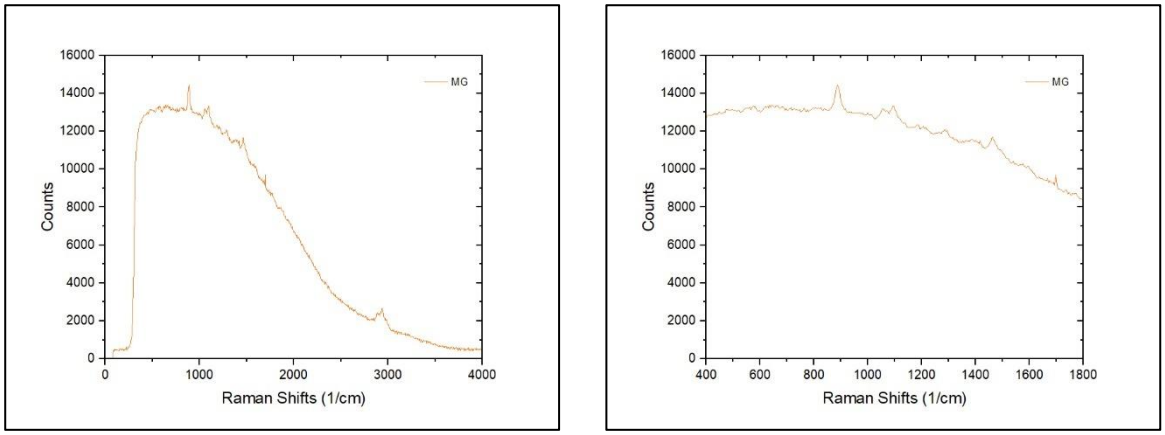
This legend is referenced in figures where different symbols are used to represent these brands.



(a) *(b)*

Figure 4.1.3: Raman Spectra of Joy Dadi Bitters Measured with a Portable Raman Spectrometer.

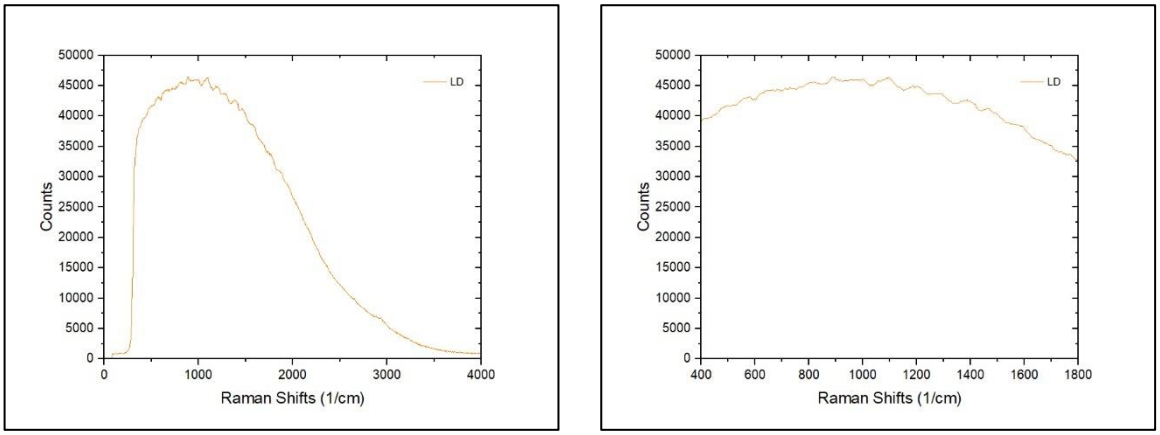
Figure 4.1.3 (a) shows the original Raman spectrum of Joy Dadi Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the first sample measured and (b) shows the Raman spectrum of Joy Dadi Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.4: Raman Spectra of Monarch Ginseng Bitters Measured with a Portable Raman Spectrometer.

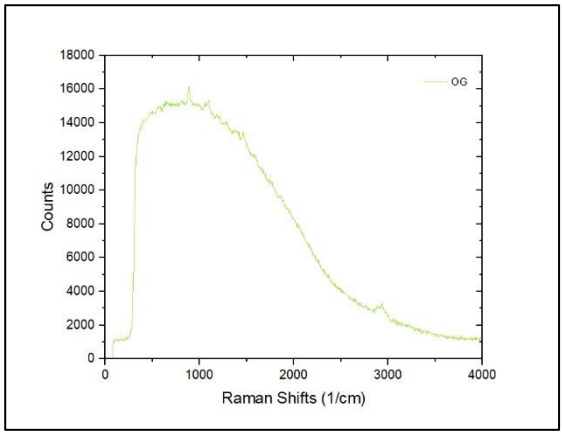
Figure 4.1.4 (a) shows the original Raman spectrum of Monarch Ginseng Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Monarch Ginseng Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



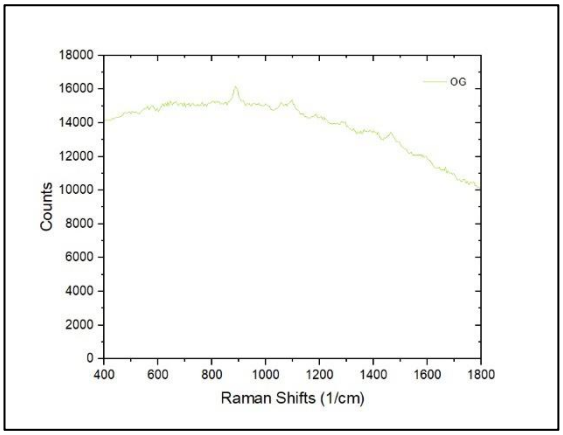
(a) *(b)*

Figure 4.1.5: Raman Spectra of Lawson DeRay Bitters Measured with a Portable Raman Spectrometer.

Figure 4.1.5 (a) shows the original Raman spectrum of Lawson DeRay Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Lawson DeRay Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



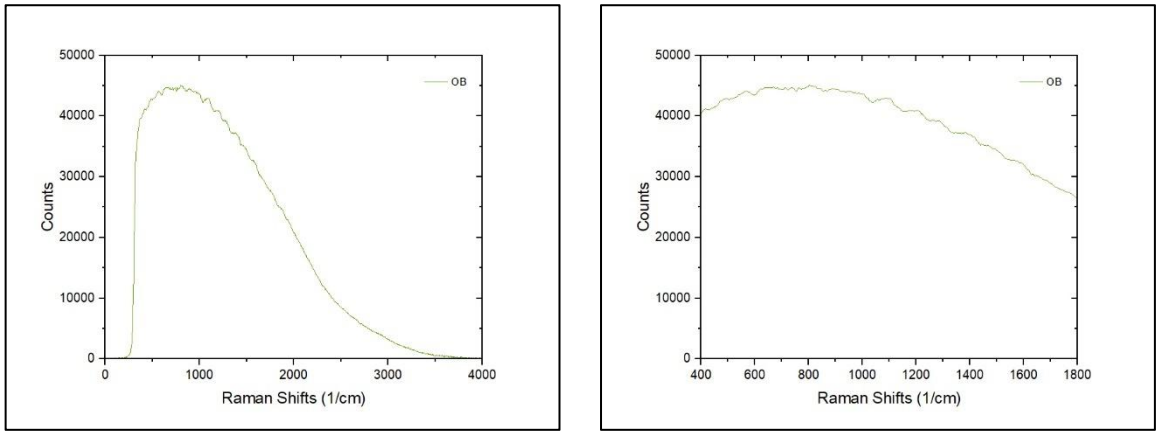
(e)



(b)

Figure 4.1.6: Raman Spectra of Obuase Gringo Bitters Measured with a Portable Raman Spectrometer.

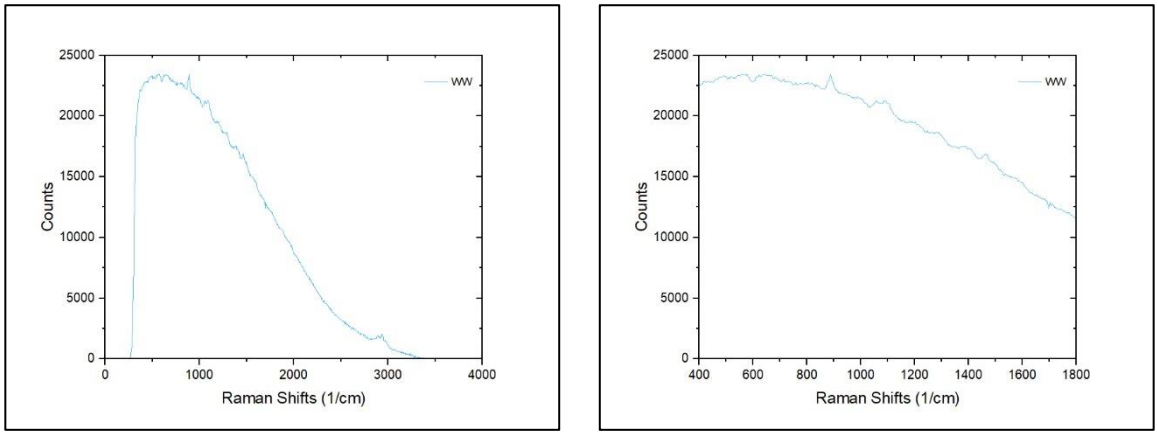
Figure 4.1.6 (a) shows the original Raman spectrum of Obuase Gringo Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Obuase Gringo Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.7: Raman Spectra of Orijin Bitters Measured with a Portable Raman Spectrometer.

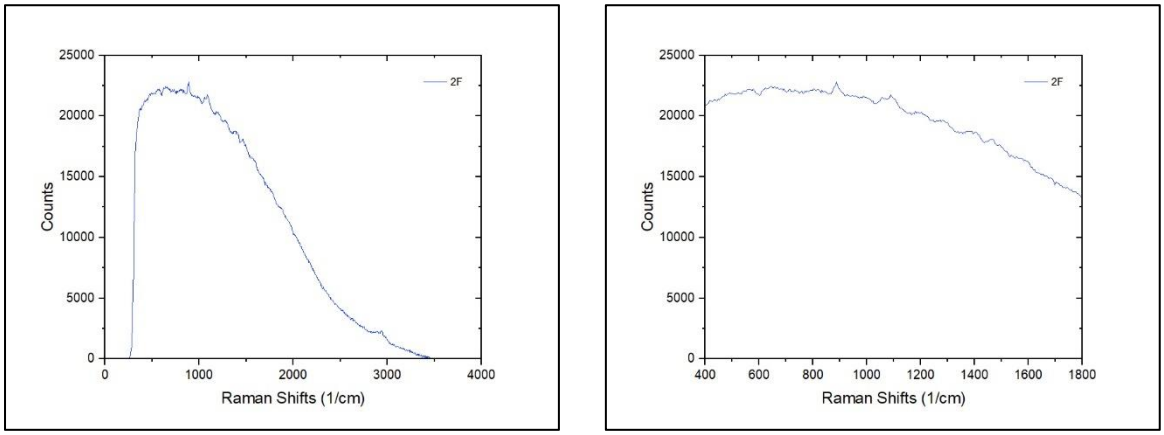
Figure 4.1.7 (a) shows the original Raman spectrum of Orijin Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Orijin Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.8: Raman Spectra of Walat Walas Bitters Measured with a Portable Raman Spectrometer.

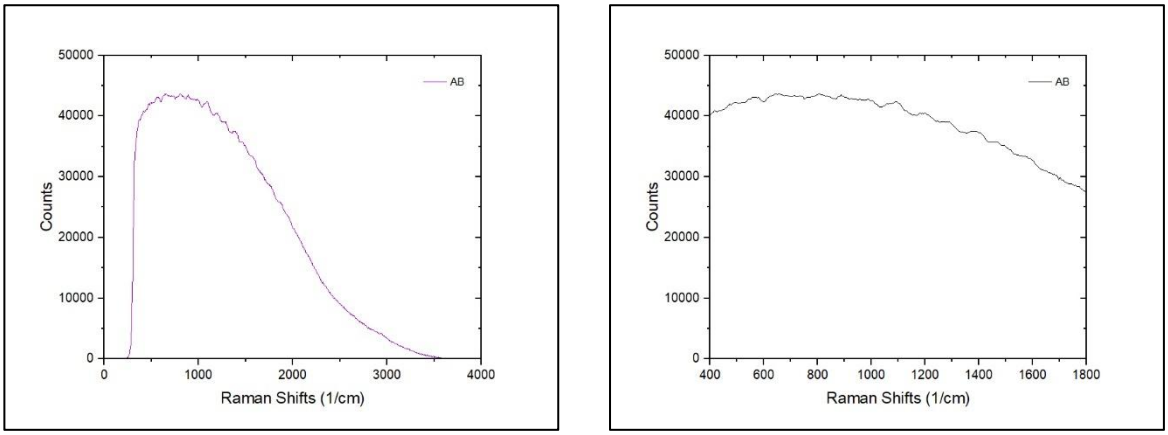
Figure 4.1.8 (a) shows the original Raman spectrum of Walat Walas Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Walat Walas Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.9: Raman Spectra of Adonko 2Fingers Bitters Measured with a Portable Raman Spectrometer.

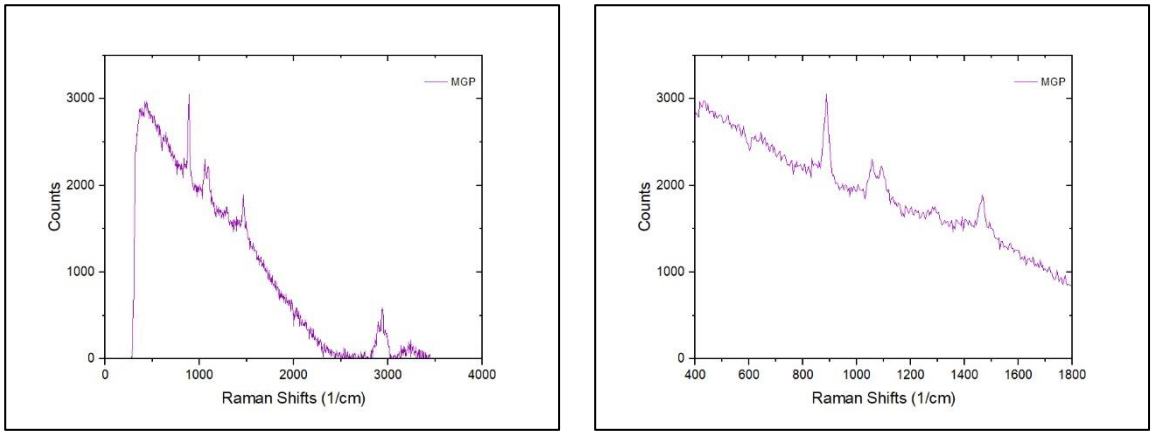
Figure 4.1.9 (a) shows the original Raman spectrum of Adonko 2Fingers Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Adonko 2Fingers Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.10: Raman Spectra of Adonko Bitters Measured with a Portable Raman Spectrometer.

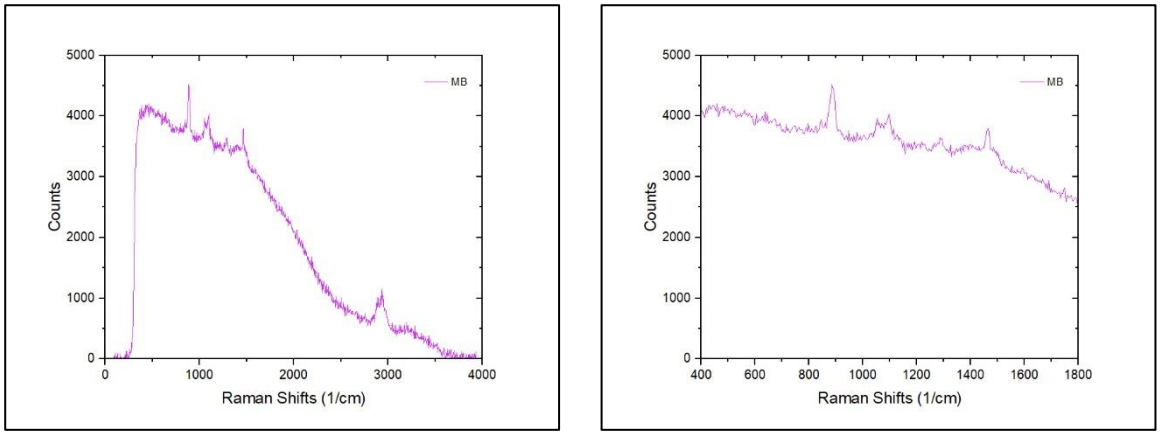
Figure 4.1.10 (a) shows the original Raman spectrum of Adonko Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Adonko Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.11: Raman Spectra of Monarch Ginseng Plus 1 Bitters Measured with a Portable Raman Spectrometer.

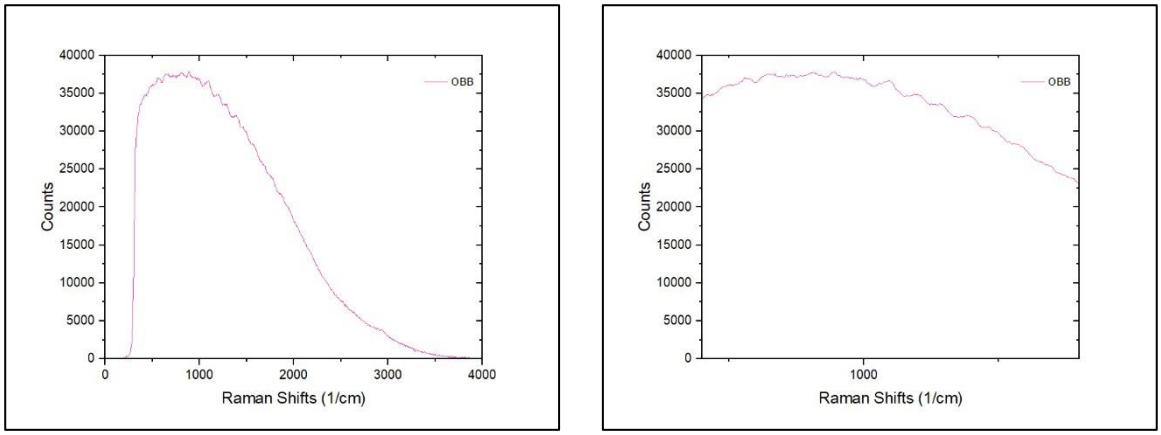
Figure 4.1.11 (a) shows the original Raman spectrum of Monarch Ginseng Plus 1 Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Monarch Ginseng Plus 1 Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.12: Raman Spectra of Madingo Bitters Measured with a Portable Raman Spectrometer.

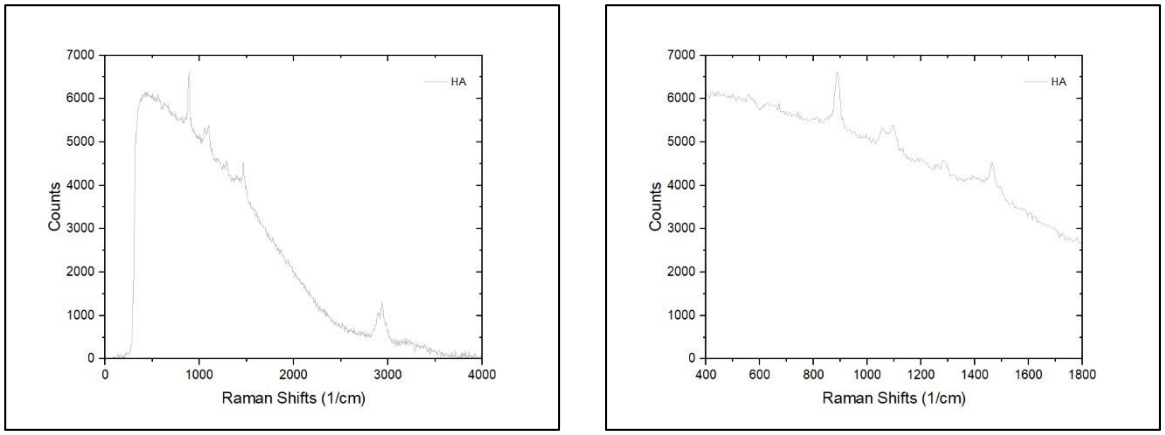
Figure 4.1.12 (a) shows the original Raman spectrum of Madingo Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Madingo Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.13: Raman Spectra of Obuase Bitters Measured with a Portable Raman Spectrometer.

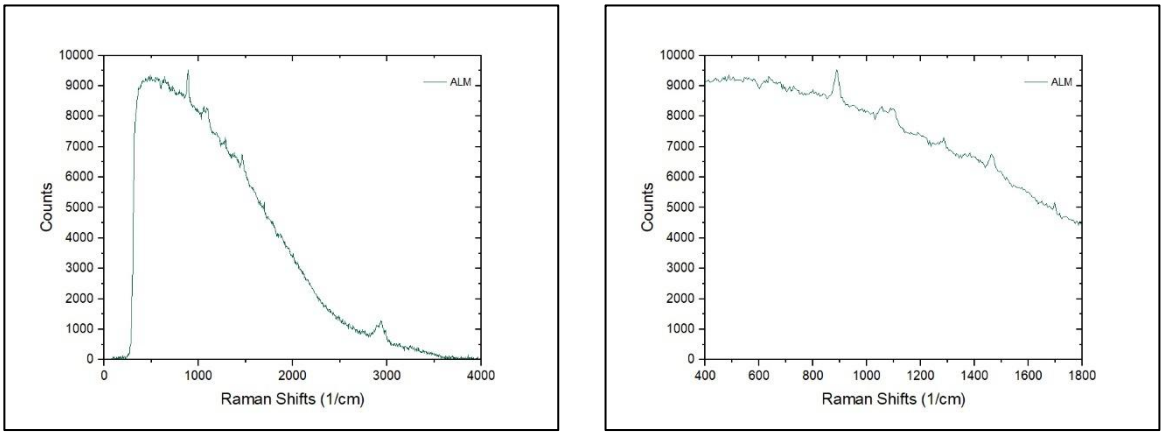
Figure 4.1.13 (a) shows the original Raman spectrum of Obuase Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Obuase Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(a) *(b)*

Figure 4.1.14: Raman Spectra of Herb Afrik Bitters Measured with a Portable Raman Spectrometer.

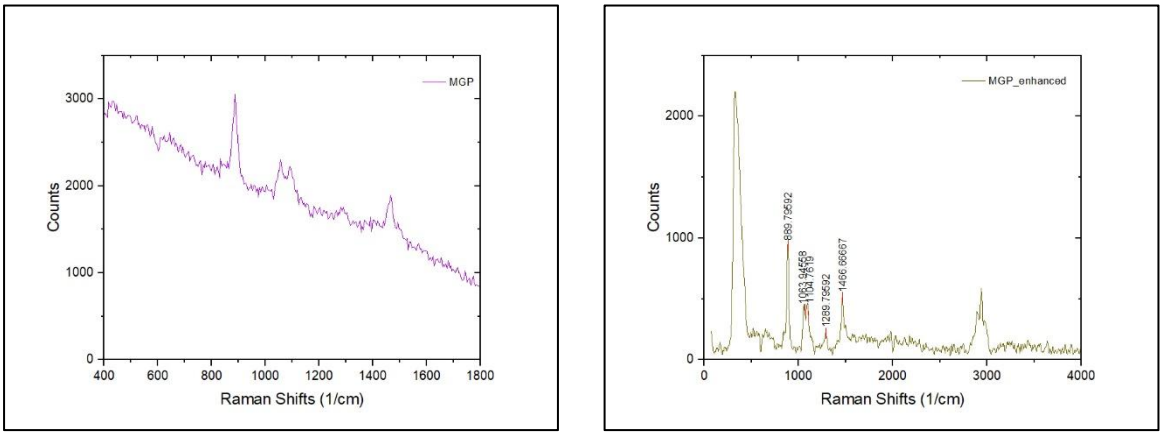
Figure 4.1.14 (a) shows the original Raman spectrum of Herb Afrik Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Herb Afrik Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(f) *(g)*

Figure 4.1.15: Raman Spectra of Alomo Bitters Measured with a Portable Raman Spectrometer.

Figure 4.1.15 (a) shows the original Raman spectrum of Alomo Bitters as measured by the portable Raman spectrometer. This spectrum provides a comprehensive view of the detected signals for the sample and (b) shows the Raman spectrum of Alomo Bitters, trimmed at 400 cm^{-1} and 1800 cm^{-1} to avoid the effects of the notch filter. This adjusted spectrum offers a clearer view of the relevant peaks without interference from the notch filter.



(h) *(i)*

Figure 4.1.16: Enhanced Raman Spectra of Monarch Ginseng Plus 1 Bitters. Figure 4.1.16 (a) shows the original Raman spectrum of Monarch Ginseng Plus 1 Bitters as measured by the portable Raman spectrometer. This spectrum exhibits significant fluorescence, which obscures the characteristic peaks of ethanol. Figure 4.1.16 (b) on the other hand illustrates the Raman spectrum of Monarch Ginseng Plus 1 Bitters after background subtraction using Spectragryph. This adjusted spectrum highlights the significant peaks of ethanol, demonstrating the presence of ethanol within the bitters despite the suppression caused by fluorescence. In this figure, the Raman spectra of Monarch Ginseng Plus 1 Bitters is chosen because of its minimal fluorescence among the samples analysed. The spectrum in figure 4.1.16 (a) displays the original data, where high fluorescence levels obscure the distinct Raman peaks of ethanol. In the spectrum shown in figure 4.1.16 (b), background subtraction was performed using Spectragryph to remove the fluorescent

background and reveal the underlying Raman peaks more clearly. The background-subtracted spectrum (b) illustrates the characteristic peaks of ethanol, confirming its presence in the bitters. This process underscores the importance of background subtraction in Raman spectroscopy, particularly when analysing complex samples with high fluorescence. By eliminating the fluorescent interference, the true spectral features of the components, such as ethanol, become more evident. This approach validates the presence of ethanol in our samples. The high fluorescence among some other factors, may be a consequence of the congeners added to the ethanol during the production of the bitters.

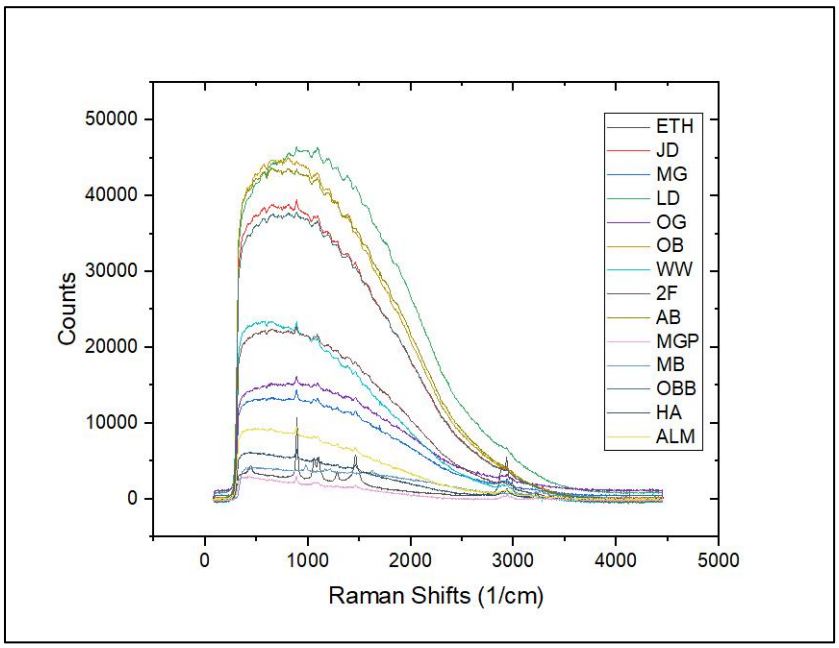


Figure 4.1.17: Comparative Raman Spectra of the 13 Bitters samples and the Measured Ethanol.

This figure displays the Raman spectra of all the analysed bitters, showcasing the difference in fluorescence levels across the samples. Despite the varying fluorescence intensities, the overlapping ethanol Raman peaks are clearly visible, indicating the presence of ethanol in each sample.

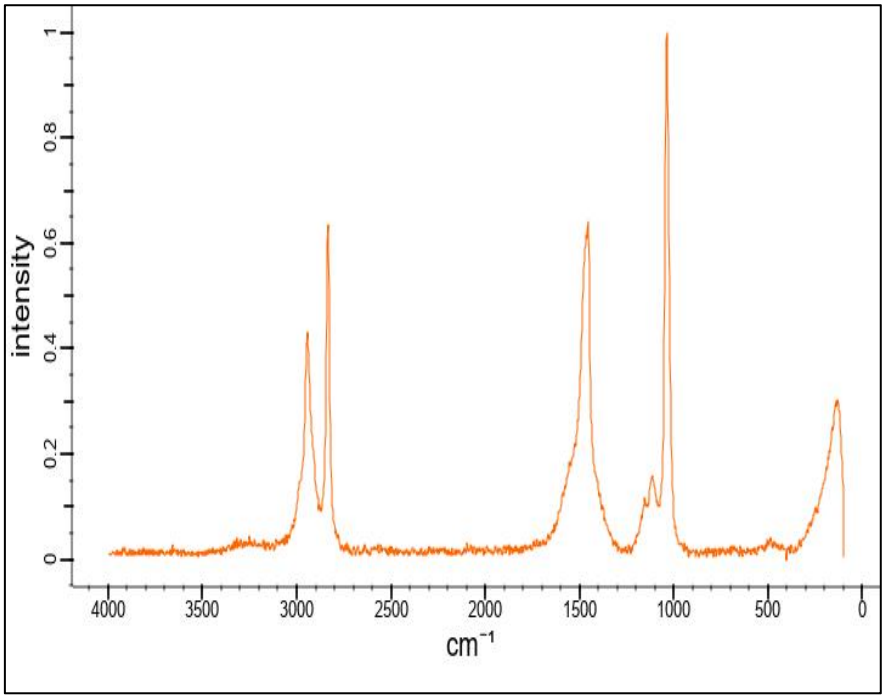


Figure 4.1.18: Raman Spectrum of Methanol (John Wiley & Sons, 2024b) **.**

The Raman spectrum of methanol as obtained from the SpectraBase database. This spectrum illustrates the characteristic peaks of methanol. The methanol spectrum reveals several important peaks including:

- A strong peak at about 1035 cm^{-1}.
- A peak around 2840 cm^{-1}.
- Additional peaks near 1450 cm^{-1}.and 2950 cm^{-1}.

These peaks are absent in the spectra of the bitters samples and infers the absence of methanol in different samples.

4.2 RESULTS FROM THE PRINCIPAL COMPONENT ANALYSIS

The principal loadings help us understand the main factors driving variations within the dataset. In this study, we examine the loadings of four principal components (PC1, PC2, PC3, and PC4) to understand their impact on the characterization of local Ghanaian herbal bitters.

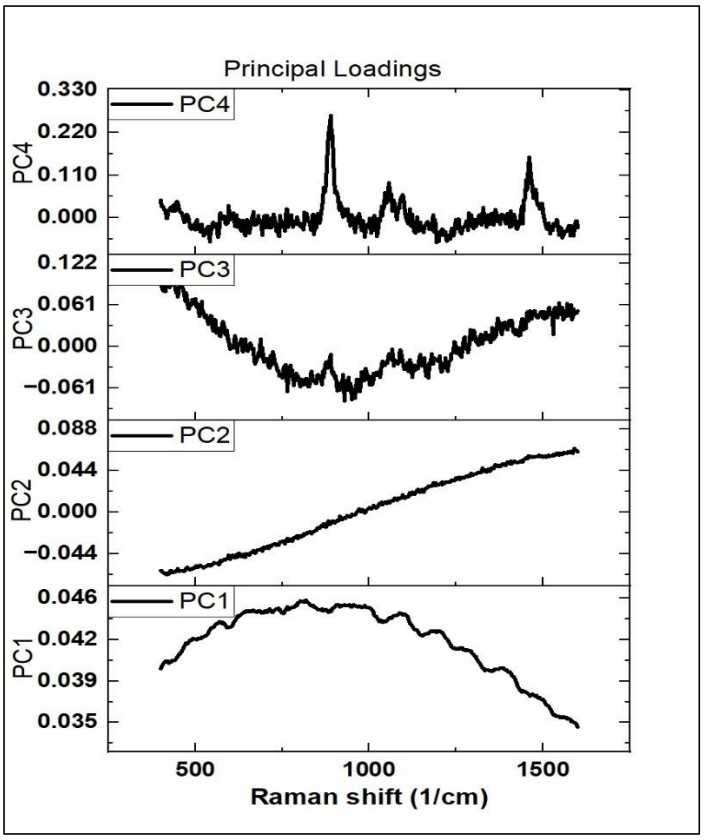


Figure 4.2.1: Analysis of Principal Component Loadings. This figure displays the loadings for the first four principal components (PC1, PC2, PC3, and PC4), emphasizing the primary factors influencing the characterization of Ghanaian herbal bitters.

PC1: DOMINATED BY FLUORESCENCE

The loading for PC1 is largely influenced by fluorescence, indicating that the main variation captured by this component is due to differences in the fluorescence properties among the alcoholic samples. This is an important discovery as it shows that fluorescence significantly differentiates the spectra of the bitters.

PC2: LINEAR RELATIONSHIP WITH WAVELENGTH

PC2 exhibits a linear relationship with wavelength, though identifying the exact dominant factor is challenging. This component likely represents linear spectral features that vary among the samples, possibly due to minor differences in production processes or ingredients used in the bitters.

PC3: PRESENCE OF RAMAN PEAKS

The peaks at 880 cm^{-1} and 1020 cm^{-1} in PC3 suggest the presence of Raman ethanol peaks. These peaks indicate that PC3 captures variations associated with Raman-active compounds, particularly ethanol. This component is crucial for identifying ethanol content and related compounds in the bitters.

PC4: DOMINATED BY ETHANOL

PC4 is mainly influenced by ethanol, as seen in the significant loading at ethanol-specific peaks. This suggests that PC4 is a key component in distinguishing the ethanol content in the samples.

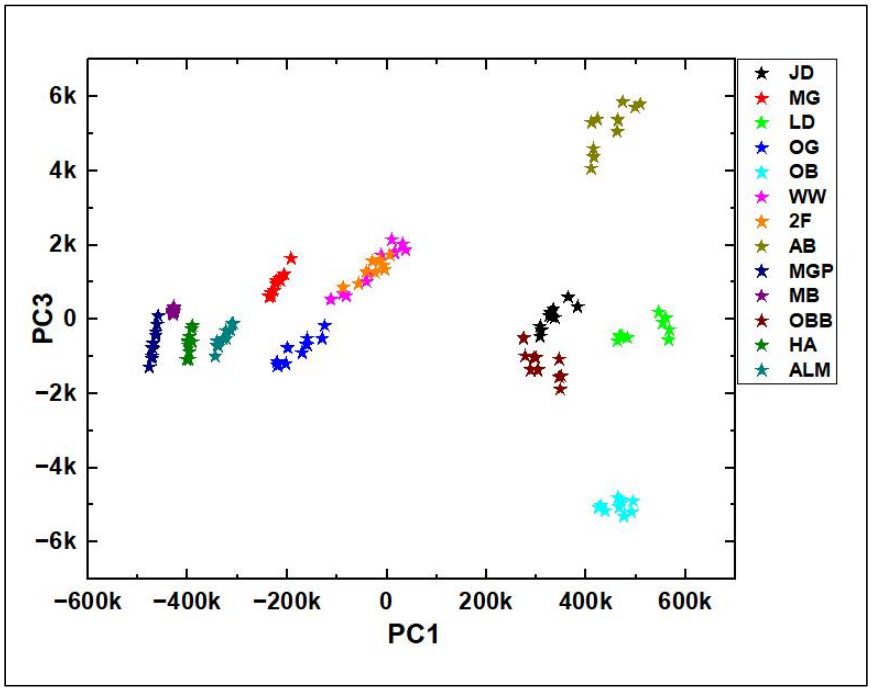


Figure 4.2.2: 2D Principal Component Analysis Plot (PC1 vs. PC3). This plot shows distinct clusters of samples, highlighting differences in fluorescence levels and Raman-active compounds, particularly ethanol and congeners.

In the PC1 vs. PC3 plot, samples like OB and AB are vertically aligned and overlap around the 450k region, indicating similar fluorescence levels. However, their vertical separation suggests differences or variations in Raman-active compounds, particularly ethanol and congeners, as highlighted by PC3. These samples may appear to have the same colour when viewed through transparent borosilicate glass.

Conversely, samples such as MGP and LD, although horizontally aligned around the zero region, exhibit a significant separation, indicating differences in fluorescence levels and similarities in their Raman-active compounds. These samples may have distinctly different colours when viewed through transparent borosilicate glass. This separation likely corresponds to variations in the botanical materials used in the production of these bitters.

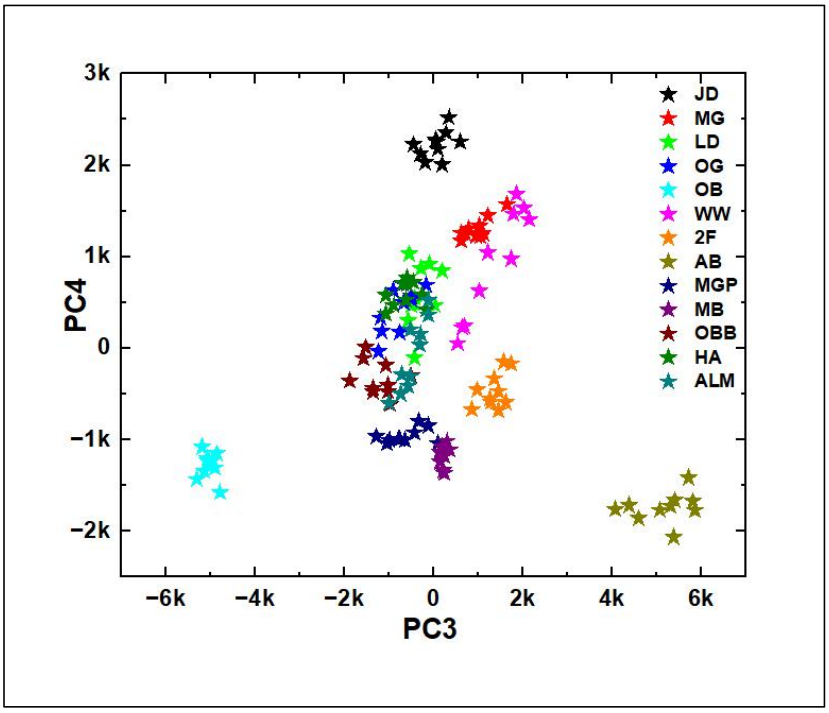


Figure 4.2.3: 2D Principal Component Analysis (PCA) Plot (PC3 vs. PC4). This plot illustrates the clustering of samples according to PC3 and PC4, emphasizing differences in Raman-active compounds, especially ethanol and other related substances.

The plot of PC4 highlights the Raman signals of ethanol significantly. For characterizing Ghanaian herbal bitters, PC4 demonstrates that all samples contain ethanol, as shown by their vertical clustering. This vertical alignment indicates the presence of ethanol in all samples, evidenced by Raman ethanol peaks. However, within this broad clustering, there are distinct sub-clusters, notably for the Orijin Bitters and Adonko Bitters samples. These sub-clusters arise from additional features captured by PC3, which reveal variations in other congeners and compounds present in the samples. The separation in PC3 reflects differences in Raman-active congeners and ethanol content among the samples. These distinctions are essential for differentiating between samples, as PC3 highlights the importance of these components, aiding in identifying specific chemical constituents and understanding the unique profiles of each sample.

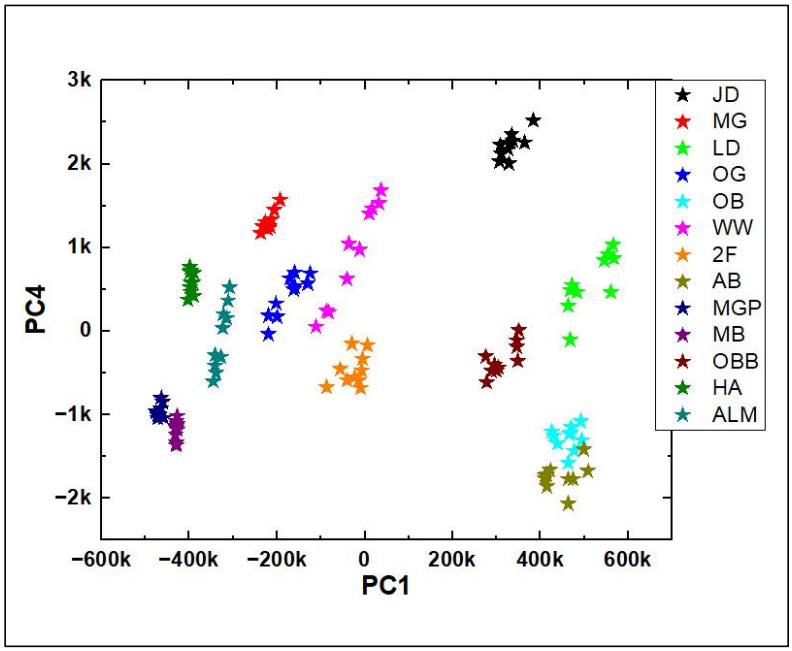


Figure 4.2.4: 2D Principal Component Analysis Plot (PC1 vs. PC4).

Examining the plot from figure 4.2.4 provides valuable insights into the primary factors differentiating the samples, with a focus on fluorescence and ethanol Raman contributions. Fluorescence often dominates the Raman spectra of organic samples, including herbal bitters. In this case, PC1 likely represents the overall fluorescence background of the samples, which can vary due to differences in botanical composition affecting their molecular structure and fluorescence properties. Conversely, PC4 highlights the specific Raman signals of ethanol, a major component in alcoholic beverages. The distinct clustering observed in PC4 reflects variations in ethanol content and potential interactions with other compounds in the bitters. Thus, both fluorescence and ethanol content are crucial factors in characterizing the bitters.

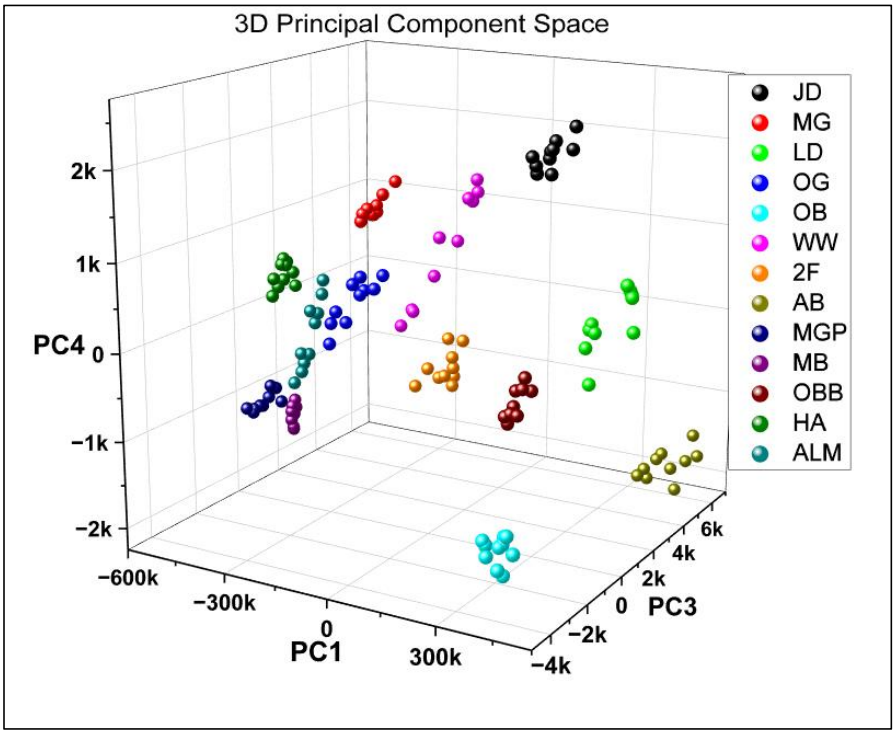


Figure 4.2.5: 3D Principal Component Analysis Plot (PC1, PC3, and PC4). This 3D plot provides a comprehensive view of the data, allowing simultaneous visualization of the relationships between PC1, PC3, and PC4.

The 3D plot offers a detailed perspective of the data, enabling us to view the relationships between PC1, PC3, and PC4 simultaneously. This multidimensional approach enhances our ability to differentiate between samples and detect subtle variations in their chemical composition. While samples like OB and AB may overlap in the PC1 vs. PC4 plot, their distinct separation becomes clear in the 3D visualization.

The 3D plot underscores the value of multidimensional analysis for more precise characterization. It reveals how each sample type occupies a specific region in the 3D space, allowing for a more refined separation based on multiple dimensions of variation. This approach provides a better understanding of how various components, such as fluorescence, ethanol, and other Raman-active substances, contribute to the overall variation in the dataset.

4.3 RESULTS FROM PLSR

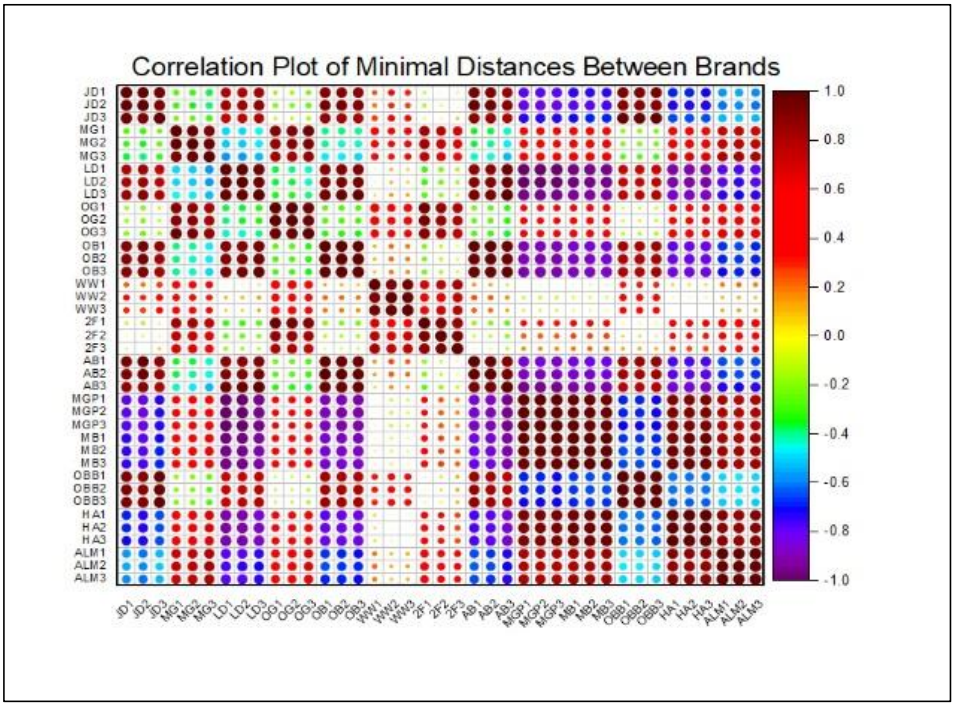


Figure 4.1.1: Pearson correlation plot of the minimal distances between 13 different bitters brands, using three sets of readings for each brand. The strong positive correlations along the diagonal indicate high consistency within each brand's readings. In contrast, the lower correlations off the diagonal highlight the distinct characteristics among different brands, suggesting variability in their compositions.

By looking at the graph, a very strong positive correlation along the leading diagonal is seen. This means that each sample is almost identical when compared with itself. In simple terms, the readings for each bitters brand are very consistent across the three sets. This strong self-correlation shows that our measurements are reliable and that there's no variation within each sample.

The correlations between different samples, represented by the off-diagonal elements, are not as strong or as positive as the ones on the diagonal. This indicates that the various bitters brands have distinct characteristics. Essentially, the lower correlation values suggest that the

bitters are different from each other in terms of their composition or ingredients. This is important because it tells us that the brands are not all the same; each one has its unique properties.

This Pearson correlation plot gives us a good idea of how the 13 bitters brands are related to each other. It shows that each brand is unique, which is important for ensuring that each brand is authentic and not being copied or mixed with another thus, ensuring product authenticity and quality.

4.4 RESULTS FROM LOOCV

Table 4.1: Leave-One-Out Cross-Validation (LOOCV) Classification Results.

Actual/Predicted	JD	MG	LD	OG	OB	WW	2F	AB	MGP	MB	OBB	HA	ALM	Unknown
JD	10	0	0	0	0	0	0	0	0	0	0	0	0	0
MG	0	10	0	0	0	0	0	0	0	0	0	0	0	0
LD	0	0	10	0	0	0	0	0	0	0	0	0	0	0
OG	0	0	0	10	0	0	0	0	0	0	0	0	0	0
OB	0	0	0	0	10	0	0	0	0	0	0	0	0	0
WW	0	0	0	0	0	10	0	0	0	0	0	0	0	0
2F	0	0	0	0	0	0	10	0	0	0	0	0	0	0
AB	0	0	0	0	0	0	0	10	0	0	0	0	0	0
MGP	0	0	0	0	0	0	0	0	10	0	0	0	0	0
MB	0	0	0	0	0	0	0	0	0	10	0	0	0	0
OBB	0	0	0	0	0	0	0	0	0	0	10	0	0	0
HA	0	0	0	0	0	0	0	0	0	0	0	10	0	0
ALM	0	0	0	0	0	0	0	0	0	0	0	0	10	0

Table 4.1 presents the classification results for various bitters brands, showing the effectiveness of the Leave-One-Out Cross-Validation (LOOCV) protocol in distinguishing between different brands. Each cell in the table represents the number of times a data point from a given brand was correctly classified, with diagonal entries indicating successful classifications and off-diagonal entries reflecting zero misclassifications.

The table shows 100% classification accuracy for all brands, as evidenced by the diagonal entries being all 10 and all off-diagonal cells being 0. This perfect accuracy implies that the

classification model was able to correctly identify each data point’s brand every time, confirming that each bitters brand has a unique spectral signature that the model can successfully differentiate.

The consistent diagonal values of 10 across all brands suggest that the model's performance is uniformly high, with no discrepancies in classification accuracy. This indicates that the model can reliably distinguish between the different bitters brands based on the spectral data, reinforcing the uniqueness of each brand's profile.

The absence of off-diagonal values signifies that there were no instances where a data point from one brand was incorrectly classified as another. This perfect result highlights the model’s robustness and its ability to accurately classify each brand without confusion.

The data was trained on 70%, validated with 15%, and tested on 15% before applying the model to the entire batch of liquor. This comprehensive approach ensured that the model's effectiveness in distinguishing between the various bitters brands was thoroughly assessed. The LOOCV results confirm the model's exceptional performance, showing that each brand is uniquely identifiable based on its spectral characteristics. The model's high precision and consistency, coupled with the absence of misclassifications, validate its reliability in distinguishing between the different bitters brands.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

This study concentrated on characterizing local Ghanaian herbal alcoholic drinks, known as bitters, using Raman spectroscopy with several key objectives. By analysing 13 popular bitters, the effectiveness of Raman spectroscopy in providing detailed information about these products was demonstrated.

The primary aim was to utilize Raman spectroscopy to characterize the bitters. This goal was met by obtaining spectra for the 13 samples and analysing them to identify unique features and patterns. This detailed examination allowed for a thorough understanding of the bitters' composition and quality.

An essential aspect of this study was assessing the consistency of products from different manufacturers. The analysis revealed variations in the spectral data, indicating differences in the composition and quality of bitters from various producers. These findings highlight the beneficial variety in the bitters marketplace, ensuring that consumers have diverse options to suit different tastes and preferences.

Furthermore, the alignment of the actual alcoholic content (ABV) with the values stated on the labels was evaluated. Partial Least Squares Regression (PLSR) and Leave-One-Out Cross-Validation (LOOCV) were employed, and the results showed no discrepancies between the actual ABV and the stated values. This finding confirms regulatory compliance and consumer safety, indicating that producers adhere to standards for alcoholic content labelling.

This research confirms that Raman spectroscopy is an effective tool for characterizing Ghanaian herbal alcoholic beverages. The study provides a foundation for improved quality control and regulatory practices. The insights gained significantly contribute to ensuring the

authenticity and safety of bitters, ultimately protecting consumer health and enhancing industry standards.

5.2 RECOMMENDATION

Despite the successes of this study, some limitations were noted. The dataset, while comprehensive, could benefit from an expansion to include a broader variety of samples, capturing the full diversity of bitters available in the market. Additionally, while Raman spectroscopy proved effective, integrating it with other analytical methods could further enhance the robustness of the characterization process. Future research should aim to increase the sample size and explore additional machine learning algorithms to improve the accuracy and reliability of spectral analysis.

REFERENCES

Andre de Villiers, Alberts Phillipus, & Nieuwoudt Helene. (2012). Analytical techniques for wine analysis: An African perspective - a review. South African Journal of Enology and Viticulture, 33(1), 1–14. https://www.researchgate.net/publication/225061828_Analytical_techniques_for_wine_analysis_An_African_perspective_a_review

Andrea, O., & Filippo, F. (2021). A Comprehensive Review on Raman Spectroscopy Application. Chemosensors. https://www.mdpi.com/2227-9040/9/9/262

Anton Paar. (2024). Basics of Raman spectroscopy | Anton Paar Wiki — wiki.anton-paar.com. https://wiki.anton-paar.com/en/basics-of-raman-spectroscopy/

Balekundri, A., & Mannur, V. (2020). Quality control of the traditional herbs and herbal products: a review. Future Journal of Pharmaceutical Sciences 2020 6:1, 6(1), 1–9. https://doi.org/10.1186/S43094-020-00091-5

Barth, A. (2007). Introduction to Spectroscopy.

BRAVO Handheld Raman Spectrometer | Bruker. (2022). https://www.bruker.com/en/products-and-solutions/infrared-and-raman/raman-spectrometers/bravo-handheld-raman-spectrometer.html

Brownlee, J. (2020a, August 26). LOOCV for Evaluating Machine Learning Algorithms - MachineLearningMastery.com. https://machinelearningmastery.com/loocv-for-evaluating-machine-learning-algorithms/

Brownlee, J. (2020b, August 26). LOOCV for Evaluating Machine Learning Algorithms - MachineLearningMastery.com. https://machinelearningmastery.com/loocv-for-evaluating-machine-learning-algorithms/

Bumbrah, G. S., & Sharma, R. M. (2016). Raman spectroscopy – Basic principle, instrumentation and selected applications for the characterization of drugs of abuse. In Egyptian Journal of Forensic Sciences (Vol. 6, Issue 3, pp. 209–215). Egyptian Forensic Medicine Authority. https://doi.org/10.1016/j.ejfs.2015.06.001

Cheriyan Ashok, P. (2011). INTEGRATION METHODS FOR ENHANCED TRAPPING AND SPECTROSCOPY IN OPTOFLUIDICS. http://research-repository.st-andrews.ac.uk/

Chi, H. W., Hu, S. W., & Lin, D. Z. (2023). Development of an automated Raman system and use of principal component analysis to classify real and counterfeit liquors. RSC Advances, 13(47), 33288. https://doi.org/10.1039/D3RA06057H

Cleveland, D., Carlson, M., Hudspeth, E. D., Quattrochi, L. E., Batchler, K. L., Balram, S. A., Hong, S., & Michel, R. G. (2007). Raman spectroscopy for the undergraduate teaching laboratory: Quantification of ethanol concentration in consumer alcoholic beverages and qualitative identification of marine diesels using a miniature raman spectrometer. Spectroscopy Letters, 40(6), 903–924. https://doi.org/10.1080/00387010701525638

D.J. Gardiner, P. R. G. H. J. B. (1989). Practical Raman Spectroscopy. Springer-Verlag.

Dr. Shinde, V. (2023, December 5). Quantitative and Qualitative Analysis in Pharmaceuticals - Veeprho. https://veeprho.com/quantitative-and-qualitative-analysis-in-pharmaceuticals/

Ghana: Ghana Alcoholic Beverages Report | USDA Foreign Agricultural Service. (2024, June 18). https://fas.usda.gov/data/ghana-ghana-alcoholic-beverages-report

Global Bitters Market – Market Reports World. (2021, January 12). https://www.marketreportsworld.com/global-bitters-market-17138387

Granite. (2021). What is Raman Spectroscopy? | Raman Spectroscopy Principle. Edinburgh Instruments. https://www.edinst.com/de/blog/what-is-raman-spectroscopy/

Hassing, S., Jernshøj, K. D., & Christensen, L. S. (2012). Raman spectroscopy: a non-destructive and on-site tool for control of food quality. Food Qual, 27, 47–73.

I.T. Jolliffe. (2002). Principal Component Analysis, Second Edition (2nd ed., p. 32). Springer.

John Wiley & Sons, Inc. (2024a). Ethanol Raman Spectrum. John Wiley & Sons, Inc. https://spectrabase.com/spectrum/64WcXAR1x1s

John Wiley & Sons, Inc. (2024b). Methanol Raman Spectrum. John Wiley & Sons, Inc. John Wiley \& Sons, Inc.

Lapierre, C., Erlandson, L. W., Stoneroad, R., Rhiner, A., Gosnell, R., Barber, J., & Pham, L. (2024). Substances of health concern in home-distilled and commercial alcohols from Texas. Heliyon, 10(11), e32317. https://doi.org/10.1016/J.HELIYON.2024.E32317

Lestyo, W., Yuni, R., Nuri, N., & Hilmia, L. (2016, July). The leave-one-out cross-validation (LOOCV) of PLS (NIR model). | Download Scientific Diagram. https://www.researchgate.net/figure/The-leave-one-out-cross-validation-LOOCV-of-PLS-NIR-model_fig3_305670586

Luo, R., Popp, J., & Bocklitz, T. (2022). Deep Learning for Raman Spectroscopy: A Review. Analytica 2022, Vol. 3, Pages 287-301, 3(3), 287–301. https://doi.org/10.3390/ANALYTICA3030020

Lussier, F., Thibault, V., Charron, B., Wallace, G. Q., & Masson, J. F. (2020). Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering. TrAC Trends in Analytical Chemistry, 124, 115796. https://doi.org/10.1016/J.TRAC.2019.115796

Marisia, A. F., Igor, K. L., Shelby, R. K., Ewelina, M., Yasmine, A., Lenka, H., & Justin, B. (2019). Forensics: Evidence Examination via Raman Spectroscopy | National Institute of Justice. https://nij.ojp.gov/library/publications/forensics-evidence-examination-raman-spectroscopy

McHugh, M. (2021, January 21). Compositional Analysis of Wine and Alcoholic Beverages | MarqMetrix Process Raman Applications. https://marqmetrix.com/raman-compositional-analysis-of-wine-and-alcoholic-beverages/

McReynolds, N. (2017). ADVANCED MULTIMODAL METHODS IN BIOMEDICINE: RAMAN SPECTROSCOPY AND DIGITAL HOLOGRAPHIC MICROSCOPY [PhD, University of St Andrews]. http://research-repository.st-andrews.ac.uk/

Mobili, P., Londero, A., De Antoni, G., Gómez-Zavaglia, A., Araujo-Andrade, C., Avila-Donoso, H. ´, Ivanov-Tzonchev, R., & Moreno, I. (2010). Multivariate analysis of Raman spectra

applied to microbiology: Discrimination of microorganisms at the species level. In REVISTA MEXICANA DE F ´ ISICA (Vol. 56, Issue 5). http://www.redalyc.org/articulo.oa?id=57019827004

MyJoyOnline. (2016). Proliferation of bitters products and the risk to consumers. https://www.myjoyonline.com/proliferation-of-bitters-products-and-the-risk-to-consumers/

Pandiscia, L. (2022, August 25). Gemstone Identification Using Raman Spectroscopy - JASCO. https://jascoinc.com/applications/gemstone-identification-using-raman-spectroscopy/

Qi, Y., Hu, D., Jiang, Y., Wu, Z., Zheng, M., Chen, E. X., Liang, Y., Sadi, M. A., Zhang, K., & Chen, Y. P. (2023). Recent Progresses in Machine Learning Assisted Raman Spectroscopy. Advanced Optical Materials, 11(14), 2203104. https://doi.org/10.1002/ADOM.202203104

ScienceDoze. (2022, October 2). Raman Spectroscopy: Principle, Instrumentation and Applications. https://www.sciencedoze.com/2022/10/raman-spectroscopy-principle.html

Sector industry study Ghana food and drink industry by strategy and research department. (2019).

Skoog. (2020, July 14). Principles of Instrumentation - Raman Spectroscopy. Chemistry LibreTexts. https://chem.libretexts.org/Under_Construction/Purgatory/Principles_of_Instrumental_Analysis_(Skoog_et_al.)_-_Under_Construction/18%3A_Raman_Spectroscopy

Smith, E., & Dent, G. (2005). Modern Raman Spectroscopy-A Practical Approach. John Wiley & Sons Ltd.

Speight, J. G. (2017). Chapter Five - Sources and Types of Inorganic Pollutants. In J. G. Speight (Ed.), Environmental Inorganic Chemistry for Engineers (pp. 231–282). Butterworth-Heinemann. https://doi.org/https://doi.org/10.1016/B978-0-12-849891-0.00005-9

Tabago, M. K. A. G., Calingacion, M. N., & Garcia, J. (2021). Recent advances in NMR-based metabolomics of alcoholic beverages. Food Chemistry: Molecular Sciences, 2, 100009. https://doi.org/10.1016/J.FOCHMS.2020.100009

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Wu, J. Bin, Lin, M. L., Cong, X., Liu, H. N., & Tan, P. H. (2018). Raman spectroscopy of graphene-based materials and its applications in related devices. Chemical Society Reviews, 47(5), 1822–1873. https://doi.org/10.1039/C6CS00915H

Zhang, S., Qi, Y., Tan, S. P. H., Bi, R., & Olivo, M. (2023). Molecular Fingerprint Detection Using Raman and Infrared Spectroscopy Technologies for Cancer Detection: A Progress Review. Biosensors 2023, Vol. 13, Page 557, 13(5), 557. https://doi.org/10.3390/BIOS13050557